



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Maestría y Doctorado en Ciencias Bioquímicas

El paisaje conformacional de las proteínas:
estudio de barriles TIM naturales y diseñados *de novo*

TESIS

QUE PARA OPTAR POR EL GRADO DE:

Doctor en Ciencias

PRESENTA:

Sergio Romero Romero

TUTOR PRINCIPAL:

Dr. Daniel Alejandro Fernández Velasco
[Facultad de Medicina](#)

MIEMBROS DEL COMITÉ TUTOR:

Dr. Enrique García Hernández
[Instituto de Química](#)

Dr. Rogelio Rodríguez Sotres
[Facultad de Química](#)

Ciudad Universitaria, Ciudad de México. Diciembre 2017



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Para todos los que abandonaron el trayecto con el fin de integrarse a otros caminos: el de los sueños colectivos y la resistencia de conciencias contemporáneas. Para ustedes, nuestro corazón abajo y a la izquierda.

Gracias todos los mexicanos que financian la Ciencia y Tecnología en el país.

Para todas aquellas mujeres quienes han sido violentadas de una u otra forma. No más. Ni una más.

Dedicada a todas las víctimas de los sismos del 11 y 19 de septiembre de 2017 en México.

In memoriam: Dr. Armando Gómez Puyou.

Niquitoa

Niquitoa ni Nezahualcoyotl:
¿Cuix oc nelli nemohua in talticpac?
An nochipa talticpac:
zan achica ye nican.

Tel ca chalchihuitl no xamani,
no teocuitlatl in tlalpani,
no quetzalli poztequi.

An nochipa talticpac:
zan achica ye nican.

Yo lo pregunto

Yo, Nezahualcóyotl, lo pregunto:
¿Acaso deveras se vive con raíz en la tierra?
No para siempre en la tierra:
sólo un poco aquí.

Aunque sea de jade se quiebra,
aunque sea de oro se rompe,
aunque sea plumaje de quetzal se desgarrá.

No para siempre en la tierra:
sólo un poco aquí.

You will never understand
How it feels to live your life
With no meaning or control
And with nowhere left to go
You are amazed that they exist
And they burn so bright
Whilst you can only wonder why

Rent a flat above a shop
Cut your hair and get a job
Smoke some fags and play some pool
Pretend you never went to school
But still you'll never get it right
Cause when you're laid in bed at night
Watching roaches climb the wall
If you called your dad he could stop it all
Yeah!

Never live like common people
Never do what common people do
Never fail like common people
Never watch your life slide out of view
And we dance, and drink, and screw
Because there's nothing else to do...

Common People

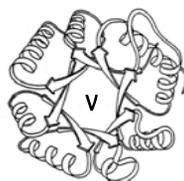
Pulp. 1995

RECONOCIMIENTOS INSTITUCIONALES

Este trabajo se realizó bajo la sobresaliente dirección del Dr. Daniel Alejandro Fernández Velasco en la anarcocomuna del Laboratorio de Físicoquímica e Ingeniería de Proteínas (LFQIP) perteneciente a la Facultad de Medicina, UNAM.

Durante el desarrollo de la investigación se contó con los siguientes apoyos financieros:

- i.* Beca doctoral del Consejo Nacional de Ciencia y Tecnología (CONACyT) con número de CVU/Becario: 387653/253724.
- ii.* Beca Mixta 2016-2017 de movilidad en el extranjero por parte del CONACyT para participar en una estancia de investigación en la Universidad de Bayreuth, Alemania.
- iii.* Beca de ayudante de investigador del CONACyT con expedientes de ayudante y de investigador: 14401/4312.
- iv.* Beca de obtención del grado doctoral por parte del Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT), UNAM-DGAPA, con clave del proyecto: IN220516.
- v.* Beca para la asistencia al curso: *Cristalografía en Biología Estructural 2017*. Financiamiento otorgado por el Centro Latinoamericano de Formación Interdisciplinaria (CELFI) del Ministerio de Ciencia, Tecnología e Innovación Productiva de Argentina.
- vi.* Beca para la asistencia al congreso: *4th International Workshop: Frontiers in Protein Folding, Evolution and Function*. Financiamiento otorgado por parte del CONACyT.

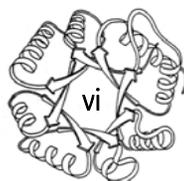


- vii. Beca para la asistencia al curso: *CCP4-IPMontevideo Macromolecular Crystallography School 2015 "From data processing to structure refinement and beyond."* Financiamiento otorgado por parte del Collaborative Computational Project Number 4 in Protein Crystallography (CCP4) del Reino Unido y el Instituto Pasteur de Montevideo, Uruguay.
- viii. Beca para la asistencia al curso: *Rapid Data 2014 Collection and Structure Solving at the NSLS: A Practical Course in Macromolecular X-Ray Diffraction Measurement.* Financiamiento otorgado por parte de la International Union of Crystallography y el Brookhaven National Laboratory de New York, EE. UU.
- ix. Financiamiento para la asistencia a congresos nacionales e internacionales, cursos y una estancia de investigación por parte del Programa de Apoyo a los Estudios del Posgrado (PAEP) y del Programa de Maestría y Doctorado en Ciencias Bioquímicas, UNAM.
- x. La Red de Estructura, Función y Evolución de Proteínas (REFEP) financió parcialmente la asistencia a una estancia de investigación en la Universidad de Bayreuth, Alemania.

El trabajo de investigación fue financiado por los proyectos PAPIIT IN220516, CONACyT 254514 y la Facultad de Medicina de la UNAM.

La tesis fue defendida mediante el examen profesional presentado el día 14 de diciembre en el auditorio Antonio Peña Díaz del Instituto de Fisiología Celular, CU-UNAM. Se agradece en demasía el apoyo para la revisión del manuscrito y la asistencia del jurado de examen conformado por:

- i. Presidente: Dra. Adela Rodríguez Romero, Instituto de Química, UNAM.
- ii. Secretario: Dr. Miguel Antonio Costas Basín, Facultad de Química, UNAM.
- iii. Vocal: Dra. Lilian González Segura, Facultad de Química, UNAM.
- iv. Vocal: Dra. Leticia Haydeé Ramírez Silva, Facultad de Medicina, UNAM.
- v. Vocal: Dr. Carlos Amero Tello, Centro de Investigaciones Químicas, UAEM.



RECONOCIMIENTOS ACADÉMICOS

Se reconoce y agradece el trabajo del Comité Tutorial que asesoró el desarrollo de esta tesis, el cual estuvo conformado por:

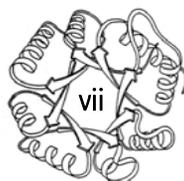
- i.* Dr. Enrique García Hernández: Instituto de Química, UNAM.
- ii.* Dr. Rogelio Rodríguez Sotres: Facultad de Química, UNAM.
- iii.* Dr. Daniel Alejandro Fernández Velasco: Facultad de Medicina, UNAM.

El trabajo se desarrolló con el amplio e invaluable apoyo técnico de los miembros del laboratorio: M. en C. Laura Ileana Álvarez Añorve y Dra. Ma. Isabel Velázquez López.

La investigación fue enriquecida día a día por los investigadores del LFQIP: Dra. Georgina Garza Ramos, Dr. Ismael Bustos Jaimes, Dr. Alejandro Sosa Peinado, Dr. Mario Calcagno Montans y Dr. D. Alejandro Fernández Velasco.

Los experimentos de Calorimetría Diferencial de Barrido y Calorimetría de Titulación Isotérmica se realizaron gracias al destacado apoyo intelectual y académico del Dr. Miguel Costas Basín y el técnico académico M. en C. Jessica Díaz Salazar del Laboratorio de Biofísicoquímica de la Facultad de Química, UNAM.

La resolución de la estructura tridimensional de las proteínas fue posible gracias al trabajo arduo, notorio y constante que se realizó en colaboración con la Dra. Adela Rodríguez Romero y el técnico académico M. en C. Geogina Espinosa Pérez del Laboratorio Nacional de Estructuras de Macromoléculas sede Instituto de Química, UNAM.



Se agradece también la disponibilidad, apertura y discusión científica por parte de la Dra. Birte Höcker de la Universidad de Bayreuth, Alemania, donde se realizó una estancia de investigación para completar parte del proyecto de los barriles TIM diseñados *de novo*. A la par, a todo el apoyo académico y personal brindado por Sooruban Shanmugaratnam, Kaspar Feldmeier, Saacnicteh Toledo, Francisco Lobos, Bruce Lichtenstein, Horst Lechner y Noelia Ferruz, todos miembros de ese grupo de trabajo.

De igual manera, se reconoce en demasía el apoyo brindado por Leticia García en la realización de todos los trámites y solución de problemas durante la maestría y el doctorado; te extrañamos, Lety. Asimismo, gracias a Julio Palacios y Lilia de la Rocha por el apoyo administrativo efectuado en la coordinación del posgrado.

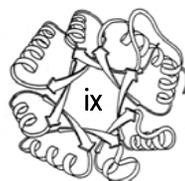


AGRADECIMIENTOS PERSONALES

Al intentar escribir esta sección me invaden dos sentimientos principales: sinceridad e incertidumbre. Sinceridad porque las palabras acá depositadas están plasmadas como naturalmente las pienso y siento (claro, con un poco de edición de estilo para no pecar de barbaján). Incertidumbre porque en este viaje lleno de sentimentalismo y cursilería tengo temor de olvidar colocar a alguien especial, pero ustedes bien saben que mi mente está lejos de ser la más apta para retener información valiosa; de verdad, disculpas anticipadas si con alguien sucede la omisión.

Pasé años haciendo una la tesis, aprendiendo de bioquímica, de biología molecular y de muchas otras cosas, pero creo que nunca aprendí a cómo agradecer lejos de usar esa palabra tan común que es "gracias", la cual engloba todo pero a la vez siempre queda vacía. Javier Marías en uno de sus tantos divagues mentales menciona que cualquiera puede contar una anécdota y, por el mero hecho de contarla, ya la está deformando y tergiversando; la lengua no puede reproducir los hechos ni por lo tanto debería intentarlo. De esta manera, ocuparé las palabras aunque éstas no reflejen completamente lo que siento por ustedes; la pregunta ahora es: ¿qué palabras usar?

Recuerdo muy bien que hace tiempo leí de un tal Adorno que la música capta la esencia misma del mundo, el Espíritu, la Infinitud. Por esto, el acercamiento a ella sólo es posible para aquel que se aproxima con el corazón, su corazón. Bien pues, quiero darles las gracias con algo que quedará ahí para siempre, incluso después de que esta tesis quede en el olvido y se generen numerosos textos que superen el contenido presente en estas hojas. Para ustedes, mis sinceros y amorosos agradecimientos mediante estas canciones que los representan, nos representan.



Madre y padre. Ojalá algún día llegue a ser la mitad de bondadoso y buen ser humano como lo son ustedes; de verdad, si lo bueno existe seguro es imagen fiel de ustedes dos. Ambos son esas gaviotas que José Luis Perales menciona y que siempre están ahí en el aire, viéndonos y cuidándonos: *Y se marchó y a su barco le llamó libertad y en el cielo descubrió gaviotas y pintó estelas en el mar...* Los amo, sin más; son el tesoro más grande que tengo en la vida.

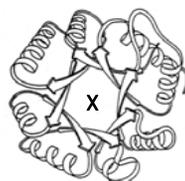
Paola, Omar y Max. En este camino siempre han sido y serán mi eco, los mismos de Pink Floyd: *And no one showed us to the land, and no one knows the where's or why's, but something stirs and something tries starts to climb towards the light.* Sé que ustedes, niños tontos, están en todas las dimensiones que nos adoramos.

Demonios: **Paola Fernanda, Luis Alfonso, Diego y Marián.** *When I stare into space I'm looking for you and I can see you at times. Sí, I'll be around.* No dejen de ver el mundo con la alegría y simpleza que lo hacen porque esa misma es la que admiro de ustedes.

Fa: *Te quiero tanto, tanto, tanto, tanto, tanto; para mí no hay nada igual, no lo hay.* Cayendo en la cursilería, esta canción siempre está cuando te pienso y me recuerda a cómo crecimos juntos en la vida y seguimos haciéndolo. Gracias por todo, primis.

Gustavo. Verga, cabrón, no sé ni qué decirte para agradecer todo lo que me has dado y lo que me has enseñado. Me quedé pensando mucho en qué rola nos identifica e iba a poner algo de Maluma pero qué turbo oso que eso quedara en mi tesis. ¿Pero sabes qué? *All the small things true care, truth brings. I'll take one lift; your ride, best trip...* ¿O no? Hermanito, te adoro.

Jess, Lina y Yer. Nada nos describe mejor que nuestra rola de The Bravery: *The memories in my head are just as real the time we spent. You always be close to me my friend... This is not the end.* Y no lo será, nunca. Los amo.



Mariana Pacheco. No puedo creer la fortuna que tuve al encontrarte en esta vida y al tener la oportunidad de seguir riendo y abrazándonos en este largo camino. Estoy seguro de que: *Yo no creo en la iglesia pero creo en tu mirada. Yo soy la vida que ya tengo, tú eres la vida que me falta... Y dame la mano y vamos a darle la vuelta al mundo, darle la vuelta al mundo.* Gracias por tanto, mi güera hermosa.

Marianita. ¿Sabes cuál canción muda imágenes y recuerdos de ti? Sí, esta misma: *Por alto que esté el cielo en el mundo, por hondo que sea el mar profundo, no habrá una barrera en el mundo que mi amor profundo no rompa por ti.* Cierra los ojos y siente cómo te abrazo, tulli.

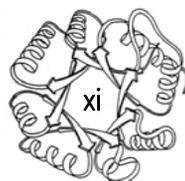
Zai. Desde aquella tarde/noche en el Guajira la vida me dio un regalo el cual cuidaré con todo mi ser: tú. Recuerda que *I don't care if Monday's black, Tuesday, Wednesday, heart attack. Thursday, never looking back. It's Friday, I'm in love...* about you. Recuerda que los mejores momentos empiezan con una mala decisión.

Isra y Fer. Sabemos lo que significamos para los otros, junto con Mariana. Un gusto tenerlos a mi lado, colegas. Sepan siempre que: *Oh well I don't mind, if you don't mind, 'cuz I don't shine if you don't shine, before you jump, tell me what you find when you read my mind...*

This is a story about a man named Vargas. Marica, tú bien sabes qué pedo en este viaje juntos, en el que empezamos y no terminaremos hasta que Puebla sea la capital del país.

"Doñitas": Roselia, Zoe, Mariana, Vanessa, Raúl, Mario, Rubinsky y Yeni. Sé que no merezco tenerlos como amigos, son el amor hecho banda. Sepan que, entre las pocas cosas que tengo para ofrecer, una de ellas es mi corazón y oído para ustedes, siempre. Ya lo dijo Mecano: *Pero la fuerza del destino nos hizo repetir, que si el invierno viene frío quiero estar junto a ~~ti~~ ustedes.*

Rafa. ¿Recuerdas?: *I'm there when your fingers snap, it's not where we left it at. I can see the love we trapped coming back.* Gracias por ser y por seguir siendo.



Ale bonita. Seré breve: *You are the dancing queen, young and sweet only seventeen... You can dance, you can jive having the time of your life. Oooh, see that girl, watch that scene digging the dancing queen!* Te pinches amo.

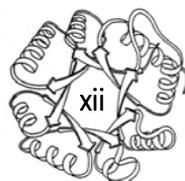
Viri, Rafa, Ale, Gerardo, Omar, Miri, Fer y Patsy. Estoy totalmente agradecidos con ustedes por dejarme compartir tantos momentos juntos, tantas risas, tanto tequila, tanta luz de la verdad. Creo que tenemos muchas canciones pero una especial dice más o menos así: *Perfume de gardenias tiene tu boca, bellísimos destellos de luz en tu mirar. Tu risa es una rima de alegres notas, se mueven tus cabellos cual ondas en el mar.* Gracias por ser lo que son y hacerme sentir tanto.

Can you hear me, Major-Tom-Chris? Can you hear me, Major-Tom-Chris? Can you... Here am I floating 'round my tin can far above the Moon. A quien diga que no se puede querer y agradecer a alguien a distancia, vil mentira, ¿no crees, **Christian?**

Marcos. *We know a place where no planes go. We know a place no space ships go.* Para que el futuro nos alcance de las maneras más sutiles e inesperadas como hasta ahora... *where we know.*

Mariana, Akram, Adhemar, Robert, Richie, Soto, Jorge y Adriana. Este viaje en el posgrado no sería lo mismo sin ustedes. Cada uno es parte de mi familia elegida y esta cumbia lo sabe: *Cómo no acordarme de ti, de qué manera olvidarte, si todo me recuerda a ti, si en todas partes estás tú...*

Mariana Schulte-Sasse. La esperanza en el mundo siempre la recobro cuando te pienso. Me haces tanta falta pero a la vez me siento tan orgulloso de ti, mujer hermosa. Recuerda: *We're burning down the highway skyline on the back of a hurricane that started turning when you were young, when you were young. And sometimes you close your eyes and see the place where you used to live... when you were young.*



Susy, Richie, Lau y Bryan. *Todavía me acuerdo de ti, todavía siento que estás junto a mí. Ni el tiempo, ni el espacio podrán borrar lo que me hiciste soñar.* Gracias por los días, las noches y la espontaneidad de disfrutar la vida y amarnos de esta forma; los adoro, hijos míos.

Chango feo, mi **Toño**. Iba a poner algo de la Chikybombom pero no quiero quemarte por acá. Por todo este tiempo, sólo puedo decirte: *Si hiciera un viaje a mis adentros y sobreviviera a los lamentos, pediría fuerzas para decir cuánto lo siento, si volviera en un viaje a mis adentros... si volviera, en todos estás tú, chango del mal.*

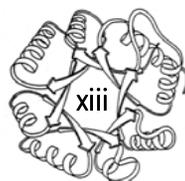
Banda socióloga. Siempre han sido un amortiguador ante el sistema que vivimos y sin duda el cristal con el que miran al mundo me ha ayudado a cuestionarme las formas en que nos desenvolvemos. Gracias por eso, por sentirme querido y a la vez vulnerable en este mundo. Para ustedes: *Viento, amárranos. Tiempo, detente muchos años...*

Francisco: *There is a light and it never goes out.*

Nicole. Qué rifado haberte conocido y seguir juntos; a un hemisferio de distancia pero juntos. Siempre... *Remember me when you're the one who's silver screened, remember me when you're the one you always dreamed. Remember me whenever noses start to bleed. Remember me, special needs...*

Ángeles: *We can live like Jack and Sally if we want, where you can always find me... ¿Sí?*

A todos mis alumnos y exalumnos biomédicos. Más que yo haberles aportado algo en sus vidas, ustedes han dejado una gran huella de conocimiento, motivación y gran cariño en mí. Sigán, sigán, porque el mundo necesita de su forma e ímpetu de aprehender la vida. Siempre: *Counting all different ideas drifting away. Past and present, they don't matter, now the future's sorted out...*



AleGutLand, Dana, Karla, Renán, Tania, Octavio, Alex, Andreaopsida, Julio y Areli. Verga, pensé mucho la canción que nos puede contener y, aunque a pocos les gusta Juan Gabriel (muy mal, eh), sólo él podía tener la respuesta: *Cómo te puedo pagar todo lo que haces por mí, todo lo feliz que soy, todo este grande amor. Solamente con mi vida, pues ten mi vida, te la doy, pero no me dejes nunca te lo pido por favor.* Arre, morros.

Geo. Porque siempre pensaré en ti cuando escuche: *Oye abre tus ojos, mira hacia arriba, disfruta las cosas buenas que tiene la vida... Un descanso en el camino, una botella de vino, un suspiro una mirada, una alegre carcajada...* Todas esas las compartí contigo, Geo, gracias por todo lo que me enseñaste, eres una persona pocamadre; sigamos bailando en esta vida.

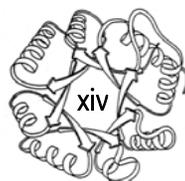
Doc Alejandro. Recuerdo que en el segundo semestre, en una de esas mañanas cafetaleras, divagamos sobre la vida y desde entonces esta pregunta es constante: *Is this the real life? Is this just fantasy? Caught in a landslide, no escape from reality.* ¿Usted qué cree? Gracias por tanto, jefe, es uno de mis ejemplos a seguir en esta vida tanto personal como académica.

Anarcocomuna LFQIP. No hay canción que mejor describa a los que estuvieron, a los que están y a los que estarán entre estas paredes: *It really, really, really could happen. Yes, it really, really, really could happen. When the days they seem to fall through you, well, just let them go...* Rífense siempre porque están en el mejor laboratorio de proteínas, a que sí.

No es la conciencia de los hombres lo que determina su ser, sino, por el contrario, es su existencia social lo que determina su conciencia.

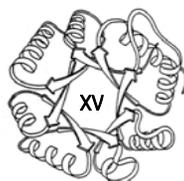
Karl Marx, La ideología alemana: 1846.

Ciudad de México, diciembre de 2017

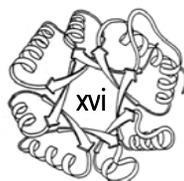


ÍNDICE GENERAL

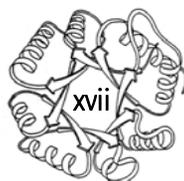
RECONOCIMIENTOS INSTITUCIONALES.....	v
RECONOCIMIENTOS ACADÉMICOS	vii
AGRADECIMIENTOS PERSONALES	ix
ÍNDICE GENERAL	xv
ÍNDICE DE FIGURAS Y TABLAS	xxi
LISTA DE ABREVIATURAS.....	xxvi
RESUMEN	xxxii
ABSTRACT	xxxiii
INTRODUCCIÓN	1
ANTECEDENTES GENERALES	5
Plegamiento de las proteínas	5
<i>El problema del plegamiento de proteínas y los embudos de energía.....</i>	5
<i>Agentes perturbantes para el estudio del plegamiento de proteínas.....</i>	7
<i>El plegamiento y la estabilidad de las proteínas:</i> <i>dos propiedades fisicoquímicas íntimamente relacionadas.....</i>	9
<i>El plegamiento en proteínas homólogas.....</i>	15
Ingeniería y diseño de proteínas	19
Barriles (β/α) ₈ o barriles TIM	25



ANTECEDENTES ESPECÍFICOS	28
La Triosafosfato Isomerasa como modelo de estudio de un barril TIM natural.....	28
<i>Triosafosfato Isomerasa: estructura y función</i>	28
<i>Triosafosfato Isomerasa: plegamiento y estabilidad</i>	38
<i>Ingeniería de proteínas aplicada en la TIM</i>	41
<i>Triosafosfato Isomerasa: estudios realizados en el dominio Bacteria</i>	44
<i>Triosafosfato Isomerasa: estudio previo realizado durante la maestría</i>	47
La proteína sTIM11 como modelo de estudio de un barril TIM diseñado <i>de novo</i>	52
JUSTIFICACIÓN	57
OBJETIVOS	58
Objetivo general.....	58
Objetivos particulares: barriles TIM naturales.....	58
Objetivos particulares: barriles TIM diseñados <i>de novo</i>	59
ESQUEMA GENERAL DE TRABAJO	60
ESTRATEGIA METODOLÓGICA	62
Reactivos y soluciones utilizadas.....	62
Parte 1: Barriles TIM naturales	62
Sobreexpresión de la proteína.....	62
Mediciones espectroscópicas.....	64
<i>Dicroísmo Circular</i>	64
<i>Fluorescencia Intrínseca</i>	65



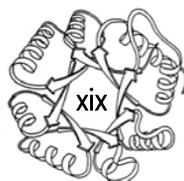
Resolución de la estructura tridimensional	65
Determinación del estado de oligomerización	67
Determinación de la actividad enzimática y los parámetros cinéticos	68
Determinación de la afinidad al PGH	69
Disociación por dilución	70
Desplegamiento proteico mediante urea	71
Desplegamiento proteico mediante temperatura.....	71
<i>Dicroísmo Circular</i>	71
<i>Calorimetría Diferencial de Barrido</i>	72
Análisis de secuencia y estructurales	73
Parte 2: Barriles TIM diseñados <i>de novo</i>	75
Diseño de las proteínas	75
Sobreexpresión de la proteína	79
Mediciones espectroscópicas	81
<i>Dicroísmo Circular</i>	81
Desplegamiento proteico mediante hidrocloreuro de guanidinio	82
Desplegamiento proteico mediante temperatura.....	82
<i>Dicroísmo Circular</i>	82
<i>Calorimetría Diferencial de Barrido</i>	83
Resolución de la estructura tridimensional	83
Parte 3: Análisis y ajuste de los datos	84



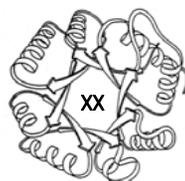
RESULTADOS Y DISCUSIÓN	92
Parte 1: El paisaje conformacional de barriles TIM naturales	92
Sobreexpresión de las BacTIMs.....	92
Purificación de las BacTIMs.....	95
Nivel estructural secundario y terciario: propiedades espectroscópicas de las BacTIMs.....	97
Nivel estructural terciario: estructura tridimensional de las BacTIMs	100
Nivel estructural cuaternario: estado oligomérico de las BacTIMs	104
Estudio de la función de la TIM: parámetros cinéticos de las BacTIM por un sustrato natural ..	106
Estudio de la función de la TIM: afinidad de las BacTIMs a un inhibidor	108
El desplegamiento de las BacTIMs: desplegamiento mediado por dilución	112
El desplegamiento de las BacTIMs: desplegamiento mediado por urea	116
El desplegamiento de las BacTIMs: desplegamiento mediado por temperatura	121
<i>BacTIMs con desplegamiento térmico irreversible: NpTIM y GoTIM.....</i>	121
<i>BacTIMs con desplegamiento térmico reversible: DrTIM, CpTIM y ScTIM.....</i>	126
Descripción fisicoquímica de la estabilidad de las RevTIMs: la curva de estabilidad	134
En búsqueda de una posible relación entre las propiedades moleculares y la reversibilidad en el plegamiento.....	136
<i>Propiedades relacionadas con la secuencia y las relaciones filogenéticas.....</i>	137
<i>Propiedades relacionadas con los estados estables.....</i>	143
<i>Propiedades fisicoquímicas de la transición desplegamiento/plegamiento.....</i>	148



Búsqueda de nuevas secuencias para estudiar	152
Discusión final sobre el plegamiento de los barriles TIM naturales	154
Parte 2: El paisaje conformacional de barriles TIM diseñados <i>de novo</i>	163
En búsqueda de incrementar la estabilidad de sTIM11: diseño de nuevos barriles TIM.....	163
Sobreexpresión de las ReTIMs	168
Purificación de las ReTIMs	170
Propiedades espectroscópicas de las ReTIMs: espectro de dicroísmo circular.....	172
Propiedades fisicoquímicas de los diseños del núcleo interno: ReTIM1-4	173
Propiedades fisicoquímicas de los diseños del núcleo periférico A: ReTIM5-7.....	176
Propiedades fisicoquímicas de los diseños del núcleo periférico B: ReTIM8-10.....	178
Propiedades fisicoquímicas de los diseños con mutaciones puntuales: ReTIM11-16	180
Propiedades fisicoquímicas de los diseños quimeras: ReTIM17-28.....	183
Descripción fisicoquímica de la estabilidad de las ReTIMs: la curva de estabilidad.....	186
Estructura tridimensional de las ReTIMs.....	189
Discusión final sobre el plegamiento de los barriles TIM diseñados de novo.....	193
Parte 3: Lo que hemos aprendido sobre el paisaje conformacional de los barriles TIM. Una comparación entre barriles TIM naturales y diseñados <i>de novo</i>	Una 197
CONCLUSIONES	203
PERSPECTIVAS	206
BIBLIOGRAFÍA	207

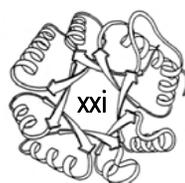


APÉNDICE A.....	226
Origen y evolución del dominio Bacteria.....	226
Filogenia y taxonomía del dominio Bacteria	229
APÉNDICE B.....	233
APÉNDICE C.....	242
APÉNDICE D.....	249
APÉNDICE E.....	252

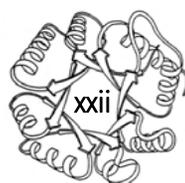


ÍNDICE DE FIGURAS Y TABLAS

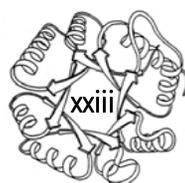
Figura	Tabla	Título	Página
1		Modelo del embudo para el plegamiento de una proteína.	3
2		Curvas de estabilidad para dos proteínas.	8
3		El panorama actual del diseño de proteínas.	20
4		Topología general de un barril (β/α) ₈ .	22
5		Reacción catalizada por la Triosafosfato Isomerasa.	25
	1	Parámetros cinéticos para las TIMs silvestres	26
6		Estructura tridimensional de la TIM.	28
	2	Especies para las cuales se ha determinado la estructura cristalográfica de la TIM.	30
7		Triada catalítica de la Triosafosfato Isomerasa.	31
8		Mecanismo catalítico de la Triosafosfato Isomerasa.	33
9		Conformaciones abierta y cerrada de la TIM.	33
10		Cambios energéticos en cada paso de la reacción catalizada por la Triosafosfato Isomerasa.	34
	3	Vías de plegamiento reportadas para la TIM de diferentes especies.	35
	4	Estudios reportados de la TIM en el dominio Bacteria.	42
	5	Propiedades fisicoquímicas de las BacTIMs.	44
11		Alineamiento de las secuencias de las BacTIMs.	45
12		Representaciones filogenéticas de las BacTIMs.	46
13		La proteína sTIM11.	51
14		Caracterización estructural y fisicoquímica de sTIM11.	52
15		Determinación del estado oligomérico de las proteínas.	63
16		Reacción acoplada en dirección G3P a DHAP.	64
	6	Archivos PDB diméricos utilizados para la comparación estructural entre RevTIMs e IrrevTIMs.	70
17		Algoritmo utilizado para el diseño de los barriles TIM.	73



18	Gel de poliacrilamida SDS-PAGE al 12% con las pruebas de sobreexpresión de las BacTIMs.	88
19	Gel de poliacrilamida SDS-PAGE al 12% con las pruebas de sobreexpresión para <i>AvTIM</i> .	89
20	Cromatogramas de la purificación de <i>DvTIM</i> .	92
21	Gel de poliacrilamida SDS-PAGE al 12% con los pasos de la purificación de <i>DvTIM</i> .	92
7	Tabla de purificación para <i>DvTIM</i> .	93
22	Propiedades espectroscópicas de las BacTIMs.	95
8	Contenido de estructura secundaria para las BacTIMs.	96
9	Condiciones exitosas de cristalización para las BacTIMs.	97
10	Estadísticas tras la colecta, integración, reducción, escalamiento y afinamiento de los datos para las BacTIMs.	98
23	Estructuras tridimensionales de las BacTIMs.	99
24	La estructura tridimensional de la TIM.	100
25	Cromatograma de elución para <i>DvTIM</i> .	101
11	Estado oligomérico de las TIMs bacterianas.	102
26	Cinética enzimática de las BacTIMs.	103
12	Parámetros cinéticos de las TIMs bacterianas.	103
27	ITC de unión de PGH para las BacTIMs.	106
13	Determinación de los parámetros termodinámicos de unión TIM-PGH usando ITC.	107
28	Firma termodinámica de las BacTIM para la unión de PGH.	107
29	Cálculo del ΔC_p mediante ITC para <i>DvTIM</i> .	108
30	Experimentos de estabilidad a la dilución de las BacTIMs.	109
31	Desplegamiento mediado por dilución para las BacTIMs.	111
14	Parámetros termodinámicos derivados de los experimentos de dilución para las BacTIMs.	112
32	Experimento de reversibilidad para el despliegamiento químico de las BacTIMs.	113
33	Curvas de despliegamiento químico de las BacTIMs.	115
15	Parámetros termodinámicos del despliegamiento químico de las BacTIMs.	116
34	Desplegamiento térmico de las IrrevTIMs.	118



35	Análisis termodinámico para el desplegamiento térmico de las IrrevTIMs.	120
16	Parámetros termodinámicos del desplegamiento térmico de las IrrevTIMs.	121
36	Desplegamiento térmico de las RevTIMs seguido por CD.	124
37	Desplegamiento térmico de las RevTIMs seguido por DSC.	125
38	Dependencia de la T_m con respecto a la velocidad de barrido para las BacTIMs.	126
17	Parámetros termodinámicos del desplegamiento térmico de las RevTIMs.	129
39	Curvas de estabilidad de barriles TIM naturales: $D\bar{T}IM$, $C\rho TIM$ y $S\bar{c}TIM$.	131
18	Valores de estabilidad a partir de las curvas construidas para $D\bar{T}IM$, $C\rho TIM$ y $S\bar{c}TIM$.	132
19	Matriz con el porcentaje de identidad y el promedio de los valores de RMSD para las IrrevTIMs y RevTIMs.	134
20	Composición aminoacídica de las RevTIMs e IrrevTIMs.	135
40	Alineamiento de las secuencias de las TIM usadas para la comparación.	136
21	Parámetros fisicoquímicos de las RevTIMs e IrrevTIMs.	138
41	Espectro de Dicroísmo Circular desnaturalizado.	140
22	Elementos de estructura secundaria e interacciones estabilizantes de las RevTIMs e IrrevTIMs.	141
23	Propiedades estructurales utilizadas para la comparación de las RevTIMs e IrrevTIMs.	141
42	Cavidades internas en la estructura tridimensional de las BacTIMs.	143
43	Comparación estructural entre RevTIMs e IrrevTIMs.	144
44	ΔG total y de asociación de las TIMs bacterianas y de otras TIMs eucariotes reportadas hasta el momento.	146
45	Gráfica 3D que muestra la relación de las 3 propiedades cuantitativas elegidas que se relacionan con la reversibilidad.	152
46	Desplegamiento térmico de $MmTIM$.	154
47	Desplegamiento térmico de $TtTIM$.	156
48	Regiones de sTIM11 que fueron rediseñadas para aumentar la estabilidad.	160
24	Quimeras de ReTIMs.	161
25	Mutaciones de las ReTIMs.	162



49	Gel de poliacrilamida SDS-PAGE al 12 % con las pruebas de sobreexpresión de las ReTIMs.	165
50	Cromatogramas de la purificación de ReTIM1.	167
51	Gel de poliacrilamida SDS-PAGE al 12 % con las ReTIMs purificadas.	167
52	Espectro nativo de dicroísmo circular para algunas ReTIMs.	168
53	Modelos estructurales de las ReTIM1-4.	169
54	Caracterización fisicoquímica de las ReTIM1-4.	171
26	Parámetros termodinámicos de las ReTIM1-4.	171
55	Modelos estructurales de las ReTIM5-7.	172
56	Caracterización fisicoquímica de la ReTIM5 y ReTIM6.	173
27	Parámetros termodinámicos de la ReTIM5 y ReTIM6.	173
57	Modelos estructurales de las ReTIM8-10.	174
58	Caracterización fisicoquímica de las ReTIM8-10.	175
28	Parámetros termodinámicos de las ReTIM8-10.	176
59	Modelos estructurales de las ReTIM11-16.	177
60	Caracterización fisicoquímica de la ReTIM13 y ReTIM15.	178
29	Parámetros termodinámicos de la ReTIM3 y ReTIM15.	178
61	Modelos estructurales de las ReTIM17-28.	179
62	Caracterización fisicoquímica de las ReTIM17-28.	181
30	Parámetros termodinámicos de las ReTIM17-28.	182
31	Parámetros termodinámicos para el desplegamiento de ReTIM17, ReTIM22, sTIM11 y sTIM11-SS.	184
63	Curvas de estabilidad de barriles TIM diseñados <i>de novo</i> : ReTIM17, ReTIM22, sTIM11 y sTIM11-SS.	185
32	Estadísticas tras la colecta, integración, reducción, escalamiento y afinamiento de los datos para ReTIM18.	187
64	Estructura tridimensional de ReTIM18.	188
65	Pruebas de cristalización de las ReTIMs.	189
66	Curvas de estabilidad comparadas entre barriles TIM naturales y diseñados <i>de novo</i> .	197

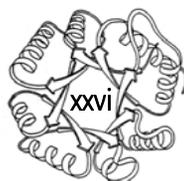


67	Cladograma que muestra la relación genética entre los distintos grupos bacterianos y su relación con los dominios Archaea y Eukarya.	222
68	Micrografía de la bacteria <i>Deinococcus radiodurans</i> .	229
69	Micrografía de un filamento de la bacteria <i>Nostoc punctiforme</i> .	231
70	Micrografía de la bacteria <i>Gemmata obscuriglobus</i> .	232
71	Micrografía de la bacteria <i>Azotobacter vinelandii</i> en estado quiste.	233
72	Micrografía de la bacteria <i>Clostridium perfringens</i> .	234
73	Micrografía de la bacteria <i>Streptomyces coelicolor</i> .	235
74	Micrografía de la bacteria <i>Moritella marina</i> .	236
75	Micrografía de la bacteria <i>Thermus thermophilus</i> .	237



LISTA DE ABREVIATURAS

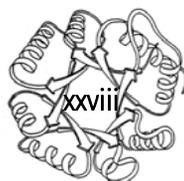
°C	Grados Celsius.
Å	Angstroms.
ADN	Ácido desoxirribonucleico.
αGDH	α-glicerol 3-fosfato deshidrogenasa.
AHA	α-amilasa de <i>Pseudoalteromonas haloplanktis</i> .
ARN	Ácido ribonucleico.
ASA	Área accesible al solvente.
<i>A</i> vTIM	TIM de <i>Azotobacter vinelandii</i> .
BacTIMs	TIMs bacterianas.
BCA	Ácido bicinconínico.
CME	Centro de masa espectral.
C_p	Capacidad calorífica.
C_{pN}	Capacidad calorífica del estado nativo.
C_{pD}	Capacidad calorífica del estado desplegado.
<i>C</i> pTIM	TIM de <i>Clostridium perfringens</i> .
D	Estado desplegado o desnaturalizado.
D'	Intermediario monomérico.
DC	Dicroísmo Circular.
DHAP	Dihidroxiacetona fosfato.
DMA	α-amilasa de <i>Drosophila melanogaster</i> .
DO	Densidad óptica.
<i>D</i> vTIM	TIM de <i>Deinococcus radiodurans</i> .
DSC	Calorimetría Diferencial de Barrido.
DTT	Ditiotreitol.
ΔASA	Cambio en el área accesible al solvente.
Δ C_p	Cambio de la capacidad calorífica a presión constante.
ΔDO	Cambio de la densidad óptica.
Δ <i>G</i>	Cambio de la energía libre de Gibbs.
Δ <i>G</i> _{asoc}	Cambio de la energía libre de Gibbs de la asociación.



ΔG_{dis}	Cambio de la energía libre de Gibbs de la disociación.
ΔG_{pleg}	Cambio de la energía libre de Gibbs del plegamiento.
ΔG_{tot}	Cambio de la energía libre de Gibbs total.
ΔH	Cambio de la entalpía a presión constante.
ΔH_b	Cambio de la entalpía de unión.
ΔH_{cal}	Cambio de entalpía calorimétrica.
ΔH_{TS}	Cambio de la entalpía a la T_s .
ΔH_{vH}	Cambio de entalpía de van't Hoff.
ΔS	Cambio de la entropía a presión constante.
ε	Coefficiente de absortividad molar.
E_A	Energía de activación.
EC	Número de la Comisión Internacional de Enzimas.
<i>E. coli</i>	<i>Escherichia coli</i> .
<i>EcTIM</i>	TIM de <i>Escherichia coli</i> .
EDTA	Ácido etilendiaminotetraacético.
<i>EtTIM</i>	TIM de <i>Entamoeba histolytica</i> .
E_T	Estado de transición.
F	Estado proteico final que es incapaz de replegarse.
f_D	Fracción desnaturalizada.
f_N	Fracción nativa.
FPLC	Cromatografía rápida líquida de proteínas.
g	Gramo.
G3P	Gliceraldehído-3-fosfato.
GdnHCl	Hidrocloreuro de guanidino.
<i>GgTIM</i>	TIM de <i>Gallus</i> .
<i>GtTIM</i>	TIM de <i>Giardia lamblia</i> .
<i>GoTIM</i>	TIM de <i>Gemmata obscuriglobus</i> .
<i>GsTIM</i>	TIM de <i>Geobacillus stearothermophilus</i> .
H	Hora.
<i>HpTIM</i>	TIM de <i>Helicobacter pylori</i> .
<i>HsTIM</i>	TIM de <i>Homo sapiens</i> .
IF	Fluorescencia Intrínseca.



IPTG	Isopropil- β -D-1-tiogalactopiranosido.
IrrrevTIMs	TIMs con desplegamiento térmico irreversible.
K	Kelvin.
K_b	Constante de unión.
k_{cat}	Constante catalítica.
K_{dis}	Constante de disociación.
K_{eq}	Constante de equilibrio.
K_m	Constante de Michaelis-Menten.
LB	Caldo de lisogenia.
λ_{max}	Longitud de onda de máxima emisión.
$LdTIM$	TIM de <i>Leishmania donovani</i> .
LLG	Probabilidad de ganancia log.
$LmTIM$	TIM de <i>Leishmania mexicana</i> .
M	Molar.
m	Pendiente de una transición.
m^{\ddagger}	m cinética en urea.
m_{eq}	m de urea al equilibrio.
mg	Miligramo.
$MjTIM$	TIM de <i>Methanocaldococcus jannaschii</i> .
mM	Milimolar.
$MmTIM$	TIM de <i>Moritella marina</i> .
monoTIM	TIM con estado de asociación monomérico.
$MtTIM$	TIM de <i>Mycobacterium tuberculosis</i> .
N	Estado nativo.
N_2	Dímero nativo.
N_2^*	Intermediario dimérico.
NAD^+	Dinucleótido de nicotinamida y adenina oxidado.
NADH	Dinucleótido de nicotinamida y adenina reducido.
nm	Nanómetros.
$NpTIM$	TIM de <i>Nostoc punctiforme</i> .
$OcTIM$	TIM de <i>Oryctolagus cuniculus</i> .
PAGE	Electroforesis en gel de poliacrilamida.



P_c	Concentración total de proteína.
PDB	Base de datos de proteínas.
<i>Pf</i> TIM	TIM de <i>Plasmodium falciparum</i> .
<i>Pfu</i> TIM	TIM de <i>Pyrococcus furiosus</i> .
PGA	Ácido 2-fosfoglicólico.
PGH	Fosfoglicolohidroxamato.
PM	Peso molecular.
PPA	α -amilasa de <i>Sus scrofa</i> .
<i>Pw</i> TIM	TIM de <i>Pyrococcus woesei</i> .
R	Constante universal de los gases.
ReTIM	Barril TIM rediseñado.
RevTIMs	TIMs con desplegamiento térmico reversible.
RMSD	Desviación cuadrática media de la raíz.
<i>Rm</i> TIM	TIM de <i>Rhipicephalus microplus</i> .
R_s	Radio de Stokes.
σ	Desviación estándar.
<i>Sc</i> TIM	TIM de <i>Streptomyces coelicolor</i> .
<i>Sc</i> TIM	TIM de <i>Sus scrofa</i> .
SDS	Dodecilsulfato sódico.
<i>Sj</i> TIM	TIM de <i>Schistosoma japonicum</i> .
sTIM	Barril TIM diseñado <i>de novo</i> .
T	Temperatura de referencia.
<i>Tb</i> TIM	TIM de <i>Trypanosoma brucei</i> .
<i>Tc</i> TIM	TIM de <i>Trypanosoma cruzi</i> .
TFA	α -amilasa de <i>Thermobifida fusca</i> .
TFZ	Z-score de la función de traslación.
TIM	Triosafosfato Isomerasa.
T_m	Temperatura media de desnaturalización en caliente.
T_m'	Temperatura media de desnaturalización en frío.
<i>Tm</i> TIM	TIM de <i>Thermotoga maritima</i> .
<i>Tmo</i> TIM	TIM de <i>Tenebrio molitor</i> .
TOC	Temperatura óptima de crecimiento.



T_0 TIM	TIM de <i>Thermococcus onnurineus</i> .
T_5	Temperatura de máxima estabilidad donde $\Delta S=0$.
T_5 TIM	TIM de <i>Taenia solium</i> .
T_7 TIM	TIM de <i>Thermus thermophilus</i> .
T_{tx} TIM	TIM de <i>Thermoproteus tenax</i> .
T_{VT} TIM	TIM de <i>Trichomonas vaginalis</i> .
μ	Promedio.
UA/AU	Unidades arbitrarias.
μg	Microgramo.
μL	Microlitro.
μM	Micromolar.
UVA	Absorción de luz ultravioleta.
V_{max}	Velocidad máxima de catálisis.
V_0	Velocidad inicial.
y_7 TIM	TIM de <i>Saccharomyces cerevisiae</i>



RESUMEN

El paisaje conformacional de las proteínas: estudio de barriles TIM naturales y diseñados *de novo*

Llamamos paisaje conformacional de las proteínas a la descripción fisicoquímica, esto es, en términos estructurales, termodinámicos y cinéticos, de los conformeros que la proteína visita al plegarse al estado nativo y al interactuar con otras moléculas para llevar a cabo su función. Un conocimiento a profundidad de este fenómeno es esencial para la predicción de la estructura proteica a partir de la secuencia de aminoácidos, el diseño de proteínas con propiedades nuevas y útiles, la comprensión de las teorías evolutivas, y el conocimiento de los orígenes moleculares de enfermedades debido al plegamiento anómalo de las proteínas.

Dentro de las diferentes conformaciones tridimensionales que pueden tomar las secuencias de aminoácidos, la superfamilia de los barriles TIM o barriles $(\beta/\alpha)_8$ es una de las más abundantes en la naturaleza, ya que aproximadamente el 10% de todas las proteínas conocidas adquieren esta estructura. Debido a su enorme variedad en funciones, su origen evolutivo común y a su ubicuidad en distintos tipos de proteínas, el barril TIM representa una topología de gran interés fisicoquímico para su estudio. El conocimiento a profundidad del paisaje conformacional que engloba las distintas propiedades de esta topología, así como la relación energético-estructural que existe entre éstas, tanto de proteínas naturales como diseñadas *de novo*, nos permitirá comprender mejor los determinantes globales que modulan la estabilidad y el plegamiento de los barriles TIM. Debido a todo lo anterior, el objetivo del presente trabajo fue conocer, explorar y caracterizar el paisaje conformacional de barriles TIM naturales y diseñados *de novo*.

Como miembro de un barril TIM natural se eligió a la Triosafosfato Isomerasa (TIM), una enzima glucolítica que cataliza la isomerización reversible del gliceraldehído-3-fosfato y la dihidroxiacetona fosfato. Se estudiaron TIMs pertenecientes a diversos fila bacterianos (BacTIMs) representados por las siguientes especies: *Deinococcus radiodurans* (*D α TIM*), *Nostoc punctiforme* (*N ρ TIM*), *Azotobacter vinelandii* (*A ν TIM*), *Gemmata obscuriglobus* (*G α TIM*), *Clostridium perfringens* (*C ρ TIM*) y *Streptomyces coelicolor* (*S τ TIM*).

Todas las proteínas fueron expresadas en *Escherichia coli* y purificados a homogeneidad, a excepción de *A ν TIM* que fue descartada del estudio debido a que la enzima se depositó en cuerpos de inclusión. Las BacTIMs caracterizadas poseen propiedades estructurales, catalíticas, espectroscópicas e hidrodinámicas compatibles con barriles TIM plegados nativamente y similares a otras Triosafosfato Isomerasas silvestres previamente estudiadas. La estructura tridimensional de las proteínas fue determinada mediante cristalografía con difracción de rayos X. Las estructuras de *G α TIM*, *D α TIM*, *C ρ TIM* y *S τ TIM* demuestran una topología canónica de un barril $(\beta/\alpha)_8$, con un RMSD menor a 1 Å entre ellas. La orientación de los monómeros con respecto a su contraparte en el dímero es también similar; de hecho, esta es una propiedad conservada a través de todas las TIMs en los distintos dominios de la vida.

El desplegamiento de las BacTIMs fue estudiado mediante tres agentes: dilución, urea y temperatura. Los experimentos de dilución arrojaron transiciones simples que fueron reversibles para *D α TIM*, *C ρ TIM* y *S τ TIM* e irreversible para las otras dos proteínas. Los experimentos de desplegamiento químico mostraron procesos reversibles de dos estados para *D α TIM*, *S τ TIM* y *G α TIM*, mientras que para *N ρ TIM* se encontró la presencia de un intermediario monomérico. Para *C ρ TIM*, no hubo cambio en la señal espectroscópica, por lo cual no se pudieron determinar los parámetros termodinámicos correspondientes.



El comportamiento en el desplegamiento térmico de las BacTIMs se agrupa en dos bloques: para *Np*TIM y *Go*TIM el proceso es irreversible, controlado cinéticamente, y para *Dn*TIM, *Sc*TIM y *Cp*TIM es un proceso reversible bajo control termodinámico, un fenómeno nunca antes descrito para una TIM. Los valores de energía de activación para *Np*TIM y *Go*TIM se encuentran entre los reportados previamente para otras TIMs eucariontes y es diferente entre ambas proteínas, lo que se refleja en una estabilidad cinética mayor para *Go*TIM con respecto a *Np*TIM. La estabilidad conformacional a 25 °C para *Dn*TIM, *Sc*TIM y *Cp*TIM es de la mitad en comparación con *Np*TIM, *Go*TIM y todas las otras TIMs estudiadas anteriormente; esta característica pudiera ser un “pago fisicoquímico” que otorgan las proteínas para tener reversibilidad en su proceso de desplegamiento térmico. La forma de la curva de estabilidad para *Dn*TIM, *Sc*TIM y *Cp*TIM es principalmente debida a los cambios en la T_s y el ΔH_{TS} , ya que el ΔC_p es casi idéntico entre las diferentes RevTIMs. Tras un análisis exhaustivo de distintas propiedades de las enzimas, proponemos cinco propiedades que relacionan con la reversibilidad en el desplegamiento térmico de los barriles TIM naturales estudiados: 1) bajo punto isoeléctrico, 2) ausencia o baja cantidad de estructura secundaria residual en el estado desplegado, 3) pequeño volumen en las cavidades internas del estado nativo, 4) baja T_m y 5) baja estabilidad conformacional. Dado que los tres barriles TIM naturales con desplegamiento térmico reversible provienen de organismos de diferentes fila, esto sugiere que la reversibilidad puede ser más común de lo que se conoce actualmente. Además, debido a que los barriles TIM oligoméricos son frecuentes en las células, la reversibilidad en su desplegamiento pudiera ser una ventaja evolutiva al reducir la probabilidad de agregación irreversible no productiva y la carga de la maquinaria celular del sistema de chaperonas.

Como miembro de un barril TIM diseñado *de novo* se eligió a la proteína sTIM11. Para esta proteína, uno de los objetivos fue incrementar su estabilidad con el fin de usarla como andamiaje para el diseño de funciones novedosas en el futuro. Bajo este fin, se diseñaron 28 nuevas proteínas, a las que nombramos como ReTIM; todas ellas fueron diseñadas incrementando el empaquetamiento en tres diferentes regiones del barril, además de la introducción de ciertas interacciones estabilizantes. Todos los diseños fueron sobreexpresados en *E. coli* y purificados a homogeneidad. Los diseños más exitosos de cada grupo fueron: ReTIM1, ReTIM6, ReTIM8, ReTIM17, ReTIM18 y ReTIM24. Para esta familia de proteínas también existen proteínas con desplegamiento térmico reversible e irreversible. De los diseños exitosos, tres de ellos cumplieron el objetivo de incrementar la estabilidad de sTIM11: ReTIM17, ReTIM18 y ReTIM24, quienes poseen los parámetros fisicoquímicos (T_m , ΔH , ΔC_p y ΔG) más altos obtenidos y cercanos a los estimados paraméricamente para una proteína de 184 residuos. De ellos, ReTIM17 tiene la estabilidad incrementada un 50% con respecto de sTIM11 y ReTIM24 posee una T_m casi 50 °C más que la proteína base. Para el caso de los barriles diseñados, T_s , ΔH_{TS} y ΔC_p cambian considerablemente entre las distintas ReTIMs, a diferencia de los barriles TIM naturales donde sólo los primeros dos parámetros son los que cambian entre las BacTIMs. El cambio en estos parámetros tiene un efecto específico en la forma de las curvas de estabilidad de las proteínas estudiadas. Además, se determinó la estructura tridimensional de una de ellas, ReTIM18; su estructura demuestra que las interacciones fueron correctamente diseñadas, ya que el RMSD entre el modelo arrojado por Rosetta y la estructura es de tan sólo 0.87 Å.

Los resultados presentados en este trabajo indican que para distintos miembros, ya sea naturales o diseñados *de novo*, de una familia que adopta una misma topología, en este caso el barril $(\beta/\alpha)_8$ o barril TIM, el desplegamiento puede estar bajo control cinético o termodinámico. Debido a todo esto, es interesante encontrar que una topología tan común, como lo es el barril TIM, pueda desplegarse/replegarse como propuso Anfinsen en 1973, es decir, sin la ayuda de la maquinaria celular.



ABSTRACT

The conformational landscape of proteins: study of natural and *de novo* designed TIM barrels

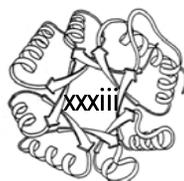
We call conformational landscape of proteins to the physicochemical description, that is, in terms of structural, thermodynamic and kinetic features, of the conformers that the protein visits when it folds to the native state and when it interacts with other molecules to carry out its function. An in-depth knowledge of this phenomenon is essential for protein structure prediction from amino acid sequence, protein design of novel and useful properties, understanding of evolutionary theories, and the knowledge of molecular origins of many diseases due to anomalous protein folding.

Within the different three-dimensional conformations that amino acid sequences can adopt, the superfamily of TIM barrel or $(\beta/\alpha)_8$ barrel is one of the most abundant in nature, since approximately 10% of all known proteins acquire this structure. Due to its enormous variety in functions, its common evolutionary origin and its ubiquity in different types of proteins, the TIM barrel represents a topology of great physicochemical interest for its study. The knowledge of the conformational landscape that contains the different properties of this topology, as well as the energetic-structural relationship that exists among these, both natural and *de novo* designed proteins, will allow a best understanding of the global determinants that modulate stability and folding of TIM barrel proteins. Due to all of the above, the aim of this work was to know, to explore and to characterize the conformational landscape of natural and *de novo* designed TIM barrels.

Triosphosphate Isomerase (TIM) was chosen as a member of a natural TIM barrel, a glycolytic enzyme that catalyzes the isomerization of glyceraldehyde-3-phosphate and dihydroxyacetone phosphate. Here, we study TIMs belonging to diverse bacterial phyla (BacTIMs): *Deinococcus radiodurans* (*D*TIM), *Nostoc punctiforme* (*Np*TIM), *Azotobacter vinelandii* (*Av*TIM), *Gemmata obscuriglobus* (*Go*TIM), *Clostridium perfringens* (*Cp*TIM) and *Streptomyces coelicolor* (*Sc*TIM).

All proteins were expressed in *Escherichia coli* and purified to homogeneity, except for *Av*TIM which was discarded from the study because the enzyme was deposited in inclusion bodies. BacTIMs have structural, catalytic, spectroscopic and hydrodynamic properties compatible with TIM barrels natively folded and similar to other previously studied wild type TIMs. The three-dimensional structure of the proteins was determined by crystallography with X-ray diffraction. The structures of *Go*TIM, *D*TIM, *Cp*TIM and *Sc*TIM show a canonical topology of a barrel $(\beta/\alpha)_8$, with an RMSD of less than 1 Å among them. The orientation of the monomers with respect to its counterpart in the dimer is also similar; in fact, this is a conserved property of all TIMs in three domains of life.

The unfolding of BacTIMs was studied by three different agents: dilution, urea and temperature. Dilution experiments showed simple transitions that were reversible for *D*TIM, *Cp*TIM and *Sc*TIM and irreversible for the other two proteins. Chemical unfolding experiments showed two state reversible processes for *D*TIM, *Sc*TIM and *Go*TIM, whereas for *Np*TIM the presence of a monomeric intermediate was found. For *Cp*TIM there was no change in the spectroscopic signal, so the corresponding thermodynamic parameters could not be determined.



The thermal unfolding behaviour in BacTIMs is grouped into two blocks: for *Np*TIM and *Go*TIM the process is irreversible, kinetically controlled, and for *Dt*TIM, *Sc*TIM and *Cp*TIM is a reversible process under thermodynamic control, a phenomenon previously never described for a TIM. Activation energy values for *Np*TIM and *Go*TIM are similar to those reported previously for other eukaryotic TIMs and the values are different between both proteins, which results in a higher kinetic stability for *Go*TIM with respect to *Np*TIM. The conformational stability at 25 °C for *Dt*TIM, *Sc*TIM and *Cp*TIM is half compared to *Np*TIM, *Go*TIM and all other TIMs so far studied; this characteristic could be a "physicochemical fee" that the proteins grant to have reversibility in their thermal unfolding process. The shape of the stability curve for *Dt*TIM, *Sc*TIM and *Cp*TIM is mainly due to changes in T_s and ΔH_{TS} , since ΔC_p is almost identical among different RevTIMs. After an exhaustive analysis of different properties of the enzymes, we propose five properties that can correlate with the reversibility in the thermal unfolding of the natural TIM barrels studied: 1) low isoelectric point, 2) absence or low amount of residual secondary structure in the unfolded state, 3) small volume in the internal cavities of the native state, 4) low T_m and 5) low conformational stability. Since the three natural TIM barrels with reversible thermal unfolding come from organisms in different phyla, this suggests that reversibility may be more common than what is currently known. In addition, because oligomeric TIM barrels are frequent in cells, reversibility in their folding/unfolding process can be an evolutive advantage because it reduces the chance of non-productive irreversible aggregation and the load of the cellular chaperonin machinery.

As a member of a *de novo* designed TIM barrel we chose the sTIM11 protein. For this protein, the aim was to increase the stability in order to use this scaffold for its molecular functionalization in a future. We designed 28 new proteins, which we named as ReTIM; all of them were designed to increase the packing in three different regions of the barrel and to introduce favorable interactions. All the designs were overexpressed and purified to homogeneity. The most successful designs of each group were: ReTIM1, ReTIM6, ReTIM8, ReTIM17, ReTIM18 and ReTIM24. For this protein family also exists proteins with reversible and irreversible thermal unfolding. Of the successful designs, three of them fulfilled the aim of increasing the stability of sTIM11: ReTIM17, ReTIM18 and ReTIM24, who possess the highest physicochemical parameters obtained (T_m , ΔH , ΔC_p and ΔG) and close to those estimated parametrically for a 184 residue protein. Of them, ReTIM17 has the stability increased by 50% with respect to sTIM11 and ReTIM24 possesses a T_m almost 50 °C more than the mother protein (sTIM11-SS). For the case of the designed barrels, T_s , ΔH_{TS} and ΔC_p change considerably on the different ReTIMs, unlike the natural TIM barrels where only the first two parameters change among the BacTIMs. The modification on these parameters has a specific effect on the shape of the protein stability curves determined. In addition, the three-dimensional structure of one of them, ReTIM18, was determined; its structure shows that the interactions were correctly designed, since the RMSD between the Rosetta model and the structure is only 0.87 Å.

Results presented in this work indicate that for different members, either natural or *de novo* designed, of a family that adopts the same topology, in this case the $(\beta/\alpha)_8$ barrel or TIM barrel, the unfolding may be under kinetic or thermodynamic control. It is therefore very thought-provoking to find that a common topological ensemble, such as the TIM barrel, can unfold/refold as proposed by Anfinsen in 1973, *i.e.* without the help of the cellular machinery.

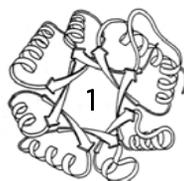


INTRODUCCIÓN

El primer deber de un hombre es pensar por sí mismo

José Martí, Hombre de Campo: 1885

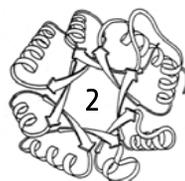
Uno de los métodos creativos que hemos utilizado los seres humanos para explorar la naturaleza es el tomar cualquier conocimiento de nuestra existencia y aplicar una mirada totalmente diferente a la establecida: cambiar la perspectiva. De todos los conceptos que la humanidad ha explorado, uno de los que más se han examinado y que guía muchas de nuestras acciones es el referente a la vida. Lejos de preguntarnos qué es la vida, una realidad de ésta es que posee una belleza peculiar que puede ser apreciada desde diferentes aristas. La belleza de la vida biológica impone incesantemente en nosotros su presencia; el concepto de belleza ha sido un rompecabezas para biólogos, teólogos, filósofos, historiadores y hasta para poetas. Tanta es su importancia, que Sócrates hace más de 2400 años llegó a preguntarse sobre las manifestaciones de la belleza en sus distintas formas: en una rosa, en la música, en una obra teatral o en el conocimiento mismo. *La belleza es difícil*, afirmaba Platón: por qué un cuerpo humano es hermoso y otro no lo es; por qué un paisaje golpea dulcemente y otro causa repulsión; por qué una pintura atrae y otra ocasiona rechazo; por qué algunas composiciones musicales, literarias, arquitectónicas o naturales, nos hacen exclamar ¡qué bello!, mientras tantas otras pasan desapercibidas o sencillamente desagradan. Igualmente, el mismo Platón en su libro clásico, *Fedro*, describe cómo Sócrates lo inicia en los misterios de la eterna belleza. Sócrates invita a Platón a contemplar con él aquellas ciencias bellas e imperfectas. Le enseña a alimentar su inteligencia con lo verdadero, lo bello y lo bueno, justo para merecer un día tomar sus alas y descubrir la esencia de la belleza en este mundo imperfecto. A mi parecer, una forma de descubrir la esencia de este mundo imperfecto es lo que el hombre ha construido bajo el campo de la química, la física y la biología, áreas con una belleza innegable.



La biología contemporánea está llena de disciplinas que se encargan de estudiar las muy diversas formas y funciones de los organismos vivos y de aquellos que se extinguieron. Ya sea en una parte anatómica, un gen, un proceso bioquímico o un conjunto de biomoléculas, la biología se encarga del estudio de los seres vivos y, gracias al conocimiento acumulado, hoy sabemos que parte esencial de los procesos vitales es posible debido a la presencia de unas macromoléculas llamadas proteínas. Las proteínas median los procesos fundamentales de la vida, y las maneras hermosas y variadas en que lo hacen ha sido el foco de décadas de investigación biomédica en el área.

Las proteínas son las biomoléculas encargadas de un gran número de funciones en las células, forman parte de la estructura básica de los tejidos como músculos, cabello, piel, mucosas, etc., y desempeñan funciones metabólicas tan diversas dentro y fuera de la célula como la asimilación de nutrientes, transporte de moléculas, protección inmunitaria, acciones hormonales, almacenamiento y reserva de compuestos químicos, reacciones enzimáticas, entre otras. De las proteínas depende en gran parte el buen funcionamiento celular, siendo elementos que definen la identidad de los seres vivos, pues con ellas se expresa el código genético contenido en el genoma; sus funciones forman parte del fenotipo de los seres vivos (Pace, 1996; Berg *et al.*, 2002; Voet *et al.*, 2006; Mathews *et al.*, 2012).

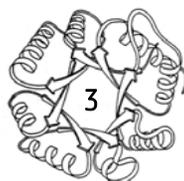
Llamamos paisaje conformacional a la descripción fisicoquímica, esto es, en términos estructurales, termodinámicos y cinéticos, de los conformeros que la proteína visita al plegarse al estado nativo y al interactuar con otras moléculas para llevar a cabo su función. Las proteínas son biopolímeros de origen genético, por lo tanto, al describir el paisaje conformacional de proteínas homólogas y variar la secuencia de aminoácidos, es posible estudiar la evolución de sus propiedades moleculares. El proceso por el cual la cadena polipeptídica de una proteína alcanza su estructura tridimensional biológicamente funcional —conocida como la estructura nativa— se denomina plegamiento y es posiblemente el ejemplo más fundamental de autoensamblaje molecular en un contexto biológico. Un conocimiento a profundidad de este proceso es de gran importancia para (i)



acercarse a alcanzar el objetivo de poder predecir la estructura proteica a partir de la secuencia de aminoácidos, (ii) poder realizar el diseño de proteínas con propiedades nuevas y útiles, (iii) profundizar en muchos aspectos de las teorías evolutivas y (iv) aumentar el conocimiento de los orígenes moleculares de muchas enfermedades que son atribuidas al plegamiento anómalo de proteínas (Sánchez-Ruiz, 2011).

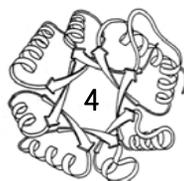
El proceso de plegamiento/desplegamiento de una proteína implica la formación y ruptura de numerosas interacciones débiles, no covalentes. Desde una perspectiva teórica, el plegamiento de una proteína se concibe como el movimiento de un punto representando una conformación en una superficie de energía libre interna frente a los grados de libertad conformacional. Dicha energía es un paisaje de alta dimensionalidad (correspondiente a las miles de posibles conformaciones) y, en principio, podría ser considerablemente rugoso, esto es, con un gran número de mínimos y máximos locales (Dill y Chan, 1997). Sin embargo, para las proteínas, se espera que el paisaje conformacional tenga la forma general de un embudo con un pequeño grado de rugosidad, de tal manera que la proteína sea rápidamente dirigida hacia el fondo del embudo (el estado nativo), aunque las moléculas individuales comiencen en diferentes lugares en el paisaje y sigan trayectorias distintas para llegar a la parte inferior. En otras palabras, un paisaje relativamente suave evita el plegamiento de una forma aleatoria y hace que las proteínas se plieguen en tiempos biológicamente compatibles con la vida (Bryngelson *et al.*, 1995).

Los avances en el estudio de las proteínas, hasta el momento, han permitido establecer una conexión entre el paisaje de energía del plegamiento de las proteínas y los datos experimentales cinéticos y al equilibrio. El análisis de la estabilidad de una proteína permite conocer la energía que diferencia al estado nativo de otras conformaciones "no nativas" o desnaturalizadas; comprender esta propiedad de las proteínas permite, entre muchas otras cosas, conocer el paisaje conformacional del plegamiento de una proteína y saber las barreras energéticas que rigen este proceso (Gruebele, 2005). La estabilidad cinética se refiere al hecho de que las funciones biológicas de una proteína



deben ser mantenidas durante una escala de tiempo fisiológica aun si el estado nativo no es termodinámicamente estable con respecto a las formas no funcionales (estados desplegados, agregados o proteína proteolizada). Dicha estabilidad es requerida en varios casos porque las proteínas a veces trabajan bajo condiciones que favorecen las alteraciones irreversibles, por lo que la existencia de grandes barreras energéticas que prevengan estas formas no funcionales es un mecanismo de selección a través del tiempo. Una estrategia para conocer dichas barreras es realizar experimentos de desnaturalización térmica que permiten describir la energía que rige el proceso y la escala de tiempo en que sucede éste. Además, la estabilidad cinética determinada *in vitro* puede ser un reflejo de la selección natural para la estabilidad cinética *in vivo*, que asegura el correcto funcionamiento de las proteínas y la homeostasis celular (Sánchez-Ruíz, 2010).

A partir de esta discusión inicial, a lo largo del presente escrito se abordarán más ampliamente algunos de los aspectos antes mencionados. Se iniciará con una descripción general actual del problema del plegamiento de las proteínas y cómo se ha abordado desde un enfoque fisicoquímico y evolutivo. Posteriormente, se describirá la topología modelo que se estudiará en el presente trabajo, los barriles $(\beta/\alpha)_8$ o barriles TIM, además de las razones de su importancia en el mundo de las arquitecturas moleculares de las proteínas. Enseguida, se describirá el barril TIM natural modelo del presente trabajo, la Triosafosfato Isomerasa, una proteína con características interesantes que ha ayudado a responder algunas de las preguntas que se tienen acerca del plegamiento proteico. Finalmente, se discutirá sobre la ingeniería de proteínas con el objeto de estabilizar un barril TIM diseñado *de novo*. En conjunto, se tratará de describir la relación entre las propiedades fisicoquímicas de las proteínas, sus mecanismos de plegamiento y evolución y cómo estos temas pueden ser abordados en una topología específica como lo son los barriles TIM.



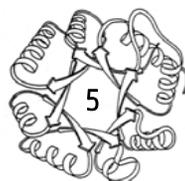
ANTECEDENTES GENERALES

Plegamiento de las proteínas

El problema del plegamiento de proteínas y los embudos de energía

El que estructuras proteicas complejas se formen espontáneamente en periodos cortos de tiempo es una necesidad evolutiva (Wolynes *et al.*, 1996). El plegamiento es el proceso fisicoquímico mediante el cual las cadenas polipeptídicas adquieren su estructura nativa o conformación de mínima energía. La concepción actual del plegamiento se deriva de los estudios realizados por Christian Boehmer Anfinsen a mediados del siglo XX, quien a través de experimentos con la ribonucleasa pancreática bovina estableció que la desnaturalización es, por lo general, un proceso reversible y que la información necesaria para que una proteína adquiriera su conformación nativa se encuentra codificada en la secuencia de aminoácidos (Anfinsen *et al.*, 1961; Anfinsen, 1973).

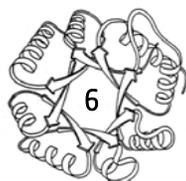
En 1969, Cyrus Levinthal mostró que la búsqueda de la estructura nativa a través de todo el universo conformacional es prácticamente imposible en un tiempo biológicamente relevante. Si suponemos que hay dos conformaciones posibles por aminoácido, una proteína de 100 residuos podría adquirir hasta 2^{100} conformaciones diferentes. Si el ir de una conformación a otra tardara un picosegundo, muestrear todo el universo conformacional tomaría un tiempo de 10^{18} segundos, lo que es igual a 10^{10} años, tiempo mayor que la edad del universo mismo. Esta cifra es completamente irreal en un contexto biológico donde el tiempo de plegamiento es del orden de segundos o minutos. A esto se le conoce como la paradoja de Levinthal (Bryngelson, 1995). Como solución a la paradoja, Levinthal propuso la existencia de vías de plegamiento. Esta forma de ver el proceso implica que el estado desnaturalizado (formado por un conjunto de conformaciones específicas) llega al estado nativo generalmente a través de una serie de pocos pasos secuenciales (Dill y Chan, 1997).



De los resultados de Anfinsen se postuló la hipótesis termodinámica, donde se establece que la estructura tridimensional de una proteína en su medio fisiológico es aquella en la cual la energía libre de Gibbs del sistema es menor (función que describe la energía del sistema proteína-disolvente). Es decir, la conformación está determinada por la totalidad de las interacciones interatómicas y, por tanto, por la secuencia de aminoácidos en un medio determinado (Creighton, 1990).

Cada conformación que la proteína puede adoptar tiene asociada una energía determinada y al conjunto de las energías correspondientes a todas las conformaciones posibles se le conoce como paisaje energético. Actualmente, se propone con base en diversos estudios teóricos y experimentales que el paisaje energético de las proteínas tiene forma de embudo, donde el estado desnaturalizado se compone de todo un conjunto de conformaciones de alta energía y entropía. El espacio conformacional se reduce conforme la proteína se acerca al estado de menor energía (nativo) en el pico del embudo (figura 1). En muchos casos este proceso ocurre *in vitro*, sin la ayuda de otro componente en el sistema, lo que reduce al plegamiento a un problema fisicoquímico, gobernado por las diferencias en energía y las barreras cinéticas de los estados conformacionales que adopta la cadena en su camino hacia el estado nativo (modelo del embudo multiestado-energético, figura 1) (Dill, 1990; Dill y Chan, 1997; Goodsell y Olson, 2000).

El plegamiento sigue siendo un problema abierto. El modelo más comúnmente usado para analizar este proceso es el que postula únicamente dos estados (nativo y desnaturalizado). Sin embargo, la descripción completa del mecanismo de plegamiento involucra no sólo el estudio de los estados inicial y final, sino también el de los estados intermedios y de los estados de transición que los conectan. En algunas condiciones, es posible observar conformaciones intermedias parcialmente estructuradas. La caracterización de estas especies proporciona valiosa información acerca de las vías y los mecanismos que dirigen a este proceso (Jaenicke, 1991).



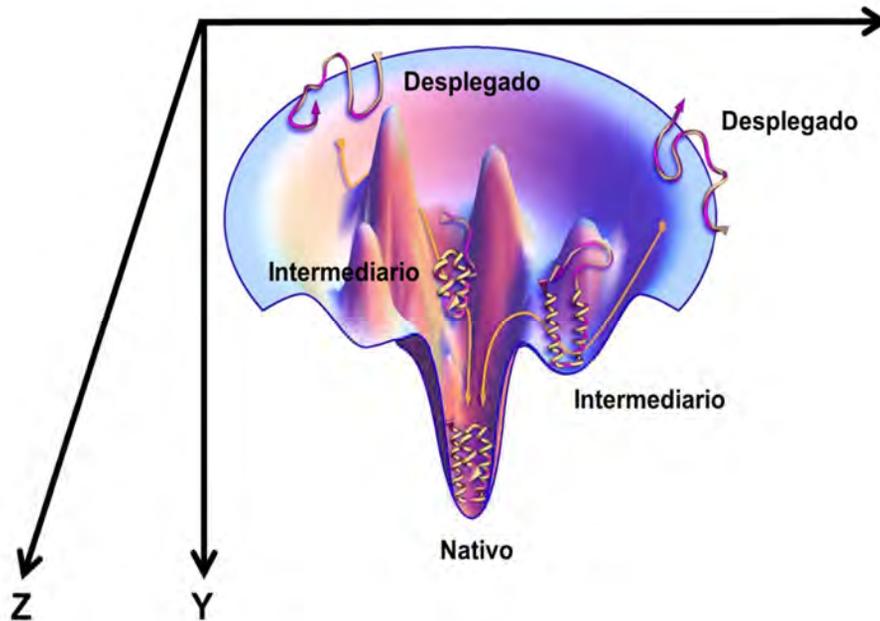
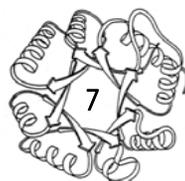


Figura 1. Modelo del embudo para el plegamiento de una proteína. Se muestra el estado nativo con el mínimo de energía que se alcanza mediante un mecanismo de plegamiento multiestado (empezando con distintas moléculas desplegadas y poblando ninguno, uno o más intermediarios en el proceso). El eje vertical del embudo (eje Y) representa la suma de las energías libres que estabilizan una cadena polipeptídica en una conformación particular. El área lateral de un paisaje de energía (eje X) a una profundidad dada (eje Z) representa el número de conformaciones que tiene la cadena polipeptídica con esa energía libre (modificado de Dill y Chan, 1997).

Agentes perturbantes para el estudio del plegamiento de proteínas

El plegamiento de proteínas es muy complejo ya que engloba diversos factores internos y externos de las proteínas. Para dilucidar cómo ocurre este proceso se realizan estudios *in vitro*, *in vivo* e *in silico*. En el proceso *in vivo* existen factores auxiliares como las chaperonas, que dirigen el proceso a una vía productiva, empero, en algunos casos, el plegamiento puede llevarse a cabo sin la necesidad de estos factores, hecho demostrado en los estudios realizados *in vitro*, los cuales han arrojado una amplia fuente de datos que han sido utilizados por los modelos computacionales para simular el plegamiento *in silico*. La meta es poder predecir en un futuro la estructura tridimensional de una

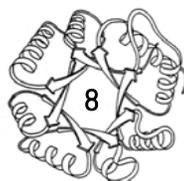


proteína a partir de su secuencia de aminoácidos sin utilizar información estadística, es decir, sin utilizar ninguna información de otra secuencia homóloga (Dobson, 2003).

Para estudiar el plegamiento *in vitro* de las proteínas, éstas se aíslan de la célula que las produce en su estado nativo de modo que el plegamiento de las mismas es un proceso que se debe estudiar por medio de reacciones de desnaturalización y renaturalización (Mathews *et al.*, 2012). La finalidad de estudiar este proceso, se basa en determinar la estabilidad de sus niveles estructurales y la ruta de plegamiento que sigue la proteína hasta el estado nativo, necesitando como primera característica para su estudio, que el proceso sea 100% reversible y que las condiciones experimentales estén al equilibrio (Dill, 1991; Creighton, 1996). Se ha observado que, en general, las proteínas pequeñas de aproximadamente 100 aminoácidos y monoméricas se desnaturalizan reversiblemente, mientras que las proteínas de mayor tamaño u oligoméricas lo hacen irreversiblemente o son parcialmente reversibles, además de que se agregan o se precipitan antes de que la condición desnaturalizante sea modificada a su estado original (Garel, 1992).

Entre los agentes más utilizados en la desnaturalización se encuentra el uso de variaciones físicas como la temperatura o la presión, o bien, agentes químicos como la variación del pH (altera los estados de ionización de las cadenas laterales de aminoácidos produciendo el desplegamiento), desnaturalizantes como el hidrocloreuro de guanidinio (GdnHCl) o la urea (permiten que la solubilidad de las cadenas laterales no polares de la proteína se incremente en agua) y detergentes como el SDS (interactúan con los residuos no polares de la proteína e interfieren con su centro hidrofóbico) (Dill y Shortle, 1991; Dobson, 2003).

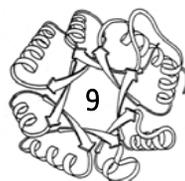
Por otro lado, uno de los agentes desnaturalizantes más utilizados es la temperatura, básicamente porque es un agente físico a cuyas variaciones estamos sometidos todos los seres vivos. De todas las condiciones ambientales a las que nos enfrentamos, existen algunas condiciones que fácilmente pueden ser controladas mediante la presencia de una membrana, como la osmolaridad, el pH o ciertos tipos de presión. Sin embargo, existen otras condiciones que son imposibles de apartar



del medio celular, ejemplo de esto es la temperatura. Al utilizar la temperatura como agente desnaturante, aumenta la energía cinética de las moléculas con lo que se desorganiza la proteína y se desnatura. Asimismo, un aumento de la temperatura destruye las interacciones débiles y desordena la estructura de la proteína, de forma que el interior hidrofóbico interactúa con el medio acuoso y se produce la desnaturación y, en algunos casos, la agregación y precipitación de la proteína desnaturada (Pain, 2000; Razvi y Scholtz, 2006).

El plegamiento y la estabilidad de las proteínas: dos propiedades fisicoquímicas íntimamente relacionadas

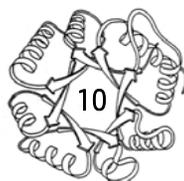
La estabilidad de una proteína es una medida de la energía que diferencia al estado nativo de otros estados "no nativos" o desnaturados. La estabilidad termodinámica se obtiene cuando es posible calcular la diferencia de energía entre el estado nativo y el desnaturado, para lo cual se requiere la reversibilidad y el equilibrio en el proceso de desnaturación. Por "proceso al equilibrio" se entiende al equilibrio termodinámico entre los estados que visita la proteína en su ruta de plegamiento/desplegamiento (nativo, desplegado, parcialmente desplegado, etc.) y que, para el caso específico del desplegamiento térmico, es establecido a través de un barrido de temperatura (Sturtevant, 1977). En otras palabras, las concentraciones de los estados poblados a cualquier temperatura están determinadas sólo por los valores de las constantes de equilibrio. Por ejemplo, para un proceso de dos estados al equilibrio con un valor de entalpía de desnaturación suficientemente alta, la constante de equilibrio de desnaturación cambia de un valor bajo que favorece el estado nativo a un valor alto que favorece el estado desnaturado en un rango de temperatura relativamente estrecho (Prabhu, 2005). Los procesos que se producen a través de una serie de estados al equilibrio se llaman reversibles. Cuando la transición del estado nativo al desplegado es reversible, es posible calcular todos los parámetros termodinámicos asociados a ella: temperatura media de desnaturación (T_m), cambio de entalpía (ΔH) y cambio en la capacidad



calorífica (ΔC_p), además de parámetros derivados como cambio en la entropía (ΔS) y cambio en la energía libre de Gibbs (ΔG), que se obtienen mediante relaciones termodinámicas básicas. Esta detallada descripción energética puede ser interpretada en términos estructurales usando correlaciones energético-estructurales que se han desarrollado con los años.

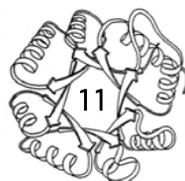
Cuando el proceso del estado nativo al desplegado no está al equilibrio (en el sentido termodinámico señalado antes) sino que es dependiente del tiempo (es decir, cinético), se dice que el proceso está controlado cinéticamente y, por tanto, el análisis de los datos experimentales debe realizarse de manera diferente a la empleada cuando es un proceso al equilibrio. En una transición que es controlada cinéticamente las concentraciones de los estados involucrados en ella están determinadas por ecuaciones de velocidad (y no por constantes de equilibrio dependientes de temperatura) (Sánchez-Ruiz, 1992; Plaza del Pino *et al.*, 2000). Las dos fuentes más comunes de control cinético en proteínas son el equilibrio lento y la presencia de alteraciones irreversibles de las proteínas (agregación, autoproteólisis, etc.). Estas alteraciones irreversibles llevan a la proteína a un estado final que es incapaz de replegarse de regreso al estado nativo. De esta forma, la estabilidad cinética se refiere a la capacidad de diferenciar energéticamente la proteína nativa del estado de transición (E^\ddagger) (el estado limitante en el proceso de desnaturalización) que da lugar al estado final. Este tipo de estabilidad es únicamente la que se puede determinar en proteínas que se desnaturalizan irreversiblemente, aunque también es posible hacerlo en las proteínas reversibles, puesto que el proceso de desnaturalización también presenta un paso limitante (Pain, 2000; Rees y Robertson, 2001).

Cuando la proteína en estudio tiene un desplegamiento que está bajo control cinético, evaluar su estabilidad -conocida como estabilidad cinética- consiste en conocer el tiempo durante el cual la proteína permanecerá en el estado nativo a una cierta temperatura T' . Este tiempo puede estimarse a partir del valor de una constante cinética k a T' empleando el gráfico de Arrhenius obtenido mediante experimentos de Calorimetría Diferencial de Barrido (DSC, por sus siglas en



inglés). Cuando la temperatura de interés T' (por ejemplo temperatura ambiente o fisiológica) está fuera del intervalo donde se construyó el gráfico de Arrhenius (que corresponde al intervalo en el que aparece la endoterma en la traza calorimétrica) la evaluación del tiempo de estabilidad implica realizar una extrapolación, de tal manera que a esa T' la estabilidad cinética de la proteína es sólo una estimación (entre más cerca esté T' del intervalo, más confiable será el valor obtenido) (Pain, 2000; Rees y Robertson, 2001).

La termodinámica proporciona un criterio claro para determinar si un proceso es o no espontáneo: si el cambio de energía libre de Gibbs es negativo ($\Delta G < 0$), el proceso ocurre de manera espontánea. Entonces, para el proceso de desplegamiento de una proteína (del estado nativo al desnaturizado: $N \rightarrow D$) cuando $\Delta G > 0$ el estado estable es el nativo. A la temperatura a la cual el $\Delta G > 0$ es máximo se le conoce como la temperatura de máxima estabilidad térmica o termodinámica de la proteína. A la representación de ΔG vs. T mostrada en la figura 2 se le conoce como la curva de estabilidad y únicamente puede ser determinada cuando el desplegamiento térmico es reversible y se está al equilibrio en las condiciones experimentales exploradas; en otras palabras, cuando es posible determinar la estabilidad termodinámica (Becktel y Schellman, 1987). Cada proteína tiene su propia curva de estabilidad y ésta es diferente de acuerdo a las condiciones experimentales usadas en la determinación de los parámetros termodinámicos (pH, solutos, amortiguador, etcétera). La curva de estabilidad muestra que la T_m a alta temperatura no es única, sino que además existe un valor de T_m a bajas temperaturas (T_m') (Pace y Tanford, 1968; Privalov, 1990). A estas dos temperaturas la proteína se desnaturaliza y, por ello, se denominan coloquialmente las temperaturas a las cuales ocurre la desnaturalización en caliente (T_m) y desnaturalización en frío (T_m'), respectivamente. Para muchas proteínas no es posible medir la segunda, debido a que esta T_m' se encuentra por debajo de 0°C (273.15 K) y, por lo tanto, la muestra se congela antes de alcanzarla.



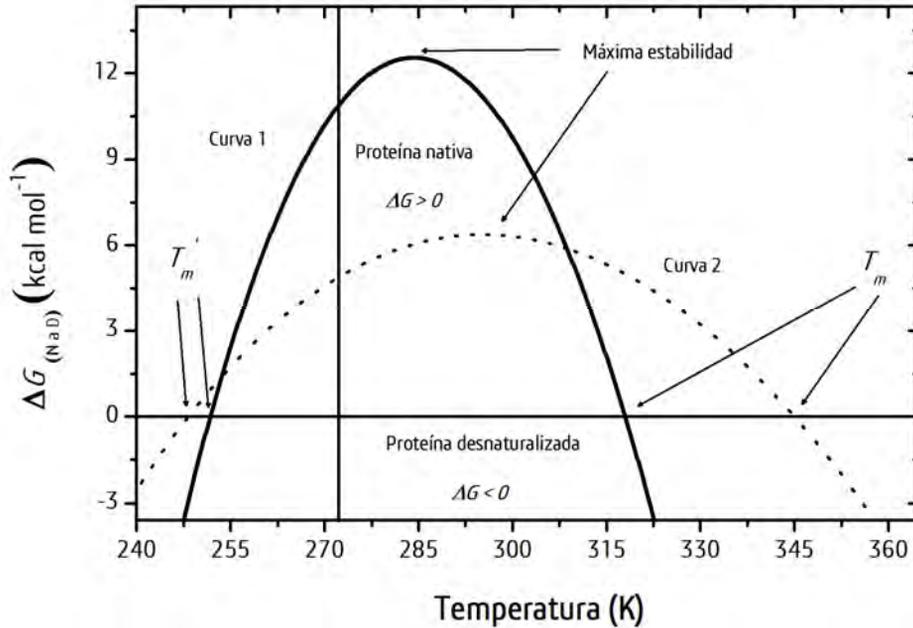
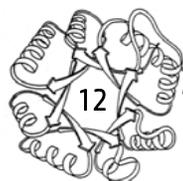


Figura 2. Curvas de estabilidad para dos proteínas. Esta curva se obtiene empleando los valores de ΔH , ΔC_p y T_m que se obtienen de los experimentos de DSC. En la zona del diagrama donde $\Delta G(N \rightarrow D) > 0$ el estado nativo es estable. Fuera de esa zona, donde el $\Delta G(N \rightarrow D) < 0$, el estado estable es el desnaturalizado. T_m y T_m' indican las temperaturas medias de desnaturalización en caliente y en frío, respectivamente. La línea vertical indica la temperatura a la cual la muestra se congela.

La evaluación del ΔG requiere que el proceso de desplegamiento esté en equilibrio y sea reversible. La expresión de la termodinámica que permite el cálculo de la curva de estabilidad es (ecuación 1):

$$\Delta G(T) = \Delta H \left(1 - \frac{T}{T_m} \right) - \Delta C_p \left(T_m - T + T \ln \left(\frac{T}{T_m} \right) \right) \quad (1)$$

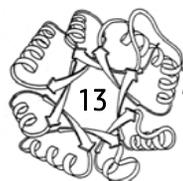
donde los valores de ΔH , ΔC_p y T_m se obtienen de los experimentos de DSC. Para obtener esta ecuación (Rees y Robertson, 2001) se ha empleado como temperatura de referencia a la T_m , es decir $\Delta G = 0$ a la T_m . Es interesante hacer notar que la curva de estabilidad sería una recta si el ΔH y el ΔS para el proceso fueran independientes de la temperatura (Fersht, 1999), es decir, si fuesen



constantes. Sin embargo, hace más de cuatro décadas, se encontró que el proceso de desplegamiento (y por tanto su inverso, el plegamiento) está asociado con cambios importantes en la capacidad calorífica, esto es, que el ΔH y el ΔS no son independientes de la temperatura. Además, con el descubrimiento del efecto hidrofóbico y su singular dependencia de la temperatura (Frank y Evans, 1945; Kauzmann, 1959) quedó claro que la simplificación de que $\Delta C_p = 0$ era una representación muy pobre de lo que en realidad sucedía en el desplegamiento de las proteínas. Por otro lado, fue el desarrollo de la DSC lo que permitió la evaluación experimental del ΔC_p . Con diversos estudios pioneros en el campo de las proteínas, se encontró que el ΔC_p de desplegamiento de una proteína es grande y positivo, y que, dentro del error experimental, éste puede tomarse como una constante para una proteína determinada, es decir es independiente de la temperatura (Brandts, 1964; Privalov y Khechinashvili, 1974; Shiao *et al.*, 1971; Privalov, 1980).

Debido a la ausencia de reversibilidad en el proceso de desplegamiento térmico de muchas proteínas, las curvas de estabilidad se han obtenido para relativamente pocas proteínas, principalmente para proteínas monoméricas pequeñas (Doyle *et al.*, 2013). Es también importante mencionar que es común en la literatura encontrar la aseveración que establece que una proteína A es más estable que otra proteína B cuando su T_m es mayor ($T_{mA} > T_{mB}$). En vista de la figura 2, este tipo de aseveraciones no son necesariamente correctas (la proteína A puede tener un ΔG menor aunque su T_m sea mayor) por no ser suficientemente claras o precisas. El lenguaje correcto debiera ser que la proteína A es más termo-resistente que la proteína B (compárense las curvas 2 y 1 en la figura 2).

Entre las propiedades termodinámicas, la capacidad calorífica a presión constante goza de una situación privilegiada. Esto es, debido a la relación que existe entre el ΔC_p y la entropía ΔS y la energía libre de Gibbs ΔG : $\Delta C_p = T(d\Delta S/dT)_p = -T(d^2\Delta G/dT^2)_p$. La relación con ΔG indica que el ΔC_p es el responsable de la curvatura en la curva de estabilidad. Considerando la interpretación microscópica de la entropía, la relación con el ΔS indica que el ΔC_p es un reflejo del nivel de

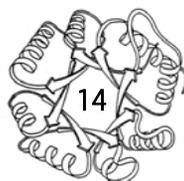


organización o estructura en el sistema. Más precisamente, el ΔC_p es una medida de la variación con temperatura de esa organización o estructura. Así, la relación entre ΔC_p y ΔS permite establecer -de manera cualitativa- un vínculo entre las propiedades termodinámicas macroscópicas determinadas experimentalmente y las características o cambios a nivel microscópico que ocurren durante el proceso, es decir explicar o racionalizar el proceso de desplegamiento (o del plegamiento) en términos microscópicos (Ibarra-Molero y Sánchez-Ruiz, 2006). Por ejemplo, el $\Delta C_p > 0$ para la desnaturalización es resultado de un $\Delta S > 0$, es decir de que la entropía del estado final (desnaturalizado) es mayor que la del estado inicial (nativo). A nivel microscópico, el $\Delta C_p > 0$ refleja la exposición al agua (hidratación) de los grupos hidrofóbicos ocultos en el estado nativo de la proteína que quedan expuestos al disolvente durante el desplegamiento (Fersht, 1999).

Los valores de ΔC_p también contienen información relevante referente a los estados estructurales de la proteína. Por ejemplo, la capacidad calorífica de una proteína en su estado desplegado (C_{PD}) puede ser calculado a partir de la secuencia de aminoácidos: si todos los aminoácidos están expuestos al agua, entonces el C_{PD} puede aproximarse a ser la suma de las contribuciones individuales de los aminoácidos constituyentes (cadenas laterales más cadena principal) (ecuación 2; Privalov y Makhatadze, 1990):

$$C_{PD} = \left(\sum_{i=1}^{20} n_i * C_{P,i} \right) + (N_{AA} - 1) * C_{P,bb} + C_{P,NH_2} + C_{P,COOH} \quad (2)$$

donde la sumatoria refiere a los 20 aminoácidos, n_i es el número de aminoácidos de cada tipo y $C_{P,i}$ es la capacidad calorífica molar de esa cadena lateral; N_{AA} es el número de aminoácidos totales en la proteína, $C_{P,bb}$ es la capacidad calorífica de la unidad básica (esqueleto) de la cadena principal (-CHCONH-), C_{P,NH_2} la capacidad calorífica del amino terminal y $C_{P,COOH}$ la capacidad calorífica del carboxilo terminal.



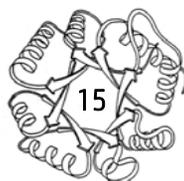
Por otro lado, la capacidad calorífica del estado nativo (C_{pN}) es una función lineal de la temperatura dentro del rango en el cual este puede ser medido. Dentro de ese rango, la capacidad calorífica de los estados nativos es muy similar para todas las proteínas cuando las trazas son normalizadas por el peso molecular. Se ha descrito que el ΔC_p entre los estados nativo y desplegado es directamente proporcional al cambio de las superficies polares y no polares accesibles al disolvente (agua) entre esos dos estados (ecuación 3; Murphy *et al.*, 1992; Murphy y Freire, 1992; Freire, 1995):

$$\Delta C_p = \Delta C_{p,ap}^\circ * \Delta A_{ap} + \Delta C_{p,pol}^\circ * \Delta A_{pol} \quad (3)$$

donde $\Delta C_{p,pol}^\circ$ y $\Delta C_{p,ap}^\circ$ son las contribuciones de los residuos polares y no polares a la capacidad calorífica, cuyos valores han sido obtenidos utilizando resultados de ΔC_p experimentales ($\Delta C_{p,ap}^\circ = 1.88 \pm 0.08 \text{ J K}^{-1} [\text{mol}\cdot\text{\AA}^2]^{-1}$ y $\Delta C_{p,pol}^\circ = -1.09 \pm 0.13 \text{ J K}^{-1} [\text{mol}\cdot\text{\AA}^2]^{-1}$), ΔA_{ap} y ΔA_{pol} son las diferencias entre las superficies polares y no polares accesibles al solvente de los dos estados calculadas a partir de su estructura.

El plegamiento en proteínas homólogas

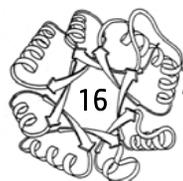
Las proteínas homólogas son aquellas que presentan la existencia de ancestros compartidos entre miembros de diferentes taxones. Existen diversos estudios cinéticos y al equilibrio para analizar las vías de plegamiento de proteínas homólogas. En todos ellos, aunque el nivel de profundidad y caracterización de los estudios es muy diferente (es decir, no todos se han realizado con las mismas técnicas ni estudiado los mismos aspectos fisicoquímicos), se han encontrado dos grandes grupos: a) las vías de plegamiento se encuentran conservadas evolutivamente, en proteínas pertenecientes a una misma familia (Gunasekaran *et al.*, 2001) y b) a pesar de conservar una topología (o dominio



estructural), la vía de plegamiento se encuentra diversificada y no se conserva evolutivamente (Zarrine-Afsar *et al.*, 2005).

Algunos ejemplos en donde las rutas de plegamiento siguen caminos similares para proteínas homólogas son:

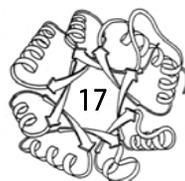
1. ***Ribonucleasa***. La vía de plegamiento se mantiene constante; al comparar proteínas de distintas especies, todas se pliegan mediante un mecanismo de dos estados, sin la presencia de intermediarios estables (Krebs *et al.*, 1983).
2. ***Proteína de unión a acetil-coA***. Al estudiar cuatro proteínas de unión a acetil-coA se encontró que sus mecanismos de plegamiento son parecidos entre ellas y que todas siguen un proceso de dos estados con solamente un paso en la conversión del estado nativo al desnaturalizado (Kragelund *et al.*, 1996).
3. ***Proteína de choque por frío***. Al comparar miembros cercanos evolutivamente, se observa que todas se pliegan por un mecanismo de dos estados sin la presencia de intermediarios estables. Incluso, al comparar proteínas derivadas de organismos mesófilos, termófilos e hipertermófilos, se encontró que las velocidades de plegamiento son similares y los estados de transición son estructuralmente parecidos en todas ellas. Por lo tanto, al parecer la vía de plegamiento en esta familia de proteínas ha sido conservada a través del tiempo (Perl *et al.*, 1998; Schindler *et al.*, 1999; Reid *et al.*, 2000).
4. ***Proteínas con topología α/β*** . Algunos miembros de bajo peso molecular (acilfosfatasa y ADA2h) mantienen una similitud en topología y en las estructuras del estado de transición de estas proteínas. Todos los miembros siguen una vía de plegamiento de dos estados poblados durante el proceso (Chiti *et al.*, 1999).



5. ***Dihidrofolato reductasa***. Al analizar proteínas de la misma familia procedentes de diferentes organismos, se observa que la vía de plegamiento es conservada, presentando un mecanismo complejo compuesto de vías paralelas de plegamiento. Dicha vía involucra al menos dos tipos de intermediarios cinéticos dentro de ella; éstos tienen las mismas características fisicoquímicas y responden a velocidades de formación semejantes. Al comparar el mecanismo de plegamiento con el del análogo p21 H-ras y encontrar que se pliega de la misma forma, sugiere que probablemente una misma vía de plegamiento ha sido conservada evolutivamente en diferentes miembros de proteínas con topología α/β (Wallace y Matthews, 2002).
6. ***Espectrina***. En homólogos de la familia de la espectrina diferentes dominios se pliegan de acuerdo al modelo de dos estados al equilibrio con un probable intermediario cinético de alta energía, el cual es semejante entre ellos (Scott *et al.*, 2004; Wensley *et al.*, 2009).

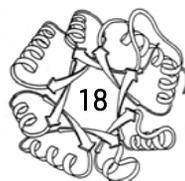
A continuación se muestran algunos ejemplos para el caso de proteínas homólogas que se pliegan por vías y mecanismos diferentes a pesar de conservar la misma topología entre sus diferentes miembros:

1. ***Proteína inmunitaria Im7/Im9***. Para dos proteínas inmunitarias homólogas (Im7 y Im9) se observó que estructuralmente el estado nativo no presenta diferencias significativas, pero que el plegamiento de ambas sigue diferentes mecanismos cinéticos. Por un lado, Im9 no presenta intermediarios estables, mientras que en Im7 un intermediario altamente compacto es poblado en la fase explosiva del plegamiento. Este fue el primer ejemplo de distintos mecanismos de plegamiento en familias de proteínas relacionadas evolutivamente (Ferguson *et al.*, 1999).
2. ***Proteína de unión a lípidos***. Al estudiar diferentes miembros de la familia de proteínas de unión a lípidos, con una identidad del 93% en secuencia de aminoácidos, se encontró que



sus vías de plegamiento son distintas. En una de ellas se observa la presencia de al menos dos intermediarios cinéticos que no presenta otra proteína de la familia, la cual tiene un intermediario al equilibrio (Banaszak *et al.*, 1994; Dalessio y Ropson, 2000; Ropson *et al.*, 2008).

3. **Proteína G/L.** A pesar de que la proteína L en su estado nativo es altamente simétrica, durante el proceso de plegamiento la simetría se pierde: en el estado de transición (E_T) una misma parte de la proteína se pliega y la otra permanece desestructurada hasta después del paso limitante del plegamiento (Kim *et al.*, 2000). Si se compara con otra proteína similar en estructura, como lo es la proteína G, se puede notar que a pesar de conservar una misma estructura tridimensional, el E_T de ambas proteínas es totalmente distinto, ya que en la proteína L una sección es formada en las etapas primarias del plegamiento (primera horquilla β), mientras que en la proteína G ésta se forma hasta al final de la vía (McCallister *et al.*, 2000).
4. **Globinas.** En la familia de globinas, las vías de plegamiento no están conservadas a pesar de tener una misma estructura terciaria (Lecomte *et al.*, 2005). Aunque presentan algunas similitudes, como la presencia de una fase explosiva al inicio del proceso, los intermediarios al equilibrio son diferentes. Para la apomioglobina, el intermediario estable se estructura de forma diferente en comparación con la apoleghemoglobina, donde esa misma parte de la proteína permanece desestructurada en el intermediario (secuencia de hélices distintas). Por lo anterior, las características fisicoquímicas de los intermediarios son diferentes en cada proteína (Nishimura *et al.*, 2000).
5. **Proteínas con puentes disulfuro.** En el caso de proteínas pequeñas ricas en puentes disulfuro se puede observar que existe una heterogeneidad grande en las vías de plegamiento, con la presencia o no de intermediarios; además se sabe que los intermediarios dependen

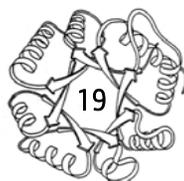


importantemente de la formación temprana de los puentes y la creación del andamiaje que guían el plegamiento global de la proteína (Arolas *et al.*, 2006).

6. **Lactoalbúmina/Lisozima.** Al comparar el plegamiento de dos proteínas homólogas, la lactoalbúmina y lisozima, se puede observar que sus mecanismos de plegamiento son distintos. A pesar de seguir una vía de plegamiento de tres estados, y tener un comportamiento cinético semejante, la estructuración de las proteínas es diferente tanto en el E_T como en el intermediario estable (Nakamura *et al.*, 2010).
7. **Triosafofato Isomerasa (TIM).** Para esta enzima se han observado diferencias considerables en las vías y mecanismos de plegamiento de acuerdo a la especie de donde provenga la enzima, por lo que se ha concluido que la vía de plegamiento no está conservada en esta familia de enzimas (Vázquez-Contreras *et al.*, 2004; Peimbert *et al.*, 2008a; Guzmán-Luna y Garza-Ramos, 2012). Debido a que la estructura terciaria y cuaternaria de la TIM en diferentes especies es esencialmente idéntica, esta enzima es un excelente modelo de estudio sobre la evolución de los mecanismos de plegamiento en proteínas homólogas (Chávez-Cárdenas *et al.*, 2002; Peimbert *et al.*, 2008b).

Ingeniería y diseño de proteínas

La ingeniería de proteínas consiste en el diseño de mutaciones en los sitios donde se cree que puedan tener un efecto sobre una función (Fersht, 1999). Este enfoque depende del conocimiento a priori de una gran cantidad de información estructural de la molécula y conocimientos de evolución molecular. En este enfoque se debe tomar en cuenta la relación existente entre la secuencia, estructura y función, ya que por una parte se tiene la información de la variabilidad de cada residuo y, por el otro, la posición tridimensional y la interacción con otros residuos en el espacio (Segovia y Peimbert, 2010). El conocimiento que utiliza el diseño e ingeniería de proteínas está basado en



análisis bioinformáticos de la secuencia de la proteína, reglas generales derivadas de caracterizar los efectos de las mutaciones sobre las propiedades de la macromolécula y por la implementación del modelado molecular, que permite predecir los efectos de las mutaciones sobre la estructura y estabilidad de las proteínas (Dalby, 2007).

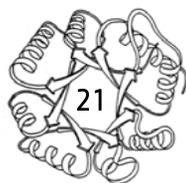
Uno de los grandes retos que enfrenta la comunidad científica es la demanda en el diseño de nuevas proteínas, principalmente de enzimas y anticuerpos (Vasserot *et al.*, 2003; Rubin-Pitel y Zhao, 2006). Esto provoca la necesidad de obtener proteínas que exhiban características mejoradas como: estabilidad, actividad catalítica específica, regulación celular, diferentes mecanismos de reacción, entre otras (Cohen *et al.*, 2001). De esta forma, el diseño de proteínas es un área encaminada a encontrar y mejorar proteínas con diferentes características a las que hoy conocemos. En este sentido, en los últimos años se han encontrado nuevos genes en diferentes metagenomas los cuales codifican para nuevas proteínas (Schloss y Handelsman, 2005; Tringe y Rubin, 2005; Lorenz y Eck, 2005). De forma sinérgica, la secuenciación masiva de nuevos organismos ha arrojado una amplia gama de proteínas nunca antes descritas (Shendure *et al.*, 2004; Venter *et al.*, 2004; Ferrer *et al.*, 2005). Pero una alternativa más reciente y con un futuro prometedor es el desarrollo de proteínas con nuevas propiedades diseñadas *de novo*, es decir, el diseño de secuencias de proteínas que no se presentan en la naturaleza.

Para una proteína de 200 residuos, existen 20^{200} secuencias de aminoácidos posibles, de las cuales el proceso evolutivo natural sólo ha muestreado un subconjunto infinitesimal. En esta perspectiva, la misma paradoja de Levinthal nos introduce a discutir la gran posibilidad de explorar otros espacios conformacionales que la naturaleza no ha explorado. La diferencia entre lo que es posible explorar en el espacio de secuencias y lo que ha sido explorado por la naturaleza -o al menos lo que hemos visto hasta ahora- ha sido denominada por diferentes nombres. El grupo de Luisi llama a estas proteínas como "nunca nacidas" (Chiarabelli *et al.*, 2009). Por otro lado, el grupo de Taylor llama a estas proteínas no exploradas como "la materia oscura del espacio proteico" (Taylor *et al.*,



2009); mientras que el grupo de David Baker las considera como parte de la "biología posterior a la evolución", que podría ser considerada análoga a la biología sintética (Bromley *et al.*, 2008) y el grupo de Christopher Woods lo nombra como "proteínas totalmente *de novo*" (Woolfson *et al.*, 2015). Sea cual sea el nombre, este espacio conformacional no explorado es un campo grande para los trabajos futuros del diseño e ingeniería de proteínas, los cuales deberán ir descubriendo los paisajes antes no conocidos.

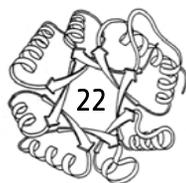
En este contexto, históricamente lo más fácil de asumir es que ya todas las secuencias de proteínas fueron muestreadas y la selección natural se quedó con lo más "óptimo", pero eso es un argumento de tipo adaptacionista. Un argumento adaptacionista interpreta a la selección natural como el agente causal más importante en toda forma, función o comportamiento orgánico, y le resta importancia a las limitaciones explicativas de esta fuerza evolutiva. Las hipótesis evolutivas propuestas desde una postura adaptacionista carecen de fundamento científico, ya que no reflejan lo que se sabe acerca del proceso evolutivo y de las diversas fuerzas que lo moldean, como lo son la deriva génica, el flujo génico y los constreñimientos ontogenéticos (Love, 2005). En el caso en cuestión, ¿las formas y funciones que no vemos en la naturaleza están ausentes porque no presentan una ventaja adaptativa o porque la naturaleza nunca ha llegado a ellas? Aquí es donde se vuelve importante el concepto de constreñimiento tomado desde la biología del desarrollo. Según el paradigma adaptacionista, la variación es completamente aleatoria e isotrópica, pero algo que saben muy bien los biólogos del desarrollo es que hay ciertas constricciones a la idea de variación, por lo que en realidad no es un proceso tan aleatorio (Minelli, 2009). El primer constreñimiento es algo que nosotros entendemos muy bien, los que nos dedicamos al mundo de las proteínas. Hay formas (secuencias) que simplemente son imposibles, porque la proteína se vuelve inestable, no adquiere el plegamiento correcto o física y químicamente no puede expresarse de esa forma; esto limita el espacio hipotético en que una secuencia puede variar. Lo segundo, son los llamados constreñimientos históricos y eso se relaciona con que sí existe una injerencia de las características fenotípicas del ancestro sobre las de los nuevos organismos (Minelli, 2009; Love, 2015). Lo anterior



es algo que en proteínas se puede observar fácilmente. ¿Cuántas secuencias de ADN y aminoácidos distintas no generan la misma topología y función (fenotipo)? Uno podría pensar que las topologías se conservan porque no hay otra estructura que lo haga tan bien (adaptacionismo); pero qué pasa si la naturaleza no sabe hacer otra cosa que no sea una topología específica, por ejemplo, un plegamiento tipo Rossmann o un barril TIM. Si la variación en secuencia puede ser tanta, por qué, hasta la fecha, no hemos encontrado algo radicalmente distinto en la naturaleza. Es así como se puede afirmar que los fenómenos biológicos no corresponden a una idea puramente adaptacionista y por esta razón el diseño de proteínas es un campo del cual se puede aprender mucho acerca de estas macromoléculas, tanto en estructura-función como en teorías evolutivas.

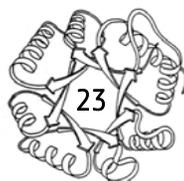
El enfoque actual que rige el diseño de proteínas se basa en la hipótesis de que éstas se pliegan en los estados de energía más bajos que son accesibles por sus secuencias de aminoácidos, como propuso Christian Anfinsen hace más de 50 años. Dado un método adecuadamente preciso para calcular la energía de una cadena polipeptídica al momento de plegarse, así como métodos para muestrear el espacio conformacional posible, debería ser posible diseñar secuencias que se plieguen en nuevas estructuras. Debido a esto, hoy en día existen dos desafíos en la aplicación de este enfoque: en primer lugar, la energía de un sistema no se puede calcular con gran exactitud, y en segundo lugar, el espacio de posibles estructuras y secuencias es muy grande y por lo tanto difícil de buscar exhaustivamente (Khoury *et al.*, 2014; Regan *et al.*, 2015; Woolfson *et al.*, 2015; Huang *et al.*, 2016b).

Uno de los principios que guían el plegamiento de las proteínas es el ocultamiento de los residuos hidrofóbicos en el núcleo de la proteína, lejos del disolvente. Para minimizar el tamaño de la cavidad que la proteína ocupa en el agua, y para maximizar las fuerzas de van der Waals, las cadenas laterales en el núcleo deben estar empaçadas muy estrechamente, pero sin sobreposiciones atómicas desfavorables energéticamente. Los grupos polares que interactúan con el disolvente en el estado desplegado que quedan ocultos en la parte interna de la proteína deben de formar puentes



de hidrógeno intra-proteína para compensar este efecto, de lo contrario el gran costo energético de separar el agua desfavorecería el plegamiento (Fleming y Rose, 2005). A partir de estas consideraciones, el diseño busca las características estructurales de las proteínas naturales: estructuras secundarias hélices α y hebras β , en las que los grupos carbonilo y amida polares del esqueleto polipéptido pueden formar puentes de hidrógeno y ensamblarse de tal manera que las cadenas laterales no polares se introduzcan en la molécula para formar núcleos densamente empaquetados. Además, el acomodo de las cadenas laterales es importante para los efectos tanto espaciales como fisicoquímicos.

Este panorama general del plegamiento de las proteínas se implementa en una función de energía que captura las interacciones de los átomos en la proteína y con el disolvente. Los principales factores que contribuyen a esta función energética son las fuerzas de van der Waals (que favorecen el empaquetamiento cercano), la repulsión estérica, las interacciones electrostáticas, los puentes de hidrógeno, la solvatación y las energías de torsión de la cadena principal y los enlaces de las cadenas laterales (Ponder *et al.*, 1987; Dahiyat y Mayo, 1996; Dahiyat y Mayo, 1997; Kuhlman y Baker, 2000). De esta forma, el desafío del diseño de proteínas puede ser englobado en 3 formas de abordarlo: predicción de estructura, diseño con la cadena principal fija y diseño *de novo*. De todas ellas, a diferencia de los problemas de predicción estructural y diseño con la cadena principal fija, en el problema del diseño *de novo*, tanto la secuencia como la estructura son desconocidas, incrementando la complejidad del método, aunque permitiendo explorar más sitios del espacio conformacional posible. Durante el proceso de diseño, se evalúa qué tanto los diseños se parecen a la estructura molde o a la arquitectura deseada en el diseño. Sea cual sea el caso, en todos los trabajos experimentales de diseño, lo que se desea demostrar es que la proteína diseñada computacionalmente es muy parecida a la estructura tridimensional determinada por algún método experimental (figura 3).



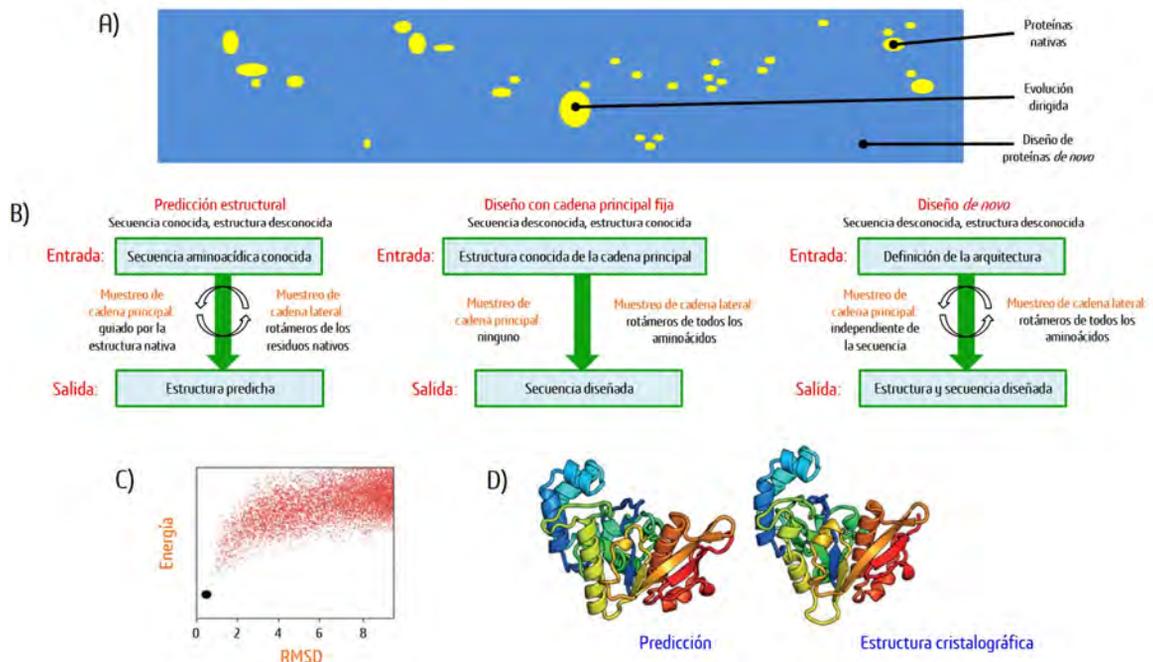


Figura 3. El panorama actual del diseño de proteínas. **A)** Un esquema del espacio explorado por las secuencias de proteínas. La evolución ha muestreado sólo una pequeña fracción del espacio total de la secuencia posible (azul), y el espacio conformacional explorado se agrupa en grupos compactos de topologías específicas (amarillo). La evolución dirigida está restringida a la región del espacio de secuencias que rodea a las proteínas nativas, mientras que el diseño *de novo* puede explorar, en teoría, todo el espacio. **B)** Los problemas de predicción estructural, el diseño con cadena principal fija y el diseño *de novo* son problemas globales de optimización con la misma función pero con diferentes grados de libertad. En la predicción estructural, la secuencia es fija y la estructura de la cadena principal es desconocida; en el diseño fijo, la secuencia es desconocida pero la estructura es fija; y en el diseño *de novo* nada es conocido. **C)** Ejemplo de un paisaje energético generado a partir de cálculos de predicción de estructura. Los puntos rojos representan las estructuras de más baja energía en cada trayectoria independiente, que se representan según la similitud con la estructura blanco (punto negro) a lo largo del eje x; la similitud estructural se mide por la desviación cuadrática media de la raíz (RMSD). **D)** Comparación de una proteína diseñada con su estructura tridimensional determinada experimentalmente (CASP11 T0806). Modificado de Huang *et al.*, 2016b.

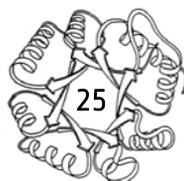
El diseño de proteínas *de novo* explora el espacio completo de las secuencias, guiado por los principios físicos que subyacen al plegamiento de las proteínas. La metodología computacional ha avanzado hasta el punto de que una amplia gama de estructuras puede ser diseñada desde cero con precisión a nivel atómico. Casi toda la ingeniería de proteínas, hasta ahora, ha implicado la modificación de las proteínas naturales; por lo que debe ser posible diseñar nuevas proteínas funcionales desde el principio para hacer frente a los desafíos actuales en biomedicina y

nanotecnología, es decir, diseñar proteínas *de novo* y justo éste es uno de los desafíos actuales en el campo de las proteínas (Huang *et al.*, 2016b).

Barriles (β/α)₈ o barriles TIM

En 1976, Levitt y Chothia, en un trabajo clásico, propusieron una clasificación de las estructuras de proteínas de acuerdo al tipo de composición de estructura secundaria en cada miembro; así surgieron cuatro clases: alfa, beta, alfa/beta y alfa + beta (Levitt y Chotia, 1976). Hasta la fecha, esta clasificación estructural de las proteínas ha prevalecido. Además, estos autores reconocieron que dentro de cada clase estructural existen varias topologías: distintos órdenes, direcciones y repeticiones de los elementos alfa o beta principales. Asimismo, con el cúmulo de información a través de los años se han diseñado representaciones diagramáticas para la topología de proteínas que se siguen utilizando hoy en día para subclasificarlas (Andreeva *et al.*, 2008). Estas topologías se pueden representar como patrones repetitivos acomodados espacialmente de forma específica y la comparación de estos patrones permite comparar gran cantidad de secuencia de proteínas y detectar semejanzas o diferencias entre los miembros, clasificándolos en super y subfamilias.

A mediados de 2017, la base de datos de proteínas (PDB) contiene más de 130 mil estructuras, de las cuales, los sistemas de clasificación sugieren que éstas son representadas por menos de 1400 diferentes topologías, dejando un mensaje claro: las proteínas naturales emplean un limitado conjunto de topologías tridimensionales. Dentro de las diferentes conformaciones tridimensionales que pueden tomar las secuencias de aminoácidos, la superfamilia de los barriles TIM o barriles (β/α)₈ es una de las más abundantes en la naturaleza, ya que aproximadamente el 10% de todas las proteínas conocidas adquieren esta estructura; incluso, esta proporción se ha mantenido a lo largo de los años, a pesar de descubrir más y más proteínas (Reardon y Farber, 1995; Sterner y Höcker, 2005; Caetano-Anollés *et al.*, 2007). De acuerdo a las bases de datos *Structural*



Classification of Proteins (SCOP), *Class (C), Architecture (A), Topology (T) and Homology (H)* (CATH) y *Families of Structurally Similar Proteins* (FSSP) (base de datos que clasifican proteínas de acuerdo a su arquitectura, topología y secuencia), la superfamilia de barriles TIM puede subclasificarse en 33, 29 y 24 subfamilias, respectivamente, indicando el gran contenido y variabilidad de estructuras pertenecientes a esta topología (Hadley y Jones, 1997). Los barriles TIM participan en una gran variedad de funciones celulares y reacciones metabólicas. La mayoría de los miembros son enzimas y existe representantes en cinco de las seis clases de enzimas definidas por la comisión internacional de enzimas (EC) (Wierenga, 2001; Nagano *et al.*, 2002).

El plegamiento tipo barril TIM consiste en un motivo β -asa- α -asa repetido ocho veces (figura 4), con las ocho hebras beta paralelas al interior del barril, cubiertas por ocho hélices alfa en el exterior (Wierenga, 2001; Höcker *et al.*, 2002; Anantharaman *et al.*, 2003). Los barriles TIM, sin importar el tipo de reacción que catalicen, siempre presentan su sitio activo en la cara C-terminal, y en muchos de ellos se ha observado un motivo de unión a fosfato en la hebra 7, lo anterior sugiere que este plegamiento tiene un origen monofilético (Copley y Bork, 2000).

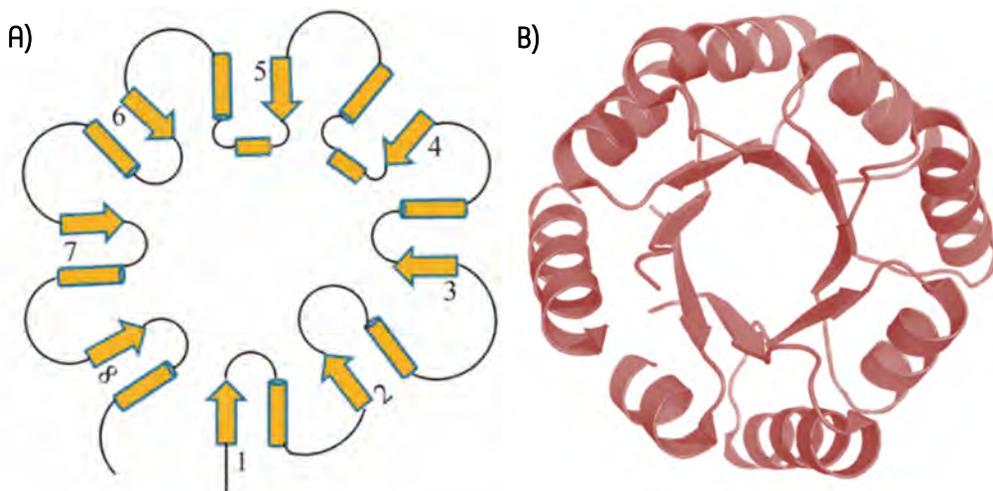
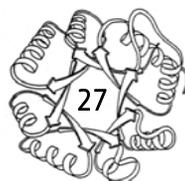


Figura 4. Topología general de un barril (β/α)₈. A) Esquema de la estructura secundaria; como flechas se señalan las hebras β y como cilindros las hélices α (modificado de Pardo-Ávila, 2011). B) Estructura tridimensional de un barril TIM (ReTIM18).

Debido a su enorme variedad en funciones, su origen evolutivo común y a su ubicuidad en distintos tipos de proteínas, el barril TIM representa una topología de gran interés fisicoquímico para su estudio. Además de la amplia distribución y gran variedad de reacciones enzimáticas que las proteínas con plegamiento de barril TIM son capaces de catalizar, este tipo de estructura representa un andamiaje muy estable para el diseño e ingeniería de proteínas, con numerosas aplicaciones en los campos de la biología sintética, diagnóstico e imagenología (Currin *et al.*, 2015).

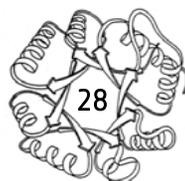


ANTECEDENTES ESPECÍFICOS

La Triosafosfato Isomerasa como modelo de estudio de un barril TIM natural

Triosafosfato Isomerasa: estructura y función

El miembro fundador de la superfamilia de los barriles TIM es la enzima Triosafostato Isomerasa (TIM, de la cual deriva el nombre de la superfamilia), una proteína con topología de barril (β/α)₈ de aproximadamente 250 aminoácidos localizada en citoplasma, cloroplastos y mitocondrias; participa en la gluconeogénesis, la síntesis de triglicéridos, el ciclo de Calvin y, principalmente, en la vía de la glucólisis. Mediante esta última vía metabólica anaerobia, la mayoría de las formas de vida degradan la glucosa en dos moléculas de piruvato, con la producción de dos moléculas de ATP a través de 10 reacciones enzimáticas. En el cuarto paso, la hexosa difosforilada fructuosa-1,6-bisfosfato se divide por la enzima aldolasa en dos triosas: la dihidroxiacetona fosfato (DHAP) y el gliceraldehído-3-fostato (G3P). Solamente un producto de la aldolasa continúa la vía de la glucólisis, el G3P. (Nelson y Cox, 2000; Voet *et al.*, 2006). La TIM cataliza la interconversión entre DHAP y G3P por una reacción de isomerización, con un intermediario enediol o enediolato (figura 5 y 8) y sin la utilización de cofactores para la catálisis (Rose y O'Connell, 1961; Knowles y Albergy, 1977; Knowles, 1991). Algunos organismos anaeróbicos son estrictamente dependientes de la actividad enzimática de la TIM, razón por la cual, esta enzima se ha utilizado como blanco en el diseño de drogas (Pérez-Montfort *et al.*, 1999; Reyes-Vivas *et al.*, 2001; Zomosa-Signoret *et al.*, 2003; López-Velázquez *et al.*, 2004; Enríquez-Flores *et al.*, 2011; Moraes *et al.*, 2011) y vacunas dirigidas en contra de endoparásitos (Zhu *et al.*, 2006).



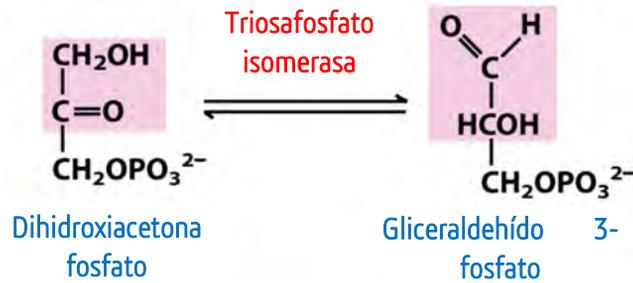


Figura 5. Reacción catalizada por la Triosafosfato Isomerasa. Modificada de: Nelson y Cox, 2000.

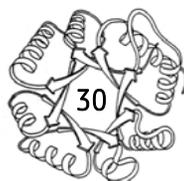
El Comité de Nomenclatura de la Unión de Bioquímica y Biología Molecular (NC-IUBMB, por sus siglas en inglés), clasifica a la TIM con el número 5.3.1.1 (5: Isomerasas; 3: Oxidorreductasas intramoleculares; 1: Interconvertidores de aldosas y cetosas; 1: Triosafosfato Isomerasa) de acuerdo a la reacción que cataliza (NC-IUBMB, 1992).

La enzima posee ciertas propiedades catalíticas interesantes: la transformación de DHAP en G3P realizada por la TIM es una reacción muy rápida, entre 10^8 y 10^9 veces más rápida que en ausencia de la enzima (Knowles y Albery, 1977; Nelson y Cox, 2000; Mathews *et al.*, 2012). Además, la TIM es un catalizador perfecto, lo cual significa que la velocidad de la reacción entre la enzima y el sustrato está controlada sólo por difusión. La formación del producto ocurre tan rápido como la enzima y el sustrato se encuentran en la solución, llegando la eficiencia catalítica (k_{cat}/K_m^{-1}) cercano de la constante de difusión, es decir, entre 10^8 y $10^9 \text{ M}^{-1} \text{ s}^{-1}$ (Knowles y Albery, 1977). Los parámetros cinéticos para las TIMs estudiadas al momento se presentan en la tabla 1. Se pueden notar las semejanzas en la afinidad por el sustrato y en la eficiencia catalítica de todas las enzimas, mientras que la constante de inhibición de sus inhibidores competitivos, fosfoglicolohidroxamato (PGH) y ácido 2-fosfoglicólico (PGA) es muy diferente entre especies.



Tabla 1. Parámetros cinéticos para las TIMs silvestres.

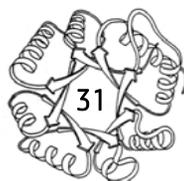
TIM	K_i PGA (mM)	K_m G3P (mM)	k_{cat} (min^{-1})	Eficiencia Catalítica ($\text{M}^{-1} \text{s}^{-1}$) [†]	Referencia
<i>Ec</i> TIM	6	1.03	5.4×10^5	2.18×10^8	Alvarez <i>et al.</i> , 1998
<i>Et</i> TIM	ND	0.61	2.4×10^5	1.64×10^8	Landa <i>et al.</i> , 1997
<i>Gg</i> TIM	0.02	0.47	2.6×10^5	2.27×10^8	Plaut y Knowles, 1972; Putman <i>et al.</i> , 1972
<i>Gl</i> TIM	0.043	0.53	2.9×10^5	2.28×10^8	López-Velázquez <i>et al.</i> , 2004; Reyes-Vivas <i>et al.</i> , 2007
<i>Hp</i> TIM	ND	3.46	8.8×10^4	1.10×10^7	Chu <i>et al.</i> , 2008
<i>Hs</i> TIM	0.0074	0.49	2.7×10^5	2.30×10^8	Mande <i>et al.</i> , 1994; Mainfroid, <i>et al.</i> , 1996
<i>Ld</i> TIM	ND	0.33	2.8×10^5	3.51×10^8	Kumar <i>et al.</i> , 2012
<i>Lm</i> TIM	ND	0.30	2.5×10^5	3.47×10^8	Kohl <i>et al.</i> , 1994
<i>Mj</i> TIM	ND	0.19	5.4×10^4	1.4×10^8	Muñiz-Trejo, 2017
<i>Mm</i> TIM	89	1.90	4.2×10^5	9.2×10^7	Alvarez <i>et al.</i> , 1998
<i>Mt</i> TIM	ND	2.10	4.1×10^6	8.13×10^8	Mathur <i>et al.</i> , 2006; Abendroth <i>et al.</i> , 2009
<i>Oc</i> TIM	14	0.32	5.2×10^4	6.77×10^7	Krietsch <i>et al.</i> , 1970; Knobeloch <i>et al.</i> , 2010
<i>PfT</i> TIM	ND	0.58	1.1×10^5	7.54×10^7	Joubert <i>et al.</i> , 2001
<i>Pfu</i> TIM	ND	1.22	9.3×10^7	1.27×10^9	Sharma y Guptasama, 2015
<i>Rm</i> TIM	ND	0.48	1.7×10^5	1.41×10^8	Moraes <i>et al.</i> , 2011
<i>Scr</i> TIM	ND	0.51	1.2×10^5	9.33×10^7	Jiménez <i>et al.</i> , 2003
<i>Sj</i> TIM	ND	0.40	2.1×10^5	2.17×10^8	Sun <i>et al.</i> , 1999
<i>Tb</i> TIM	0.024	0.25	3.7×10^5	6.17×10^8	Lambeir <i>et al.</i> , 1987; Hernández-Alcántara <i>et al.</i> , 2002



TIM	K_i PGA (mM)	K_m G3P (mM)	k_{cat} (min^{-1})	Eficiencia Catalítica ($\text{M}^{-1} \text{s}^{-1}$) [†]	Referencia
T_c TIM	ND	0.48	2.6×10^5	2.26×10^8	Reyes-Vivas <i>et al.</i> , 2001
T_m TIM	ND	2.60	1.1×10^3	1.76×10^5	Beaucamp <i>et al.</i> , 1997 [‡]
T_{mo} TIM	27	0.84	2.6×10^5	1.28×10^8	Knobeloch <i>et al.</i> , 2010
T_o TIM	ND	1.67	8.6×10^7	0.86×10^9	Sharma y Guptasama, 2015
T_s TIM	ND	0.37	7.5×10^4	8.49×10^7	Jiménez <i>et al.</i> , 2000; Jiménez <i>et al.</i> , 2003
T_{tx} TIM	ND	0.20	2.2×10^5	4.62×10^8	Walden <i>et al.</i> , 2004
T_v TIM	ND	0.21	8.2×10^4	1.63×10^8	Figuroa-Angulo <i>et al.</i> , 2012
yt TIM	0.03	1.27	1.0×10^6	3.28×10^8	Krietsch <i>et al.</i> , 1970; Lambeir <i>et al.</i> , 2000

PGA: ácido 2-fosfoglicólico, **G3P:** gliceraldehído 3-fosfato, K_m : constante de Michaelis-Menten, k_{cat} : constante catalítica, **ND:** no determinado. †: Realizando la corrección en la K_m al considerar que sólo el 4% del G3P en solución tiene el sustrato aldehído no hidratado, única forma que la TIM es capaz de catalizar (relación 1:25). ‡: K_m determinada para la DHAP, no se tienen datos reportados para el G3P. **Ec:** *Escherichia coli*, **Eh:** *Entamoeba histolytica*, **Gg:** *Gallus*, **Gl:** *Giardia lamblia*, **Hp:** *Helicobacter pylori*, **Hs:** *Homo sapiens*, **Ld:** *Leishmania donovani*, **Lm:** *Leishmania mexicana*, **Mj:** *Methanocaldococcus jannaschii*, **Mm:** *Moritella marina* (antes *Vibrio marinus*), **Mt:** *Mycobacterium tuberculosis*, **Oc:** *Oryctolagus cuniculus*, **Pf:** *Plasmodium falciparum*, **Pfu:** *Pyrococcus furiosus*, **Rm:** *Rhipicephalus microplus*, **Scr:** *Sus scrofa*, **Sj:** *Schistosoma japonicum*, **Tb:** *Trypanosoma brucei*, **Tc:** *Trypanosoma cruzi*, **Tm:** *Thermotoga maritima*, **Tmo:** *Tenebrio molitor*, **To:** *Thermococcus onnurineus*, **Ts:** *Taenia solium*, **Ttx:** *Thermoproteus tenax*, **Tv:** *Trichomonas vaginalis*, **yt:** *Saccharomyces cerevisiae*.

En la estructura de la TIM, el barril β está compuesto por ocho hebras β paralelas, las cuales forman una elipse cuyos semiejes miden 12 Å y 18 Å. El barril está situado en el interior de la proteína, las hebras están conectadas por asas a las hélices α que rodean a las hebras β (Reardon y Faber, 1995). Los aminoácidos que forman el barril β son predominantemente hidrofóbicos. Muchas de las cadenas laterales de estos aminoácidos llenan el interior del barril y forman parte del centro hidrofóbico de la enzima. El sitio activo se encuentra ubicado en el extremo carboxilo terminal de la hoja beta cilíndrica (figura 6) (Lolis y Petsko, 1990; Wierenga *et al.*, 1991; Wierenga *et al.*, 1992;).



Al comparar la secuencia de las TIMs con estructura conocida, la mayor identidad se encuentra entre *HsTIM* vs. *OcTIM* (98.5%), mientras que la menor identidad (12%) se observa para la TIM de *Pyrococcus woesei* (*PwTIM*) vs. *PfTIM*. Sin embargo, al comparar las estructuras tridimensionales, se encuentra que la topología se mantiene en todas, es decir, la arquitectura de barril (β/α)₈ ha sido conservada evolutivamente independientemente de la variación en la secuencia aminoacídica.

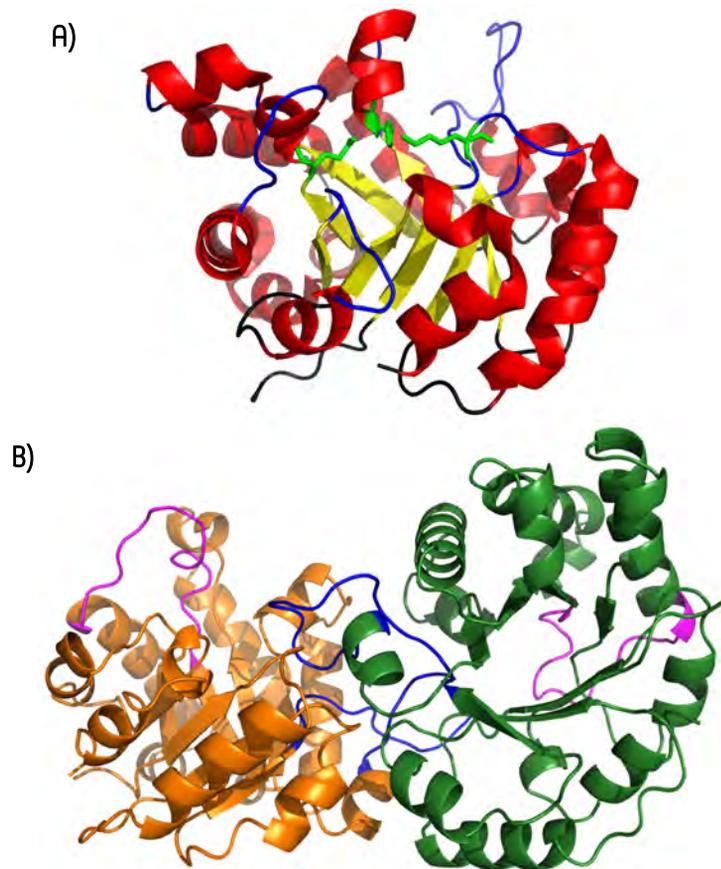


Figura 6. Estructura tridimensional de la TIM. **A)** Estructura terciaria de una subunidad de la TIM; en rojo se muestran las hélices α , en amarillo las hebras β , en negro las asas $\alpha\beta$ y en azul las asas $\beta\alpha$, en verde se muestran los residuos catalíticos: Lys12, His95 y Glu165 (numeración de la TIM de *Saccharomyces cerevisiae*). **B)** Dímero activo presente en la TIM; en naranja y verde se indican los monómeros, en magenta se señalan las asas $\beta\alpha$ -6 y en azul se indican las asas $\beta\alpha$ -3 (PDB ID: 4Y8F).

Hasta el momento (septiembre de 2017), se han reportado un total de 195 estructuras cristalográficas de la TIM (entre TIMs silvestres y mutantes), procedentes de 39 especies taxonómicas distintas y ubicadas en los 3 dominios de la vida (156 para Eukarya, 35 en Bacteria y 4 en Archaea). Las especies de las cuales se ha determinado al menos una estructura tridimensional de la TIM se resumen en la tabla 2 y se muestra el número total de estructuras determinadas para cada especie (sin incluir las estructuras tridimensionales determinadas en este trabajo).

Además del asa $\beta\alpha$ -6 que participa en la catálisis, existe el asa $\beta\alpha$ -3 que participa de manera indirecta en la estabilización del sitio activo y la intercara del dímero. En el dímero, la mayoría de los contactos de intercara entre monómeros son mediados por el asa 3, también referida como el asa de intercara (residuos 65 a 79) (figura 9), que se extiende desde de un monómero hasta una cavidad cercana al centro catalítico de la segunda subunidad (entre las asas $\beta\alpha$ -1 y $\beta\alpha$ -4) (Lolis y Petsko, 1990). Algunos residuos del asa 3 interactúan a través de contactos de van der Waals con los residuos del sitio activo de la otra subunidad. La tapa del asa 3 (residuos 74 a 77) forma una región hidrofóbica alrededor del residuo Leu13 de la subunidad adyacente (Maithal *et al.*, 2002). Las TIMs silvestres sólo son activas como dímeros (unas pocas como tetrámeros), los residuos de la intercara son cruciales para mantener la integridad de las interacciones entre ambas subunidades (figura 6) (Zabori *et al.*, 1980; Casal *et al.*, 1987; Garza-Ramos, *et al.*, 1992; Cabrera *et al.*, 2008).

En la naturaleza, esta enzima sólo es activa en su forma oligomérica, y la estructura del dímero nativo (o del tetrámero en algunas especies) sugiere que el proceso de oligomerización es importante para estabilizar la geometría de algunos residuos del sitio activo (Wierenga *et al.*, 1992; Nájera *et al.*, 2003). El sitio activo está localizado en la cara C-terminal del barril y dentro de la catálisis participan 3 aminoácidos principales, los cuales están conservados en todas las secuencias de la TIM reportadas hasta el momento: lisina 12 (asa $\beta\alpha$ -1), histidina 95 (asa $\beta\alpha$ -4) y glutámico 165 (asa $\beta\alpha$ -6) (numeración de la TIM de *Saccharomyces cerevisiae*, yfTIM, figura 7) (Berg *et al.*, 2002).

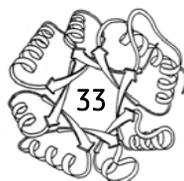
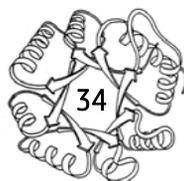


Tabla 2. Especies para las cuales se ha determinado la estructura cristalográfica de la TIM.

Código PDB	Especie	Resolución (Å)	Año [†]	Referencia de la primera estructura	Número de estructuras
1TIM	<i>Gallus gallus</i>	2.50	1976	Banner <i>et al.</i> , 1976	16
1YPI	<i>Saccharomyces cerevisiae</i>	1.90	1990	Lolis y Petsko, 1990	7
3TIM	<i>Trypanosoma brucei</i>	2.80	1991	Wierenga <i>et al.</i> , 1991	43
1TRE	<i>Escherichia coli</i>	2.60	1993	Noble <i>et al.</i> , 1993	11
1HTI	<i>Homo sapiens</i>	2.80	1994	Mande <i>et al.</i> , 1994	11
1BTM	<i>Geobacillus stearothermophilus</i>	2.80	1995	Delboni <i>et al.</i> , 1995	2
1AMK	<i>Leishmania mexicana</i>	1.83	1999	Williams <i>et al.</i> , 1999	8
1YDV	<i>Plasmodium falciparum</i>	2.20	1997	Velanker <i>et al.</i> , 1997	30
1TCD	<i>Trypanosoma cruzi</i>	1.83	1998	Maldonado <i>et al.</i> , 1998	7
1AW2	<i>Moritella marina</i>	2.65	1998	Alvarez <i>et al.</i> , 1998	2
1B9B	<i>Thermotoga maritima</i>	2.85	1999	Maes <i>et al.</i> , 1999	1
1HG3	<i>Pyrococcus woesei</i>	2.70	2001	Walden <i>et al.</i> , 2001	1
1M6J	<i>Entamoeba histolytica</i>	1.50	2002	Rodríguez-Romero <i>et al.</i> , 2002	1
1M00	<i>Caenorhabditis elegans</i>	1.70	2003	Symersky <i>et al.</i> , 2003	1
1R2T	<i>Oryctolagus cuniculus</i>	2.25	2003	Aparicio <i>et al.</i> , 2003	4
1W0M	<i>Thermoproteus tenax</i>	2.50	2004	Walden <i>et al.</i> , 2004	1
1YYA	<i>Thermus thermophilus</i>	1.60	2005	Yamamoto y Kunishima, 2005	1
2I9E	<i>Tenebrio molitor</i>	2.00	2006	Knobeloch <i>et al.</i> , 2010	1
2H6R	<i>Methanocaldococcus jannaschii</i>	2.30	2007	Gayathri <i>et al.</i> , 2007	1
2DP3	<i>Giardia lamblia</i>	2.10	2007	Reyes-Vivas <i>et al.</i> , 2007	5
2JGQ	<i>Helicobacter pylori</i>	2.30	2008	Chu <i>et al.</i> , 2008	1
3GVG	<i>Mycobacterium tuberculosis</i>	2.55	2009	Abendroth <i>et al.</i> , 2009	3
3KRS	<i>Cryptosporidium parvum</i>	1.55	2011	Nguyen <i>et al.</i> , 2011	1
3KXQ	<i>Bartonella henselae</i>	1.60	2009	Abendroth <i>et al.</i> , 2009	1
3M9Y	<i>Staphylococcus aureus</i>	1.90	2011	Mukherjee <i>et al.</i> , 2011	7
3S6D	<i>Coccidioides immitis</i>	2.20	2011	Edwards <i>et al.</i> , 2011	1
3TH6	<i>Rhipicephalus microplus</i>	2.40	2011	Moraes <i>et al.</i> , 2011	1
3QST	<i>Trichomonas vaginalis</i>	1.75	2012	Salgado-Lugo <i>et al.</i> , 2012	11



Código PDB	Especie	Resolución (Å)	Año [‡]	Referencia de la primera estructura	Número de estructuras
4G1K	<i>Burkholderia thailandensis</i>	2.35	2012	Baugh <i>et al.</i> , 2013	1
4GNJ	<i>Leishmania siamensis</i>	1.93	2013	Kuaprasert <i>et al.</i> , 2013	2
4NVT	<i>Brucella melitensis</i>	2.10	2013	Lukacs <i>et al.</i> , 2013	1
4MKN	<i>Chlamydomonas reinhardtii</i>	1.10	2014	Zaffagnini <i>et al.</i> , 2014	1
4OHQ	<i>Arabidopsis thaliana</i>	2.15	2015	Lopez-Castillo <i>et al.</i> , 2016	2
5CSR	<i>Thermoplasma acidophilum</i>	1.94	2015	Park <i>et al.</i> , 2015	2
4X22	<i>Leptospira interrogans</i>	2.08	2016	Pareek <i>et al.</i> , 2016	2
5IBX	<i>Streptococcus pneumoniae</i>	1.65	2016	Minasov <i>et al.</i> , 2016	1
5EYW	<i>Litopenaeus vannamei</i>	1.70	2016	Lopez-Zavala <i>et al.</i> , 2016	1
5UJW	<i>Francisella tularensis</i>	2.65	2017	Chang <i>et al.</i> , 2017	1
5UPR	<i>Toxoplasma gondii</i>	2.00	2017	Filippova <i>et al.</i> , 2017	1

[‡] Se indica el año de la primera estructura depositada para cada especie.

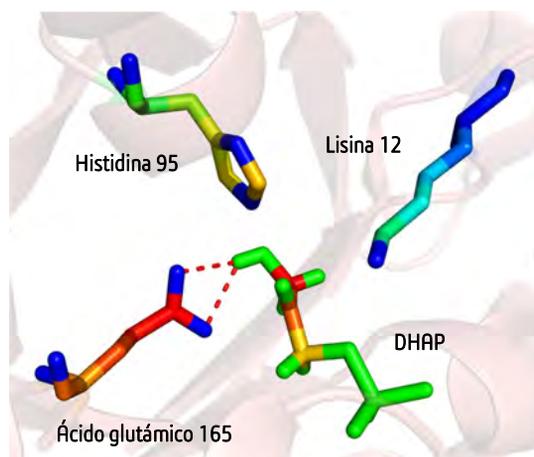
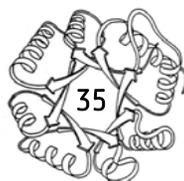


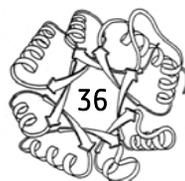
Figura 7. Triada catalítica de la Triosafosfato Isomerasa. Se indican las interacciones que estabilizan al sustrato catalizable (en este caso, DHAP). La numeración está basada en la TIM de *Saccharomyces cerevisiae*. (PDB ID: 1NEY).

El mecanismo de acción de la TIM se indica en la figura 8. Consideraciones químicas sugieren que la conversión del DHAP al intermediario enediol es catalizada por una base general, la cual sustrae un protón del C1 de la DHAP y por un ácido general el cual protona su átomo de oxígeno



carbonílico (Voet *et al.*, 2006). La cadena lateral del Glu165 se encuentra situada idealmente para sustraer el protón del C1 del DHAP, mientras que la cadena lateral de His95 está posicionada apropiadamente para protonar el oxígeno carbonílico del DHAP (Komives *et al.*, 1991; Wierenga *et al.*, 1992). La cadena lateral de la Lys12, cargada positivamente, estabiliza electrostáticamente el E_T que se encuentra negativamente cargado y, además, forma un enlace salino con la cadena lateral del carboxilato de Glu97; también se ha visto que participa en la estabilidad de la enzima (Joseph-McCarthy *et al.*, 1994; Lodi *et al.*, 1994; González-Mondragón *et al.*, 2004). Asimismo, en los pasos iniciales del mecanismo de reacción, sucede una estabilización del fosfato del sustrato por puentes de hidrógeno con la cadena principal de la proteína y la exclusión del solvente. Este último proceso es de vital importancia, ya que en solución, el intermediario enediol se rompe rápidamente con la subsecuente eliminación del fosfato del C3 formando un compuesto altamente tóxico para la célula, el metilglioxal. Para realizar la exclusión del solvente, la TIM posee un asa flexible que es el asa 6. Esta asa, a través de un movimiento de aproximadamente 7 Å, funciona como una tapa que oculta el sitio activo, permitiendo el posicionamiento del Glu165 a una distancia adecuada del sustrato para permitir su catálisis (figura 8). Esta asa protege el sitio activo del contacto con el medio acuoso y previene la formación de metilglioxal, ya que tiene al grupo fosfato del intermediario en una posición que desfavorece la reacción de eliminación de éste (Lolis y Petsko, 1990; Wierenga *et al.*, 1991); se ha observado que la ausencia de dicha asa genera la liberación temprana del intermediario y la transformación del sustrato a metilglioxal (Pompliano *et al.*, 1990).

El perfil energético calculado de todo el proceso catalítico en dirección DHAP \rightarrow G3P, considerando la formación del complejo TIM:DHAP, TIM-enediol, TIM:G3P y la liberación del producto final, se muestra en la figura 10 (Albery y Knowles, 1976).



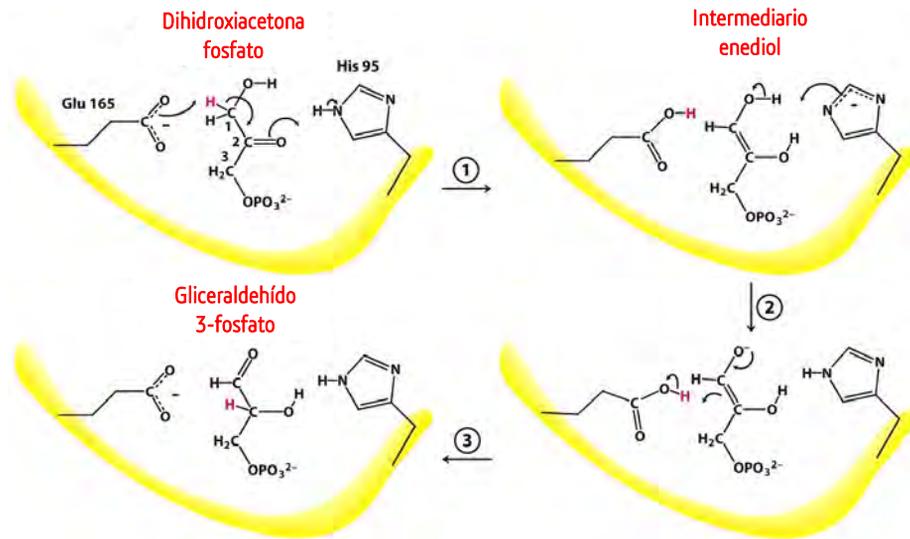


Figura 8. Mecanismo catalítico de la Triosafosfato Isomerasa. Modificado de: Berg *et al.*, 2002.

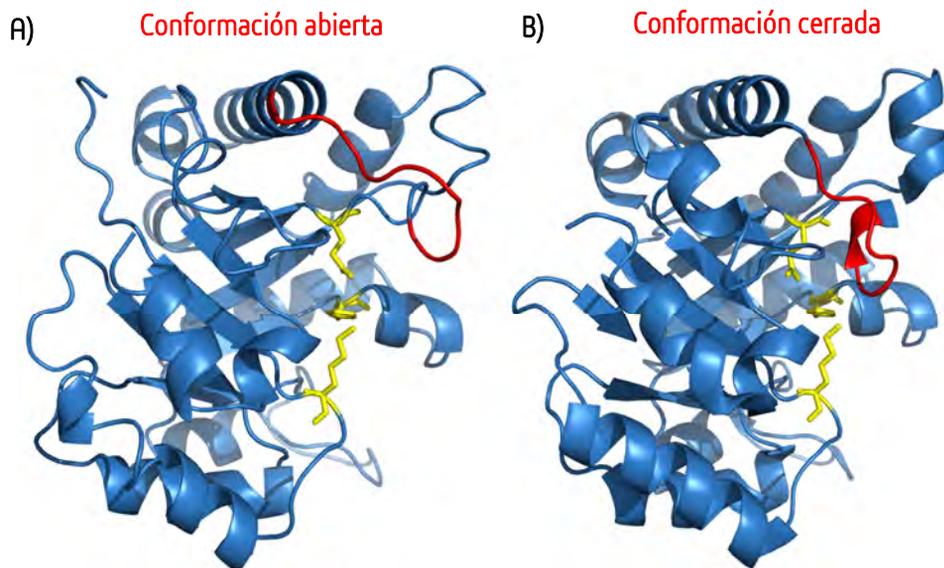


Figura 9. Conformaciones abierta y cerrada de la TIM. El asa 6 se encuentra en rojo y los aminoácidos catalíticos se señalan en amarillo, nótese el cambio de la asa 6 de la conformación abierta (A) a la conformación cerrada (B) (PDB ID conformación abierta: 4Y8F, PDB ID conformación cerrada: 2NEY).

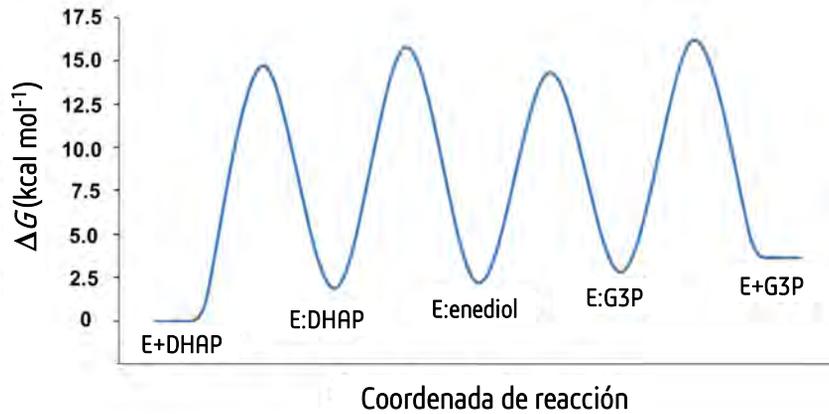


Figura 10. Cambios energéticos en cada paso de la reacción catalizada por la Triosafosfato Isomerasa. Se muestra la reacción en el sentido DHAP a G3P y se incluyen los valores de energía libre reportados para el intermediario enediol y los estados de transición (modificado de: Alberly y Knowles, 1976).

Triosafosfato Isomerasa: plegamiento y estabilidad

Se ha propuesto que a estructuras semejantes corresponden vías de plegamiento semejantes (Clarke *et al.*, 1999). Las TIMs presentan un alto grado de conservación estructural, sin embargo, existen numerosos estudios en los que se han reportado distintas vías de desnaturalización para la TIM de diferentes especies, la mayoría de ellos utilizando agentes caotrópicos desnaturalizantes como la urea o GdnHCl. Así, se ha visto que la vía de plegamiento no está conservada evolutivamente, presentando una gran diversidad en las vías de plegamiento de la TIM, desde procesos de dos estados para *GsTIM*, *OcTIM*, *GgTIM*, *LmTIM* y *HsTIM*; tres estados con un intermediario monomérico para la de *ytTIM* y *EhTIM*; tres estados con la presencia de estados agregados para *TmTIM*; y cuatro estados para *TbTIM*, *TcTIM* y *PfTIM* (tabla 3).

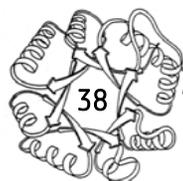
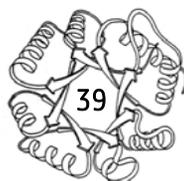


Tabla 3. Vías de plegamiento reportadas para la TIM de diferentes especies.

TIM	Vía de desplegamiento/plegamiento	Agente desnaturizante (Técnica)	Referencia
<i>Gs</i> TIM	$N_2 \rightarrow 2D$	GdnHCl (IF)	Mainfroid <i>et al.</i> , 1996a
<i>yt</i> TIM	$N_2 \rightleftharpoons 2D$	Temperatura (IF, DC, DSC)	Benítez-Cardoza <i>et al.</i> , 2001
<i>Oc</i> TIM	$N_2 \rightleftharpoons 2D$	GdnHCl (IF)	Rietveld y Ferreira, 1996
<i>Gg</i> TIM	$N_2 \rightarrow 2D$	Temperatura (IF, DC, DSC)	Shi <i>et al.</i> , 2008
<i>Lm</i> TIM	$N_2 \rightleftharpoons 2D$	GdnHCl (IF, DC)	Lambeir <i>et al.</i> , 2000
<i>Hs</i> TIM	$N_2 \rightleftharpoons 2D$	Urea (IF, DC)	Mainfroid <i>et al.</i> , 1996a
<i>yt</i> TIM	$N_2 \rightleftharpoons 2D' \rightleftharpoons 2D$	GdnHCl (IF, DC)	Morgan <i>et al.</i> , 2000; Nájera <i>et al.</i> , 2003
<i>Et</i> TIM	$N_2 \rightleftharpoons 2D' \rightleftharpoons 2D$	GdnHCl (IF)	Vázquez-Pérez y Fernández-Velasco, 2007
<i>Tm</i> TIM	$N_2 \rightarrow A \rightarrow 2D$	GdnHCl (IF)	Beaucamp <i>et al.</i> , 1997
<i>Pf</i> TIM	$N_2 \rightarrow N_2^* \rightarrow A \rightarrow 2D$	GdnHCl (IF, DC)	Gokhale <i>et al.</i> , 1999
<i>Tb</i> TIM	$\begin{array}{c} A \quad A \\ \uparrow \quad \uparrow \\ N_2 \rightleftharpoons N_2^* \rightleftharpoons 2D' \rightleftharpoons 2D \end{array}$	GdnHCl (IF, DC)	Chávez-Cárdenas <i>et al.</i> , 2002; Guzmán-Luna y Garza-Ramos, 2012
<i>Tc</i> TIM	$N_2 \rightleftharpoons N_2^* \rightleftharpoons 2D' \rightleftharpoons 2D$	GdnHCl (IF, DC)	Chávez-Cárdenas <i>et al.</i> , 2005
<i>Tv</i> TIM	$N_2 \rightleftharpoons 2D' \rightleftharpoons 2D$	GdnHCl (IF)	Lara-González <i>et al.</i> , 2014

N_2 : dímero nativo, N_2^* : intermediario dimérico, D' : intermediario monomérico, D : estado desnaturizado, A : estados agregados. **GdnHCl**: hidrocloreto de guanidino, **IF**: Fluorescencia Intrínseca, **DSC**: Calorimetría Diferencial de Barrido, **DC**: Dicroísmo Circular. **Gs**: *Geobacillus stearothermophilus*, **Et**: *Entamoeba histolytica*, **Gg**: *Gallus*, **Hs**: *Homo sapiens*, **Lm**: *Leishmania mexicana*, **Oc**: *Oryctolagus cuniculus*, **Pf**: *Plasmodium falciparum*, **Tb**: *Trypanosoma brucei*, **Tc**: *Trypanosoma cruzi*, **Tm**: *Thermotoga marítima*, **Tv**: *Trichomonas vaginalis*, **yt**: *Saccharomyces cerevisiae*.

La relación entre la estructura y la estabilidad no ha sido completamente explicada. El análisis de la estructura cristalográfica muestra que los contactos intercatenarios representan tan sólo el 5% del cambio total del área accesible al solvente (ASA). Sin embargo, los intermediarios monoméricos son poco estables y sus propiedades estructurales y parámetros termodinámicos son

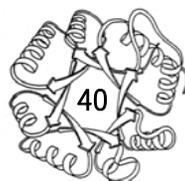


poco conocidas. Igualmente, se ha determinado que el ΔG de disociación (ΔG_{dis}) es responsable de al menos el 80% del ΔG total de estabilidad (Nájera *et al.*, 2003), por lo cual existe una paradoja importante entre estabilidad de los monómeros y la del dímero.

La primera etapa del plegamiento de la TIM involucra la formación de intermediarios monoméricos. Se han reportado este tipo de intermediarios para la desnaturalización con GdnHCl de $TcTIM$ (Chávez-Cárdenas *et al.*, 2005), $TbTIM$ (Chávez-Cárdenas *et al.*, 2002; Guzmán-Luna y Garza-Ramos, 2012), $yfTIM$ (Morgan *et al.*, 2000; Nájera *et al.*, 2003), $EhTIM$ (Vázquez-Pérez y Fernández-Velasco, 2007) y $TvTIM$ (Lara-González *et al.*, 2014). Algunas de las características comunes a todos ellos son que conservan estructura residual, son inactivos y expandidos. Posterior a la formación de los monómeros ocurre la asociación y los rearrreglos estructurales necesarios para la consolidación del estado nativo. La presencia de intermediarios diméricos ha sido reportada para la desnaturalización de $PfTIM$ (Gokhale *et al.*, 1999), $TbTIM$ (Chávez-Cárdenas *et al.*, 2002; Guzmán-Luna y Garza-Ramos, 2012) y $TcTIM$ (Chávez-Cárdenas *et al.*, 2005).

Así mismo, estudios realizados a partir de monómeros de TIM obtenidos por derivatización química, el análisis de intermediarios en la vía de plegamiento y el estudio de mutaciones que impiden la dimerización, han mostrado que los monómeros son marginalmente estables y que aproximadamente el 70% del ΔG de plegamiento está dado por la asociación de los monómeros. Las razones por las cuales la TIM es un oligómero obligado, no están totalmente entendidas (Peimbert *et al.*, 2008b). Existen dos hipótesis para explicar este comportamiento, la primera sugiere que la dimerización optimiza la geometría del sitio activo, ya que los contactos entre las subunidades disminuyen la flexibilidad de las asas catalíticas (Schliebs *et al.*, 1997). La segunda hipótesis presupone que la asociación genera una estabilidad conformacional máxima y que los monómeros son termodinámicamente inestables (Mainfroid *et al.*, 1996a).

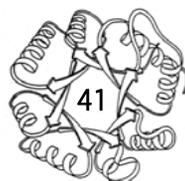
Aunado a lo anterior, el desplegamiento térmico de la TIM, principalmente de organismos eucariontes, ha sido estudiado extensivamente. En las TIMs reportadas a la fecha, el desplegamiento



térmico es un proceso irreversible, fuera del equilibrio o ambos (con presencia de histéresis: no existe coincidencia entre la rampa de desplegamiento y la de replegamiento), es decir, un proceso controlado cinéticamente, lo cual ha impedido una caracterización energético-estructural detallada. Debido a la histéresis y al alto grado de irreversibilidad del proceso, el uso de la temperatura como agente desnaturalizante se ha limitado más bien a la estimación de la T_m de algunas TIMs y sus mutantes (Mainfroid *et al.*, 1996b; Schliebs *et al.*, 1996; Beaucamp *et al.*, 1997; Schliebs *et al.*, 1997; Alvarez *et al.*, 1998; Alvarez *et al.*, 1999; Gopal *et al.*, 1999; Sun y Sampson, 1999; Williams *et al.*, 1999; Benítez-Cardoza *et al.*, 2001; Jiménez *et al.*, 2003; Mathur *et al.*, 2006; Peimbert *et al.*, 2008a; Téllez *et al.*, 2008; Costas *et al.*, 2009; Lara-González *et al.*, 2014) y la determinación de algunos parámetros termodinámicos (como el ΔS , ΔH , ΔC_p , E_a , ΔG , m^\ddagger y constantes cinéticas de la reacción, etc.) dependiendo de la TIM reportada (Mainfroid *et al.*, 1996b; Benítez-Cardoza *et al.*, 2001; Téllez *et al.*, 2008; Costas *et al.*, 2009; Lara-González *et al.*, 2014).

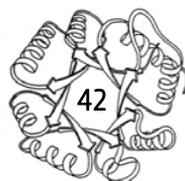
Ingeniería de proteínas aplicada en la TIM

Desde años recientes, existe gran interés en estudios para modificar el plegamiento de la TIM; estos intentos involucran tanto el diseño de nuevas proteínas, así como también modificaciones específicas de regiones de enzimas existentes. Dichos estudios se han enfocado especialmente hacia el incremento de la estabilidad, optimizar las propiedades catalíticas, formar quimeras de diferentes enzimas de barril TIM, o crear actividades enzimáticas completamente nuevas (Borchert *et al.*, 1993; Zomosa-Signoret *et al.*, 2003; Alahuhta *et al.*, 2008; Röthlisberger *et al.*, 2008). Recientemente, una revisión publicada a inicios de 2017 por Olivares-Illana y colaboradores, detalla de manera comprensiva los efectos en la catálisis, estabilidad, plegamiento y aplicaciones de fármacos de cada uno de los residuos que se han estudiado en la enzima, contribuyendo a entender mejor las propiedades de cada región de esta proteína fundamental (Olivares-Illana *et al.*, 2017).



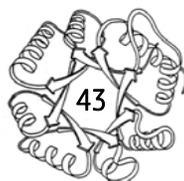
La estabilización de la proteína ha sido una tarea central en los estudios de ingeniería. El primer logro exitoso fue llevado a cabo por Klibanov y colaboradores. Ellos reemplazaron los residuos de asparagina de la intercara en γ TIM (N14T y N78I). La eliminación de los dos sitios de desamidación, aumenta la estabilidad cinética de la enzima al doble cuando es calentada a 100 °C (Aher *et al.*, 1987). Otra estrategia seguida para aumentar la estabilidad de la TIM ha sido estabilizar las hélices α . Para esto, se han creado mutantes que introducen nuevas interacciones electrostáticas, incrementan el contenido de alaninas en las hélices α y que elevan el contenido de prolinas, en todos los casos se encontró que la estabilidad de la enzima se veía incrementada varias kilocalorías (Mainfroid *et al.*, 1996). Diversos resultados apoyan la idea de que enzimas mesófilas pueden ser convertidas en enzimas superestables sin afectar las propiedades catalíticas a temperaturas mesófilas (Alvarez *et al.*, 1998; Williams *et al.*, 1999; Lambeir *et al.*, 2000)

A pesar de que el sitio activo de cada monómero tiene una arquitectura independiente, la disociación de la enzima resulta concomitantemente en la inactivación de la misma, aun cuando el monómero está correctamente plegado (Schliebs *et al.*, 1996). Debido a esto, se han realizado numerosos esfuerzos para tratar de comprender el mecanismo que subyace a la naturaleza oligomérica de la TIM. Una de las estrategias más utilizadas ha sido la desestabilización de la intercara de la enzima silvestre (ya sea acortándola o mutándola por aminoácidos desestabilizantes), a partir de la cual se han obtenido proteínas monoméricas mutantes (Borchert *et al.*, 1994; Borchert *et al.*, 1995). Otros ejemplos de monomerización de la TIM abarcan los realizados en Tb TIM (mono Tb TIM), la cual resultó ser estable pero con una k_{cat} 1000 veces mayor y una K_m 20 veces mayor respecto a la enzima silvestre, dando como resultado una disminución en la eficiencia catalítica en 4 órdenes de magnitud ($3.15 \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$) (Borchert *et al.*, 1993; Schliebs *et al.*, 1996). Otra monoTIM se logró obtener al acortar el asa 8 de unión a fosfato, con el fin de unir sustratos distintos a DHAP y G3P (Norledge *et al.*, 2001).



Así, se sabe que la catálisis presentada en todas estas mutantes es mucha más baja que la observada para la enzima silvestre. Estos monómeros de TIM se han sometido a diversos estudios para tratar de mejorar su actividad catalítica, en todos ellos, la eficiencia catalítica del monómero es aproximadamente 1000 veces menor que la de una enzima dimérica. Datos cristalográficos indican una mayor accesibilidad al solvente y movimiento de los aminoácidos catalíticos, lo que sugiere la alteración del sitio activo. También se ha mencionado que una de las razones clave de su baja actividad es la pérdida de rigidez en la intercara dimérica, lo que implica a su vez, conformaciones más flexibles para algunos de los residuos del sitio activo (Wierenga *et al.*, 1992; Thanki *et al.*, 1997), además de que la dimerización optimiza la geometría del sitio activo, ya que los contactos entre las subunidades disminuyen la flexibilidad de las asas catalíticas (Schliebs *et al.*, 1996). Esto implica que, en las formas monoméricas de la TIM con estructura conservada, la población de sitios activos competentes es muy baja. Al generar una mutante monomérica, una gran cantidad de área hidrofóbica de la intercara dimérica queda expuesta al solvente polar, lo que genera una desestabilización a la enzima (Borchert *et al.*, 1994; Zarate-Pérez *et al.*, 2009). En resumen, la monomerización invariablemente produce una drástica disminución en la actividad catalítica y una reducción en la estabilidad proteica, independientemente de la enzima estudiada.

Los acercamientos con mutagénesis masiva han sido utilizados para explicar la robustez de la TIM. Al mezclar el orden de las unidades β/α (el mejor arreglo tiene el orden de 2-4-6-8-1-3-5-7) se observa que la enzima sigue siendo dimérica, cooperativa en la desnaturalización y mantiene un contenido de estructura secundaria similar al estado original. Todo esto indica un cierto grado de autonomía en la formación de las subunidades β/α (Shukla y Guptasarma, 2004). Con la misma filosofía, también se han explorado un total de 182 mutaciones que preservan el patrón polar-hidrofóbico de la proteína, encontrado que 97 posiciones no tienen un efecto medible en la enzima. Con lo anterior mencionado, se concluye que la catálisis es preservada evolutivamente dentro de una plasticidad grande en la secuencia de la proteína (Silverman *et al.*, 2001).

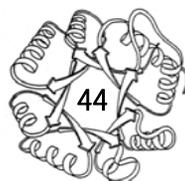


La estrategia de evolución dirigida también ha sido explorada, en la cual se ha encontrado la persistencia de las mutaciones A43P y T44A o T44S, y se postula que estas mutaciones estabilizan el asa 2 y serían la causa probable del mejoramiento de la actividad catalítica que presenta la variante de monoTIM *TbTIM* (RMMO-1) (Saab-Rincón *et al.*, 2001). En el caso de la TIM de humano, se han creado formas monoméricas en solución que son inactivas y con una baja estabilidad, cuyo estado de asociación depende de la concentración de proteína (Mainfroid *et al.*, 1996a).

Los experimentos de modificación química han mostrado que el sitio catalítico de un monómero puede ser completamente activo incluso si está acoplado a otro monómero cuyo sitio catalítico no es funcional (Sun *et al.*, 1992a; Sun *et al.*, 1992b). La formación de heterodímeros con *TcTIM* y *TbTIM* muestran que cerca de la mitad de la actividad de un monómero depende de la integridad de la interacción de la Cys15 de *TcTIM* (del monómero 1) con el asa 3 (del monómero 2) (Zomosa-Signoret *et al.*, 2007). El cambio de la Cys14 de la *TbTIM* ha mostrado su importancia en la oligomerización de la enzima, ya que al sustituirse por fenilalanina genera un monómero (Hernández-Alcántara *et al.*, 2002). Conjuntamente, mediante proteólisis limitada, en *TbTIM* pueden observarse diferencias estructurales entre los monómeros que forman el dímero y que la alteración de una subunidad puede generar cambios en la otra subunidad del dímero (Reyes-Vivas *et al.*, 2002). Asimismo, la derivatización de esta cisteína en *PfTIM* y *EcTIM* también promueve la monomerización (Maithal *et al.*, 2002; Rodríguez-Romero *et al.*, 2002).

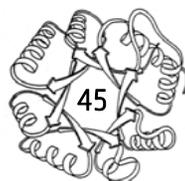
Trisafosfato Isomerasa: estudios realizados en el dominio Bacteria

Hasta el momento, la mayoría de los estudios sobre vías de plegamiento de la TIM se han realizado en enzimas procedentes de organismos eucariontes. Con respecto a la TIM en especies del dominio bacteria, hasta el momento se han reportado 35 estructuras cristalográficas, procedentes de 14 especies bacterianas distintas, en varias especies se tiene únicamente la estructura tridimensional,



sin tener algún dato sobre propiedades catalíticas, estado de oligomerización o de la ruta de plegamiento (tabla 4).

Debido a esto, el estudio de las TIMs bacterianas representa un área interesante y nueva para explorar diversos aspectos de interés sobre la enzima, ya que al tener las bacterias una gran diversidad de hábitats, se pueden encontrar condiciones distintas de adaptación que podrían reflejarse en el comportamiento diferencial de sus proteínas. Por ejemplo, en ciertas TIMs de especies bacterianas extremófilas, como *Thermotoga marítima*, *Bacillus stearothermophilus* y *Moritella marina*, se ha encontrado un incremento en el número de puentes salinos involucrados en mantener una conformación adecuada de la enzima. Igualmente, se ha observado un número mayor de interacciones hidrofóbicas en la parte interna del barril en el caso de *T. marítima* y *B. stearothermophilus*, cuando se comparan con las demás TIMs, lo que pudiera indicar que en conjunto, todos estos cambios le otorgan mayor estabilidad a la enzima bajo condiciones extremas de temperatura, osmolaridad o ambas (Maes *et al.*, 1999).





Microorganismo	Código PDB y resolución (Å)	Tipo de asociación	K_m (mM)	k_{cat} (min^{-1})	Eficiencia catalítica ($\text{M}^{-1} \text{seg}^{-1}$) [†]	T_m (°C) (técnica empleada)	Otros análisis realizados acerca de la enzima	Referencia
<i>Bartonella henselae</i>	3KXQ (1.6)	ND	ND	ND	ND	ND	ND	Abendroth <i>et al.</i> , 2009
<i>Brucella melitensis</i>	4NVT (2.1)	ND	ND	ND	ND	ND	ND	Lukacs <i>et al.</i> , 2013
<i>Burkholderia thailandensis</i>	4G1K (2.4)	ND	ND	ND	ND	ND	ND	Baugh <i>et al.</i> , 2013
<i>Geobacillus stearothermophilus</i>	2BTM (2.4)	Dímero	ND	ND	ND	76 (DSC)	Estabilidad enzimática a diferentes temperaturas	Delboni <i>et al.</i> , 1995; Alvarez <i>et al.</i> , 1999
<i>Escherichia coli</i>	1TRE (2.6)	Dímero	1.03	5.4×10^5	2.18×10^8	54 (DSC)	ND	Noble <i>et al.</i> , 1993; Alvarez <i>et al.</i> , 1998
<i>Francisella tularensis</i>	5UJW (2.7)	ND	ND	ND	ND	ND	ND	Chang <i>et al.</i> , 2017
<i>Helicobacter pylori</i>	25GQ (2.3)	Dímero	3.46	8.8×10^4	1.10×10^7	ND	Mutantes K183S, K183A, D213Q y D213A	Chu <i>et al.</i> , 2008
<i>Leptospira interrogans</i>	4X22 (2.1)	ND	ND	ND	ND	ND	Estudios estructurales	Pareek <i>et al.</i> , 2016
<i>Moritella marina</i>	1AW1 (2.7)	Dímero	1.9	4.2×10^5	9.21×10^7	41 (DSC)	Estabilidad enzimática a diferentes temperaturas	Alvarez <i>et al.</i> , 1998
<i>Mycobacterium tuberculosis</i>	3GVG (1.6)	Dímero	2.1	4.1×10^6	8.13×10^8	55 (CD)	Actividad a diferentes pH. Utilización de metales para su actividad. Estabilidad a diferentes temperaturas	Mathur <i>et al.</i> , 2006; Abendroth <i>et al.</i> , 2009
<i>Staphylococcus aureus</i>	39MY (1.9)	Dímero	ND	ND	ND	ND	Estudios estructurales	Mukherje <i>et al.</i> , 2012
<i>Streptococcus pneumoniae</i>	5IBX (1.7)	ND	ND	ND	ND	ND	ND	Minasov <i>et al.</i> , 2016
<i>Thermotoga maritima</i>	1B9B (2.9)	Tetrámero	2.6	1.1×10^3	1.76×10^5	82 (CD, IF, UVA)	Desnaturalización con GdnHCl. Actividad a diferentes pHs y temperaturas. Estudio de la enzima separada de la enzima fosfoglicerato cinasa	Beaucamp <i>et al.</i> , 1997; Maes <i>et al.</i> , 1999
<i>Thermus thermophilus</i>	1YYA (1.6)	Dímero	ND	ND	ND	ND	ND	Yamamoto y Kunishima, 1995 [‡]

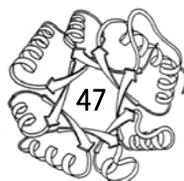
Tabla 4. Estudios reportados de la TIM en el dominio Bacteria.

†: Realizando la corrección en la K_m al considerar que sólo el 4% del G3P en solución tiene el sustrato aldehído no hidratado, única forma que la TIM es capaz de catalizar (relación 1:25) (Trentham *et al.*, 1969). ‡: Estudio únicamente estructural publicado en la base de datos de proteínas. **CD**: Dicroísmo Circular, **DSC**: Calorimetría Diferencial de Barrido, **GdnHCl**: hidrocloreuro de guanidina, **IF**: Fluorescencia Intrínseca, **ND**: no determinado, **UVA**: Absorción de UV.

Triosafofato Isomerasa: estudio previo realizado durante la maestría

Durante los estudios de maestría (Romero-Romero, 2013), se seleccionaron y estudiaron TIMs bacterianas pertenecientes a diversos fila bacterianos. La selección de las proteínas se realizó mediante un análisis filogenético del dominio Bacteria para elegir especies de interés evolutivo que nos pudiera proporcionar información de otros fila bacterianos donde no ha sido estudiada esta enzima. Para esto, se consideró el árbol filogenético publicado por Cicarelli y colaboradores (Cicarelli *et al.*, 2006) y, por otra parte, se analizó la clasificación filogenética de las bacterias, arqueas y eucariontes propuesta por Cavalier-Smith (Cavalier-Smith, 2002; Cavalier-Smith, 2006a-b) y se ubicaron los 3 supertaxones parafiléticos de las bacterias con sus respectivos taxones y fila, donde se incluyeron los estudios previos de la TIM realizados en bacterias. Posteriormente, se detectaron todos los fila bacterianos descritos al momento de la selección y se analizaron los eventos evolutivos clave ubicados en el dominio Bacteria.

Teniendo considerados estos eventos, se eligieron los fila representativos que cubrieran sucesos importantes en la evolución y que permitieran tener una descripción global de todo el árbol filogenético de las bacterias, descartando los fila donde ya se tuvieran estudios previos de la TIM. Una vez que se tuvieron los fila elegidos, se analizaron las especies representativas de éstos y con base en la importancia evolutiva (complementándola con una importancia biomédica o biotecnológica), se seleccionaron algunas especies para estudiar las características estructurales, cinéticas y de plegamiento de la TIM. Bajo estos parámetros de selección, las especies elegidas fueron: *Deinococcus radiodurans* (*DrTIM*), *Nostoc punctiforme* (*NpTIM*), *Azotobacter vinelandii* (*AvTIM*), *Gemmata obscuriglobus* (*GoTIM*), *Clostridium perfringens* (*CpTIM*) y *Streptomyces*

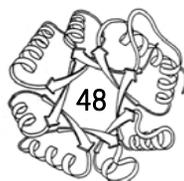


coelicolor (*Sc*TIM) (figura 12) (Romero-Romero *et al.*, 2015). Al conjunto de estas proteínas se le nombró como BacTIMs. En el apéndice A y B se presentan los datos de evolución bacteriana con los que se basó la selección de especies y también se muestran algunas de las características fisiológicas y bioquímicas de los organismos seleccionados. En la tabla 5 se enlistan algunas propiedades fisicoquímicas generales para las BacTIMs y en la figura 11 se muestra un alineamiento de las secuencias de las proteínas seleccionadas.

Tabla 5. Propiedades fisicoquímicas de las BacTIMs.

TIM (especie)	Número de residuos	Peso molecular del monómero (Da)	Punto isoeléctrico	Número de cisteínas	Número de prolinas	Número de triptófanos
<i>Deinococcus radiodurans</i> (<i>Dr</i> TIM)	247	25592.86	4.9	3	10	3
<i>Nostoc punctiforme</i> (<i>Np</i> TIM)	246	27052.72	5.0	3	11	3
<i>Azotobacter vinelandii</i> (<i>Av</i> TIM)	251	26029.84	5.8	5	10	2
<i>Gemmata obscuriglobus</i> (<i>Go</i> TIM)	252	27024.92	6.4	3	11	3
<i>Clostridium perfringens</i> (<i>Cp</i> TIM)	251	27339.26	5.1	6	7	2
<i>Streptomyces coelicolor</i> (<i>Sc</i> TIM)	261	28017.76	5.3	4	6	2

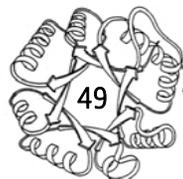
Todas estas enzimas fueron expresadas en *Escherichia coli* y purificadas a homogeneidad, a excepción de *Av*TIM que fue descartada del estudio debido a que la enzima se depositó en cuerpos de inclusión, por lo que no fue posible purificarla tras múltiples intentos en la modificación de los protocolos de sobreexpresión. La actividad enzimática mostró una cinética Michaeliana con una eficiencia catalítica similar a la de otras TIMs reportadas. La cromatografía de exclusión molecular indicó que todas las TIMs analizadas son diméricas. El espectro de Dicroísmo Circular (DC) indicó una composición de estructura secundaria propia de barriles TIM en todas las enzimas. El máximo de emisión de los espectros de Fluorescencia Intrínseca (IF) mostró que los triptófanos de todas las



enzimas estudiadas están en un centro hidrofóbico, indicando el correcto plegado de las proteínas. El desplegamiento térmico se estudió mediante DC y DSC. El desplegamiento de *Np*TIM y *Go*TIM fue irreversible y la T_m se incrementó ≈ 5 K cuando la velocidad de barrido se aumentó de 30 a 180 K h⁻¹, indicando control cinético, un comportamiento comúnmente observado en las TIMs eucariontes previamente estudiadas. El valor de la energía de activación es semejante al de otras TIMs analizadas. El estado de transición para ambas TIMs es dimérico y muy parecido al nativo en términos de exposición al solvente.

<i>NpTIM</i>	1	GSHV	-	RK	IV	AGNWK	MFK	QAET	QEF	LQG	FLPH	LEETP	-	QGRE	VIL	CP	PFTD	SVL	SKT	LHGS	61		
<i>GoTIM</i>	1	GSHMP	TR	KKFV	AGNWK	MNTT	LAE	AKAL	GAA	VA	KG	VTD	-	-	DRVT	VAV	FPP	PWL	TAV	GEV	LKGS	61	
<i>DrTIM</i>	1	GSHMQ	T	-	-	LLAL	NWK	MNK	TPT	E	AR	SWAE	EL	TKY	APA	-	-	EGV	DL	AVL	APAL	DL	59
<i>CpTIM</i>	1	GSHM	-	RTP	II	AGNWK	MHY	TIDE	AV	KL	VEEL	KPL	VKDA	-	-	KCE	V	V	CV	PT	FV	CL	59
<i>ScoTIM</i>	1	GSHMT	TR	TPL	MAGNWK	MNL	NHLEA	IA	HV	QKL	AFAL	AD	KDY	DAVE	AVL	AP	FTD	LRS	VQ	TL	VDGD	64	
<i>NpTIM</i>	62	L	-	-	I	QLGA	QNI	HWE	EFG	AY	TGEI	SGP	M	TES	G	VR	FVI	VGH	SERR	QY	FGET	DAT	123
<i>GoTIM</i>	62	P	-	-	VAL	GAQ	DVS	SE	KKG	AF	TGEV	SP	AML	L	ET	G	CK	YAL	I	GHS	ERR	HIG	123
<i>DrTIM</i>	60	-	-	-	I	AF	GGQ	DV	SA	HES	GAY	TGEI	SA	A	M	L	K	D	A	G	A	S	120
<i>CpTIM</i>	60	N	-	-	I	KV	GA	Q	N	M	H	F	E	E	K	G	A	F	T	G	E	I	121
<i>ScoTIM</i>	65	K	L	K	I	KY	GA	Q	D	I	S	A	H	D	G	A	Y	T	G	E	I	S	128
<i>NpTIM</i>	124	G	L	T	P	I	L	C	V	G	E	T	K	Q	R	D	A	G	E	T	E	S	184
<i>GoTIM</i>	124	G	L	S	V	V	L	C	M	G	E	T	L	A	E	R	E	R	G	L	Q	E	187
<i>DrTIM</i>	121	G	L	L	P	I	V	C	V	G	E	N	L	D	V	R	E	R	G	E	H	V	180
<i>CpTIM</i>	122	N	L	T	P	I	L	C	C	G	E	T	L	E	Q	R	E	N	G	T	T	N	185
<i>ScoTIM</i>	129	G	L	T	P	I	L	C	V	G	E	E	L	D	V	R	E	A	G	N	H	V	192
<i>NpTIM</i>	185	N	R	I	I	G	L	I	R	S	Q	L	S	N	-	-	-	-	-	-	P	N	241
<i>GoTIM</i>	188	Q	E	A	H	A	F	V	R	S	K	L	R	L	L	Y	G	D	K	I	A	D	251
<i>DrTIM</i>	181	E	E	L	A	A	A	I	R	G	A	L	R	E	Q	Y	G	A	R	-	A	E	243
<i>CpTIM</i>	186	N	E	T	I	A	A	I	R	A	M	V	A	E	M	F	G	Q	E	V	A	D	249
<i>ScoTIM</i>	193	Q	E	V	C	A	A	I	R	G	K	L	A	E	L	Y	S	Q	E	L	A	D	256
<i>NpTIM</i>	242	N	F	H	L	V																246	
<i>GoTIM</i>	252	K	A	A	G	-																255	
<i>DrTIM</i>	244	D	A	L	R	-																247	
<i>CpTIM</i>	250	N	Y	-	-	-																251	
<i>ScoTIM</i>	257	R	F	R	D	Q																261	

Figura 11. Alineamiento de las secuencias de las BacTIMs. Los residuos conservados se representan de azul claro a oscuro de acuerdo al incremento de conservación en secuencia.



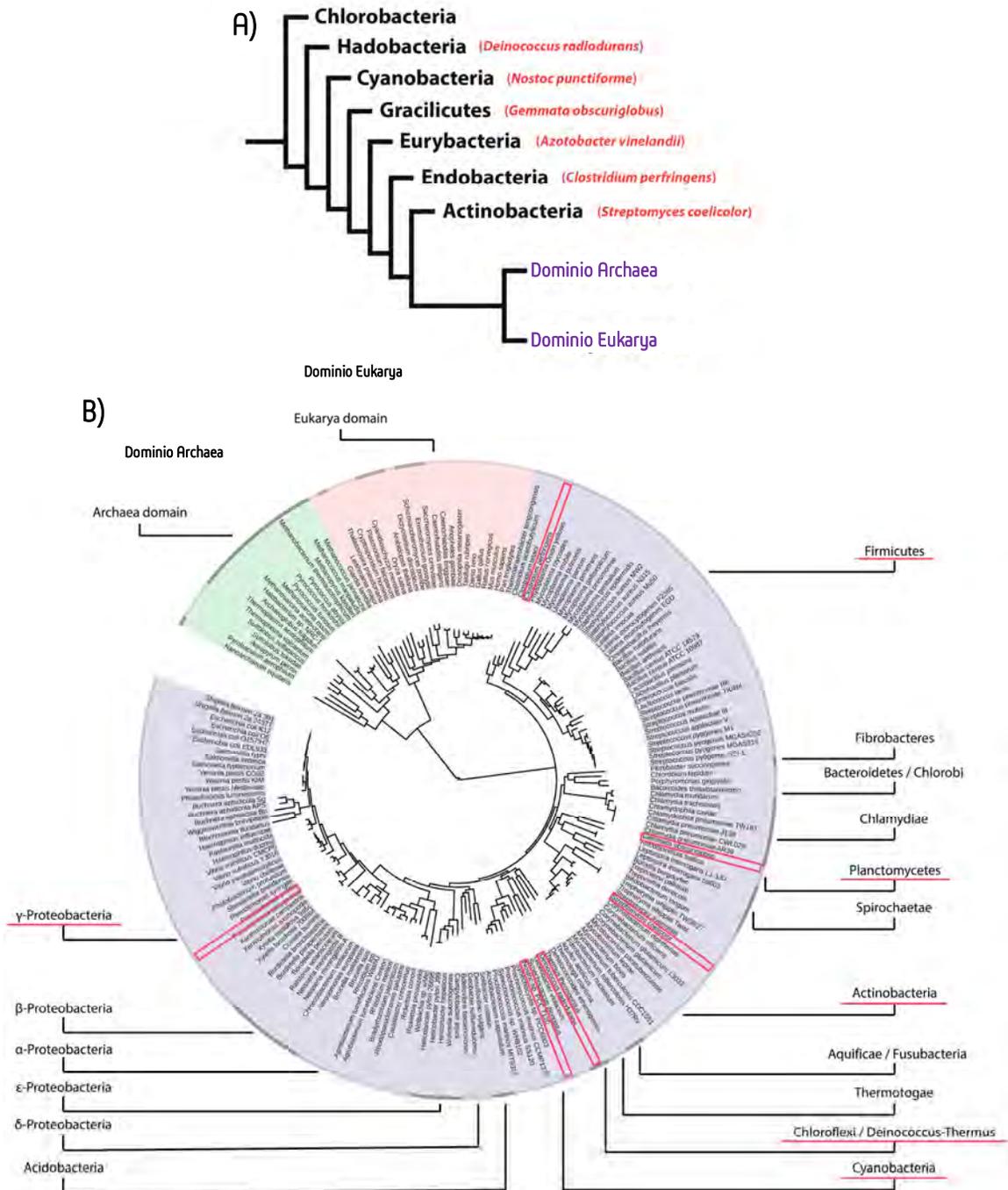
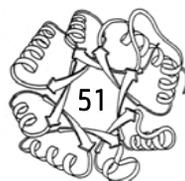


Figura 11. Representaciones filogenéticas de las BacTIMs. A) Cladograma de los supertaxa bacterianos (según Cavalier-Smith, 2006) que muestra la especie a partir de la cual se estudió la secuencia TIM indicada. B) Árbol filogenético que muestra las TIMs estudiadas durante la maestría (modificado de Letunic y Bork, 2011). Para el *phylum* de *AvTIM* (γ -Proteobacteria) sólo se muestra una ubicación aproximada en el árbol filogenético.

Todas estas enzimas fueron expresadas en *Escherichia coli* y purificadas a homogeneidad, a excepción de *AvTIM* que fue descartada del estudio debido a que la enzima se depositó en cuerpos de inclusión, por lo que no fue posible purificarla tras múltiples intentos en la modificación de los protocolos de sobreexpresión. La actividad enzimática mostró una cinética Michaeliana con una eficiencia catalítica similar a la de otras TIMs reportadas. La cromatografía de exclusión molecular indicó que todas las TIMs analizadas son diméricas. El espectro de Dicroísmo Circular (DC) indicó una composición de estructura secundaria propia de barriles TIM en todas las enzimas. El máximo de emisión de los espectros de Fluorescencia Intrínseca (IF) mostró que los triptófanos de todas las enzimas estudiadas están en un centro hidrofóbico, indicando el correcto plegado de las proteínas. El desplegamiento térmico se estudió mediante DC y DSC. El desplegamiento de *NpTIM* y *GoTIM* fue irreversible y la T_m se incrementó ≈ 5 K cuando la velocidad de barrido se aumentó de 30 a 180 K h⁻¹, indicando control cinético, un comportamiento comúnmente observado en las TIMs eucariontes previamente estudiadas. El valor de la energía de activación es semejante al de otras TIMs analizadas. El estado de transición para ambas TIMs es dimérico y muy parecido al nativo en términos de exposición al solvente.

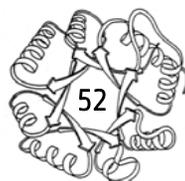
De manera sorprendente, para *DtTIM*, *CpTIM* y *ScTIM*, el desplegamiento térmico fue reversible con una histéresis menor a un grado entre las trazas de desplegamiento y replegamiento, además la ΔT_m entre las velocidades de barrido estudiadas fue menor de 1.5 grados; todas estas evidencias son indicativas de procesos reversibles y al equilibrio, esto es, de control termodinámico, un comportamiento nunca antes encontrado para este tipo de enzimas. Debido al tiempo de duración de los estudios de la maestría, el análisis termodinámico se limitó a calcular algunos parámetros fisicoquímicos de manera preliminar, por lo que uno de los objetivos para el proyecto doctoral fue la profundización de la caracterización termodinámica para estas 5 proteínas, así como también darnos a la tarea de encontrar y proponer los orígenes moleculares que subyacen la reversibilidad en el desplegamiento térmico de esta familia de proteínas.



La reversibilidad en el desplegamiento térmico es un resultado interesante y relevante, ya que nunca se había tenido una TIM con reversibilidad elevada y al equilibrio (así mismo, el proceso es robustamente reversible ya que al calentar hasta por 30 grados después de la transición calorimétrica, se sigue observando reversibilidad del proceso). La mayoría de las TIMs estudiadas hasta el momento habían presentado transiciones irreversibles en un proceso de desnaturalización térmica, y en el caso de que hubiera reversibilidad, como lo es en *ytTIM* (bajo condiciones muy especiales), *HsTIM* (y sus mutantes con baja energía de disociación *HsTIM/M14Q* y *HsTIM/R98Q*) y *EhTIM*, el proceso no estaba al equilibrio, ya que el regreso a la forma nativa muestra una histéresis alta (aproximadamente 10 °C para *ytTIM* (Benítez-Cardoza *et al.*, 2001); 2 °C para *EhTIM* -sólo la primera transición ya que la segunda es irreversible- (Téllez *et al.*, 2008). En contraste, en las TIMs bacterianas (*DfTIM*, *CpTIM* y *ScTIM*) la histéresis es menor a 1 °C entre las trazas de desnaturalización y renaturalización. La reversibilidad del proceso de desplegamiento térmico pudiera tener algún significado de acuerdo al nicho ecológico en el que habitan estas bacterias y su ubicación en el árbol filogenético, aunque esto quedó como una pregunta abierta durante el proyecto a realizarse durante el doctorado.

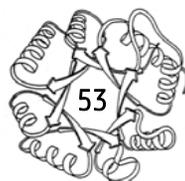
La proteína sTIM11 como modelo de estudio de un barril TIM diseñado *de novo*

El estudio de las proteínas está pasando por una nueva fase de diseño e ingeniería. Actualmente se están desarrollando nuevos enfoques y métodos que permiten el diseño de proteínas más allá de las conocidas naturalmente. Esta posibilidad de diseñar proteínas completamente nuevas abre ciertas preguntas: ¿qué topologías construimos? ¿Cómo construimos en el espacio conformacional donde hay pocas proteínas naturales que nos guíen? ¿Qué usos se pueden dar a las proteínas resultantes? ¿Cómo nos puede ayudar el conocimiento adquirido de proteínas naturales en el diseño de proteínas nuevas?



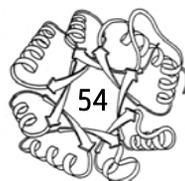
Éstas y más preguntas son la base de múltiples investigaciones recientes. Por ejemplo, ha habido un progreso en el diseño *de novo* de proteínas como Top7, toroides helicoidales, hélices transmembrana para transporte de Zn^{2+} , péptidos autoasociados, receptores agonistas/antagonistas, epítomos para el diseño de vacunas, entre otros (Kuhlman *et al.*, 2003; Huang *et al.*, 2014; Joh *et al.*, 2014; Koga *et al.*, 2012; Smadbeck *et al.*, 2014; Bellows-Paterson *et al.*, 2012; Khoury *et al.*, 2014; Correia *et al.*, 2014). En esta vertiente, como ya se mencionó anteriormente, el explorar nuevas secuencias pertenecientes a topologías ya conocidas resulta interesante, ya que podría permitirnos no sólo entender más sobre las bases que rigen el plegamiento y estabilidad de las proteínas, sino también la posibilidad el diseño de diseñar funciones novedosas y útiles.

Bajo esta premisa, la topología de barril $(\beta/\alpha)_8$ es una de las más robustas, versátiles y abundantes que existen en la naturaleza y, por esto mismo, ha sido a través de los años un objetivo claro para el diseño e ingeniería de proteínas (Höcker, 2005; Kiss *et al.*, 2013; Höcker *et al.*, 2004; Höcker *et al.*, 2009; Claren *et al.*, 2009). Tanta es su relevancia estructural, que el barril TIM fue uno de los cinco objetivos en el primer taller sobre diseño de proteínas de la Organización Europea de Biología Molecular realizado en 1987 (EMBO, Vienna 1987), pero ningún esfuerzo de diseño *de novo* hasta antes de 2016 había dado frutos precisos, es decir, proteínas con estructura terciaria claramente definida (Goraj *et al.*, 1990; Tanaka *et al.*, 1994; Houbrechts *et al.*, 1995; Offredi *et al.*, 2003; Figueroa *et al.*, 2013; Nagarajan *et al.*, 2015). Finalmente, a inicios del año 2016 fue descrito el diseño exitoso de un barril TIM *de novo*: la proteína sTIM11 (Huang *et al.*, 2016a). La importancia de este trabajo recae en que a pesar de los esfuerzos realizados durante más de 25 años, no se había podido diseñar un barril TIM *de novo*. La proteína diseñada es un barril TIM de 184 residuos con una simetría de orden 4 (figura 32a), es decir, la estructura primaria total de la proteína se conforma de 4 segmentos idénticos, los cuales si se comparan entre sí, su RMSD estructural es de 0.17 Å, lo que confirma que los fragmentos son casi idénticos (figura 13c). sTIM11 presenta un RMSD de 1.28 Å cuando se compara el modelo diseñado con la estructura tridimensional experimental determinada,



lo que comprueba que el diseño fue realizado con una precisión atómica (figura 13b). sTIM11 fue diseñada con dos cisteínas posicionadas para formar un puente disulfuro, aunque en la estructura cristalográfica este puente no se observa (ver discusión más adelante en la sección de resultados). También, las búsquedas en PSI-BLAST no identifican secuencias similares con otras proteínas barril-TIM, lo que indica que el diseño está distante a otras familias de barriles TIM naturales (figura 13d). Lo anterior demuestra lo que se ha discutido en líneas anteriores: la naturaleza ha muestreado sólo un subconjunto del espacio-secuencia disponible para formar proteínas con un plegamiento de este tipo. Estudios como este nos ayudan a comprender mejor el plegamiento de estas macromoléculas, no sólo en aspectos estructurales, sino también en aspectos termodinámicos. Además, obtener las curvas de estabilidad podría ayudarnos a entender más el cómo modular ésta racionalmente tanto diseñando proteínas como realizando ingeniería en las proteínas naturales que existen.

La caracterización fisicoquímica de sTIM11 arrojó que esta proteína se pliega adecuadamente y su espectro de DC es compatible con una proteína que contiene estructura secundaria α/β (figura 14a). Se encontró también que la proteína posee una alta termoestabilidad, presentando un desplegamiento térmico cooperativo con una T_m de aproximadamente 80 °C (figura 14b). Sin embargo, al analizar el experimento se pudo observar que la desnaturalización térmica presenta una pretransición con una pendiente muy grande y un comportamiento indicativo de algún cambio en la estructura nativa previa al desplegamiento completo, cuestión que podría ser debida a un mal empaquetamiento de la proteína. Por otro lado, cuando el desplegamiento de la proteína fue estudiado utilizando hidrocloreuro de guanidinio (desplegamiento seguido por dicroísmo circular y fluorescencia), se observan curvas cooperativas (figura 14c) y de las cuales, ajustando los datos a un modelo de dos estados, se estima un valor de ΔG a 25 °C de 4.2 kcal mol⁻¹.



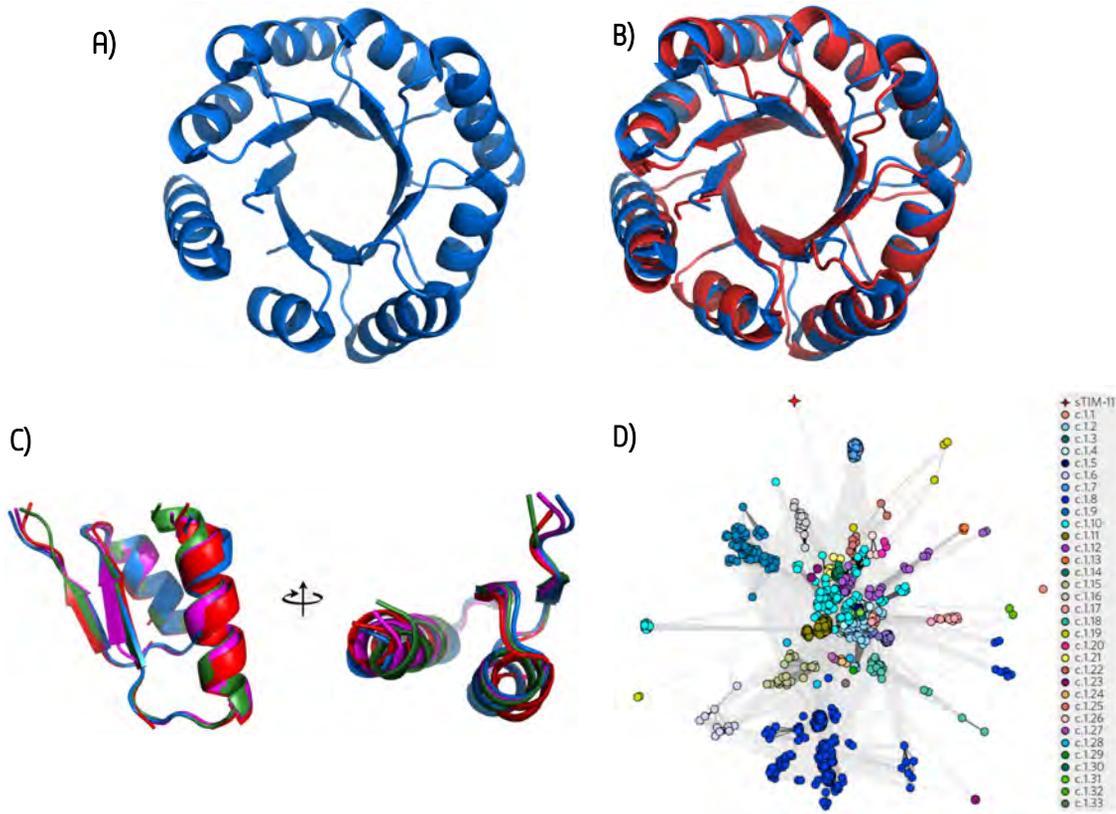


Figura 13. La proteína sTIM11. A) Estructura tridimensional de sTIM11, nótese la simetría de orden 4 de la proteína. B) Alineamiento estructural del modelo de sTIM11 (rojo) con la estructura determinada experimentalmente (azul). C) Alineamiento estructural de los 4 fragmentos repetitivos que conforman a sTIM11. D) Relaciones de las secuencias entre sTIM11 y los barriles TIM naturales. Se muestran los agrupamientos por HHsearch contra la base de datos Astral SCOPe 2.04. Las conexiones entre los grupos están mostradas con líneas grises. Modificado de Huang *et al.*, 2016a.

De los datos mencionados, sobresalen dos interrogantes por contestar: 1) ¿a qué se debe el comportamiento en la pretransición de la curva de desplegamiento de la proteína? y 2) ¿por qué la proteína posee una baja estabilidad en comparación con otros barriles TIM naturales (cuyos valores están entre 8-25 kcal mol⁻¹; Rudolph *et al.*, 1992; Cipolla *et al.*, 2012)? En especial esta última es importante porque dicha proteína diseñada se desea ocupar para el diseño de actividades enzimáticas o sitios de unión y al ser su estabilidad tan baja, el éxito en estos objetivos se reduce considerablemente. Por estas dos razones, una de las perspectivas del trabajo (y que fue retomada como se puede ver más adelante en la sección de objetivos) fue incrementar la estabilidad de sTIM11,

es decir, realizar nuevos diseños utilizando a sTIM11 como base para generar nuevas proteínas con el fin de que alguna de ellas fuera más estable y sirviera como andamiaje para trabajos futuros de diseño de actividades enzimáticas y/o sitios de unión. Este diseño racional se basará algunos principios fisicoquímicos que hemos aprendido de los estudios en barriles TIM naturales y, en específico, de lo que conocemos hasta el momento de la Triosafosfato Isomerasa.

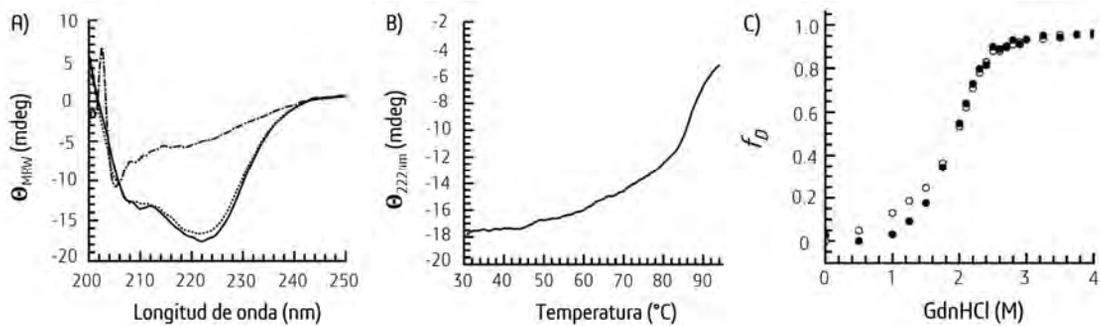
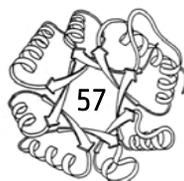


Figura 14. Caracterización estructural y fisicoquímica de sTIM11. A) Espectro de DC medido a 30 °C (línea continua), a 95 °C (línea discontinua) y a 30 °C después de desnaturalizar (línea punteada). B) Desplegamiento térmico seguido por dicroísmo circular. C) Desplegamiento químico utilizando GdnHCl como agente desnaturalizante. Se muestran las curvas seguidas por dicroísmo circular (símbolos llenos) y fluorescencia (símbolos vacíos). Modificado de Huang *et al.*, 2016a.

JUSTIFICACIÓN

La diversidad funcional que desempeñan las proteínas en los seres vivos es impresionantemente alta. Al saber que las proteínas son componentes esenciales de las células y que la vida es un escenario múltiple resultante de la actividad de las mismas, no es de extrañar que la variedad de estas macromoléculas posean la información necesaria para conocer los determinantes que han permitido el mantenimiento y la evolución de las muy diversas propiedades fisicoquímicas de las proteínas, como lo son la estabilidad y el plegamiento de las mismas. ¿Cómo han evolucionado las proteínas en términos de sus propiedades termodinámicas? Esta pregunta está en el centro de nuestros esfuerzos científicos y sobre el quehacer de los métodos bioinformáticos, bioquímicos y de biología estructural actuales.

Dentro de todas las arquitecturas moleculares descritas hasta hoy en día, el arreglo estructural $(\beta/\alpha)_8$ (también conocido como barril TIM) es una topología modelo para estudios que intentan describir el paisaje conformacional de las proteínas, tanto por su amplia distribución en la naturaleza así como por sus propiedades estructurales y fisicoquímicas que posee este tipo de plegamiento. El conocimiento a profundidad del paisaje conformacional que engloba dichas propiedades, así como la relación energético-estructural que existe entre éstas, tanto de proteínas naturales como diseñadas *de novo*, nos permitirá comprender mejor los determinantes globales que modulan la estabilidad y el plegamiento de los barriles TIM. Además, este conocimiento nos abre las puertas a un estudio evolutivo detallado de la relación estructura-función y al diseño racional de proteínas nuevas con funciones moleculares específicas.



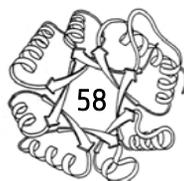
OBJETIVOS

Objetivo general

- Conocer, explorar y caracterizar el paisaje conformacional de barriles TIM naturales y diseñados *de novo*.

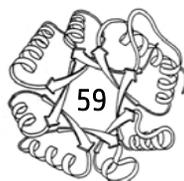
Objetivos particulares: barriles TIM naturales

- i.* Caracterizar los diferentes niveles estructurales de proteínas naturales bacterianas que pertenecen a la familia de la Triosafosfato Isomerasa.
- ii.* Conocer los parámetros cinéticos de las TIMs tanto para sus sustratos naturales como para uno de sus inhibidores.
- iii.* Estudiar el desplegamiento de las enzimas utilizando diferentes agentes desnaturizantes: dilución, agentes químicos y temperatura.
- iv.* Sugerir algunas propiedades de la Triosafosfato Isomerasa que son determinantes para que el desplegamiento térmico esté bajo control cinético o control termodinámico.



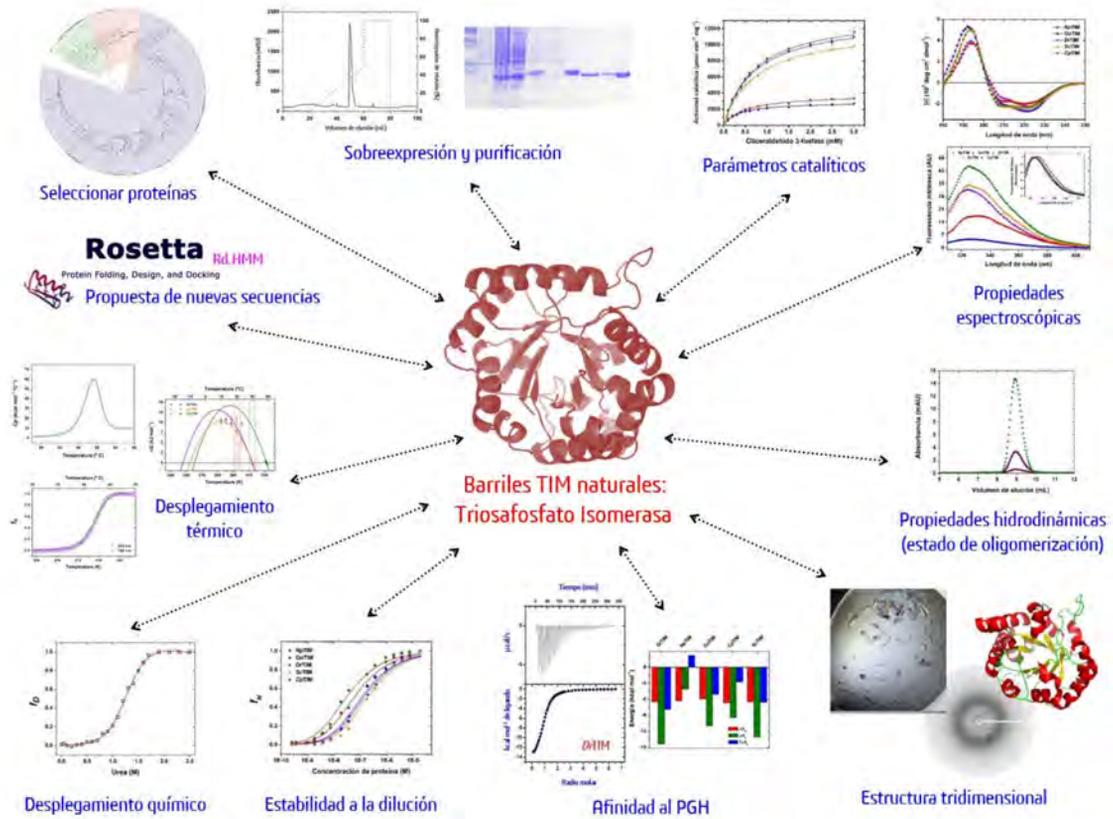
Objetivos particulares: barriles TIM diseñados *de novo*

- i.* Generar una nueva serie de diseños *de novo* con el fin de incrementar la estabilidad de la proteína sTIM11.
- ii.* Caracterizar el desplegamiento de los diseños utilizando temperatura y agentes químicos.
- iii.* Determinar la estructura tridimensional de los diseños más exitosos.
- iv.* Proponer nuevas secuencias tanto de barriles TIM naturales como diseñados *de novo* para analizar en un futuro.
- v.* Establecer una relación energético-estructural de las proteínas naturales y diseñadas para conocer los determinantes que modulan la estabilidad y el plegamiento de los barriles TIM.

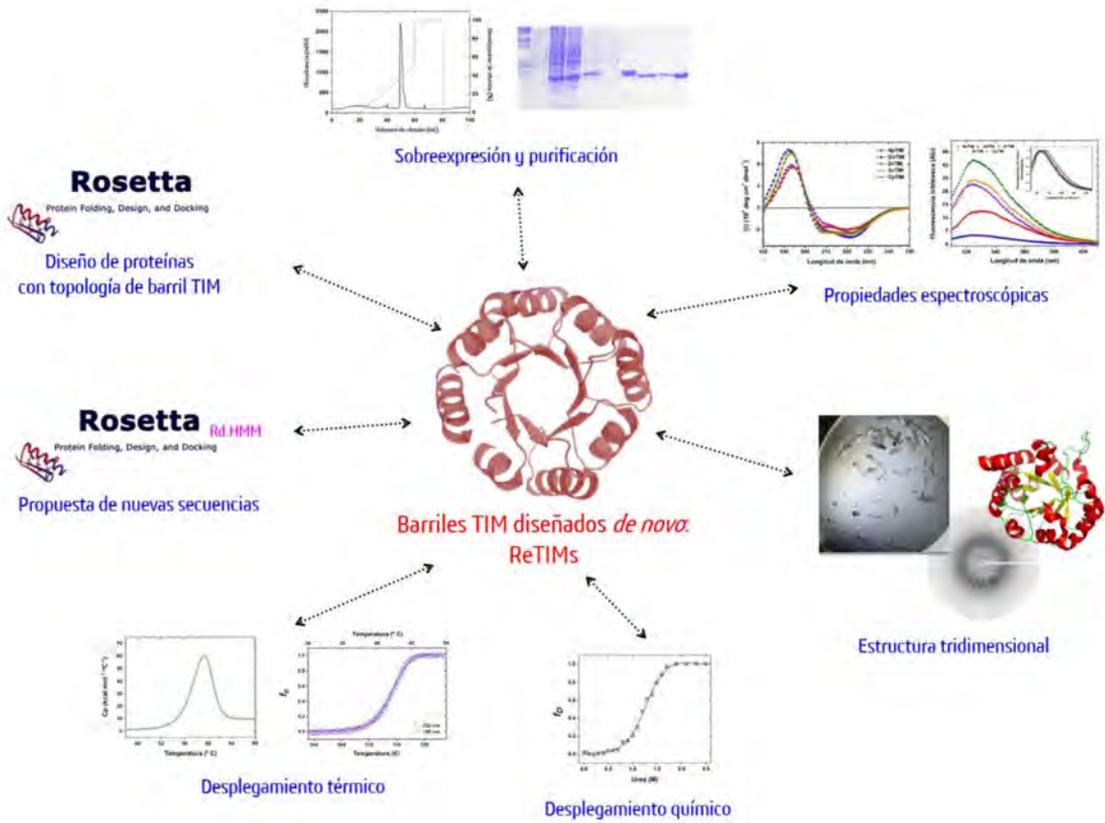


ESQUEMA GENERAL DE TRABAJO

Parte 1: Barriles TIM naturales



Parte 2: Barriles TIM diseñados *de novo*



ESTRATEGIA METODOLÓGICA

Reactivos y soluciones utilizadas

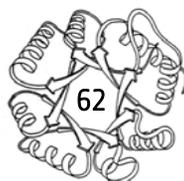
Todos los reactivos utilizados fueron de la marca Sigma-Aldrich® a excepción de aquellos donde se haga la aclaración pertinente.

Parte 1: Barriles TIM naturales

En esta parte metodológica se incluirán algunos métodos que fueron descritos ampliamente en la tesis de maestría. De igual forma, se presentarán algunos resultados que en su momento eran preliminares y que, durante el doctorado se mejoró la calidad de los datos. También se presentarán algunos pocos resultados obtenidos durante la maestría, con el fin de dar claridad y continuidad al lector de la presente tesis.

Sobreexpresión de la proteína

Los genes de las TIMs bacterianas seleccionadas durante el trabajo de maestría fueron previamente subclonados en un vector de expresión pET-28b(+) (Novagen®) con los sitios de restricción *NdeI* en el extremo 5' y *XhoI* en el extremo 3' en cada uno de ellos. Los genes subclonados se transformaron en la cepa de expresión *Escherichia coli* BL21(DE3)*pLysS* (Thermo Fisher Scientific®) mediante la técnica de choque térmico. Para analizar la sobreexpresión de las proteínas, inicialmente se realizaron pruebas ajustadas a un volumen pequeño con el fin de detectar si la proteína estaba siendo



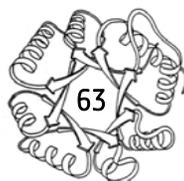
sobreexpresada, tal cual es reportado en Romero-Romero, 2013. Las pruebas de sobreexpresión se analizaron en geles de poliacrilamida al 12% en condiciones desnaturalizantes y fueron teñidos con una solución al 0.1% de azul de coomassie G-250 (diluido en 50% de metanol y 10% de ácido acético), cargando tanto el sobrenadante como la pastilla después del centrifugado para observar si la proteína se encontraba en la parte soluble o insoluble del extracto.

Purificación de la proteína

La purificación de las enzimas se realizó en volúmenes grandes mediante el protocolo que incluye primero una columna de afinidad HisTrap™ HP de 5 mL (GE Healthcare Life Sciences ®) acoplada a un sistema FPLC (GE Healthcare Life Sciences ®), después la remoción de la etiqueta de histidinas con trombina mediante el kit Thrombin CleanCleave (Sigma-Aldrich ®), siguiendo de nuevo con la columna de afinidad para retirar la parte no cortada por la trombina y finalmente se introdujo la proteína en una columna de intercambio iónico Mono Q (Pharmacia Biotech ®). El protocolo de purificación completo se reporta en Romero-Romero, 2013 y Romero-Romero *et al.*, 2015.

Los amortiguadores utilizados fueron:

- Amortiguador A: NaH₂PO₄ 35 mM, Imidazol 35 mM, NaCl 300 mM, pH: 8.0.
- Amortiguador B: NaH₂PO₄ 35 mM, Imidazol 500 mM, NaCl 300 mM, pH: 8.0.
- Amortiguador C: Tris 50 mM, CaCl₂ 10 mM, pH 8.0.
- Amortiguador D: Trietanolamina 10 mM, EDTA 1mM, DTT 1 mM, pH: 8.0.
- Amortiguador E: Trietanolamina 10 mM, EDTA 1mM, DTT 1 mM, NaCl 1000 mM, pH: 8.0.



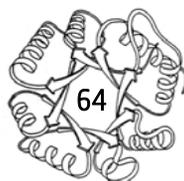
Cuantificación de la proteína

La concentración de proteína se determinó por su absorbancia a 280 nm. El coeficiente de absorptividad molar (ϵ_{280}) utilizado para la medición fue un valor teórico obtenido mediante el servidor ProtParam (Gasteiger *et al.*, 2005) considerando la secuencia de proteína y las cisteínas en su forma reducida (Pace *et al.*, 1995). Alternativamente, también se utilizó el método de cuantificación por ácido bicinconínico (BCA) (Smith *et al.*, 1985).

Mediciones espectroscópicas

Dicroísmo Circular

Los ensayos de Dicroísmo Circular (DC) se realizaron en un espectropolarímetro Chirascan™ (Applied Photophysics®) acoplado a un sistema Peltier (PolyScience®) para el control de temperatura. El espectro fue seguido de 180-260 nm de longitud de onda. La concentración de proteína utilizada para los espectros fue de 0.4 mg mL⁻¹ y se realizó en amortiguador de G (NaH₂PO₄ 10 mM, pH: 8.0) a una temperatura de 25 °C. Previamente se compararon los amortiguadores D y F para ver si su utilización provocaba algún cambio en la proteína, pero como no se observaron diferencias, se decidió utilizar en todos los experimentos el amortiguador F ya que es un amortiguador que permite alcanzar longitudes de onda menores que el amortiguador D, el cual genera más ruido en los espectros.



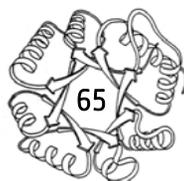
Fluorescencia Intrínseca

Las mediciones de fluorescencia intrínseca (IF) se realizaron en un espectrofluorómetro PC1 (ISS ®) acoplado a un sistema Peltier (PolyScience ®) para controlar la temperatura. Los experimentos se realizaron a 0.2 mg mL^{-1} en un amortiguador G y a una temperatura de $25 \text{ }^\circ\text{C}$. Se utilizó una longitud de onda de excitación de 295 nm y se midió la emisión en el intervalo de 310 a 410 nm , monitoreando la intensidad de fluorescencia cada 1 nm . El centro de masa espectral (CME) se calculó mediante: $\text{CME} = \frac{\sum \lambda I_\lambda}{\sum I_\lambda}$, donde I_λ es la intensidad de fluorescencia para cada longitud de onda (λ).

Resolución de la estructura tridimensional

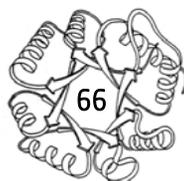
La determinación de la estructura tridimensional se realizó mediante la técnica de cristalografía de proteínas con difracción de rayos X en el laboratorio Nacional de Estructura de Macromoléculas del Instituto de Química, UNAM en colaboración con la Dra. Adela Rodríguez Romero. Como primer paso, se analizaron las condiciones de cristalización de otras TIMs depositadas en el PDB. Como primeras pruebas de cristalización, se eligieron 42 condiciones distintas de los kits Crystal Screen I y II (Hampton Research ®) para ensayar cada proteína, todas a una concentración de 10 mg mL^{-1} en amortiguador H (Trietanolamina 10 mM , EDTA 1 mM , DTT 1 mM , 150 mM de NaCl, pH: 8.0). Todas las pruebas se realizaron mediante el método de cristalización con difusión de vapor en gota colgante. Cada condición de cristalización se preparó utilizando $500 \text{ } \mu\text{L}$ de la solución cristalizante en el pozo y una relación de $1:1$ de la proteína y la solución cristalizante en un portaobjetos siliconizado; los pozos se sellaron con grasa de vacío y se incubaron en un cuarto de temperatura controlada a $18 \text{ }^\circ\text{C}$.

Después de 3 semanas de incubación, se observaron las placas que contenían las pruebas de cristalización y se eligieron un total de 12 condiciones donde se observó alguna formación cristalina



(esférulos, agujas, microcristales, maclas, etc.), las cuales se utilizaron para hacer una matriz con 7 concentraciones de proteína distintas: 2.5, 5, 7.5, 10, 12.5, 15, 17.5 mg mL⁻¹; además, en el pozo de cristalización, se colocaron 200 µL de aceite mineral para disminuir la velocidad de formación de cristales y favorecer su calidad. El método de cristalización y resto de las condiciones fueron las mismas utilizadas en las primeras pruebas colocadas. En algunas proteínas donde se observó de nuevo formación cristalina pero sin tener la calidad necesaria para difractar, se probaron otras matrices de cristalización, variando la concentración de sales, agente precipitante y pH. Además, se colocaron pruebas nuevas con soluciones ya ensayadas pero ahora con proteína recién purificada. Se probaron los cristales con buena calidad para difractar y en aquellos que produjeron un buen patrón de difracción y a buena resolución, se colectó el set completo de datos.

La colecta se realizó en un ánodo rotatorio Rigaku Micromax-007 HF (Rigaku ®) con una longitud de onda de 1.54 Å y con un detector Rigaku Raxis IV++ (Rigaku ®). Después de la colecta de datos, se realizó la integración, reducción y escalamiento de los mismos utilizando los programas XDS (Kabsch, 2010) y Scala, ubicado en la suite de CCP4 v.6.1.2 (Winn *et al.*, 2011). Todas las estructuras fueron resueltas utilizando el método de reemplazo molecular y mediante el programa Phaser (McCoy *et al.*, 2007) de la suite de Phenix (Adams *et al.*, 2010). Las TIMs de *Thermotoga maritima* (ID PDB: 1B9B), *Thermus thermophilus* (ID PDB: 1YYA), *Mycobacterium tuberculosis* (ID PDB: 3TA6) y *Geobacillus stearothermophilus* (ID PDB: 1BTM) se usaron como el modelo de partida para la sustitución del reemplazo molecular para *Go*TIM (Z-score de la función de traslación (TFZ): 42.3 y probabilidad de ganancia log (LLG): 1360.2), *D*TIM (TFZ: 32.8 y LLG: 1104.3), *S*cTIM (TFZ: 45.6 y LLG: 3064.9) y *C*pTIM (TFZ: 21.3 y LLG: 755.8), respectivamente. Los ciclos de afinamiento se efectuaron con el programa *phenix.refine* (Adams *et al.*, 2010). El modelo de la proteína se construyó con COOT (Emsley y Cowtan, 2004) utilizando los mapas F_o-F_c y 2F_o-F_c. Al finalizar de construir el modelo de la proteína, se validó en el servidor del PDB (Bernstein *et al.*, 1977) y MolProbity (Davis *et al.*, 2007). Tras obtener una validación correcta, se depositó la estructura en la base de datos de proteínas (Bernstein *et al.*, 1977), obteniendo un código PDB para cada estructura.



Todas las figuras fueron producidas utilizando el visualizador molecular PyMOL v1.5.0.4 (Schrodinger, LLC).

Determinación del estado de oligomerización

El estado de oligomerización se determinó mediante experimentos de cromatografía de exclusión molecular con una columna analítica Superdex 75 HR 10/30 (GE Healthcare Life Sciences ®) acoplada a un sistema FPLC (GE Healthcare Life Sciences ®). Previamente al corrimiento cromatográfico de las TIMs, se realizó una curva de calibración con las siguientes proteínas: albúmina (Peso molecular (PM): 66.3 kDa, Radio de Stokes (R_s): 33.9 Å), ovoalbúmina (PM: 43.5 kDa, R_s : 30.2 Å), anhidrasa carbónica (PM: 28.8 kDa, R_s : 23.6 Å), lisozima (PM: 14.3 kDa, R_s : 20.0 Å), ribonucleasa (PM: 13.7 kDa, R_s : 19.3 Å) y citocromo C (PM: 11.7 kDa, R_s : 17.0 Å). Se graficaron: 1) el volumen de elución contra el R_s (figura 15a) y 2) el volumen de elución contra el logaritmo del PM (figura 15b); ambas series de datos se ajustaron a una regresión lineal y a partir de éstas se obtuvieron los valores necesarios para la determinación del estado oligomérico.

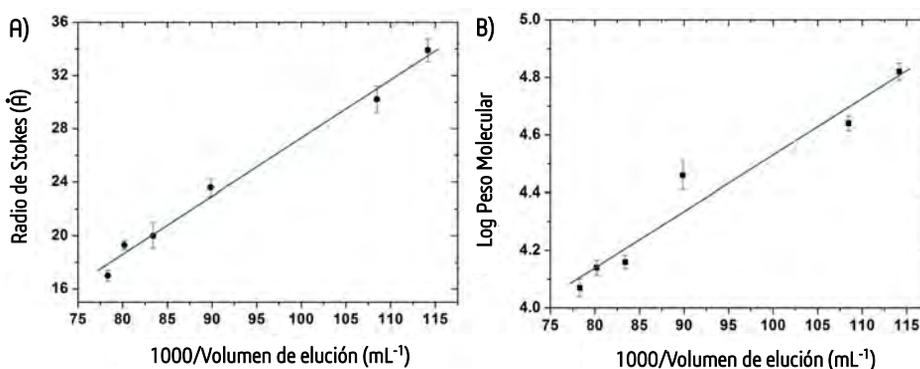
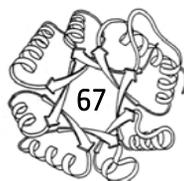


Figura 15. Determinación del estado oligomérico de las proteínas. A) Dependencia de la tasa de migración con el Radio de Stokes de los estándares de calibración utilizados para la determinación del estado de oligomerización. Valor de R^2 : 0.994. B) Dependencia de la tasa de migración con el logaritmo del peso molecular de los estándares de calibración utilizados para la determinación del estado de oligomerización. Valor de R^2 : 0.975. Amortiguador de equilibrio y elución: G.



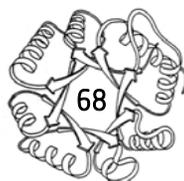
Para el caso de las TIMs bacterianas se utilizó el amortiguador D. La concentración de proteína en los experimentos fue de 0.01, 0.1 y 0.5 mg mL⁻¹ y se realizaron 3 repeticiones para cada una de ellas, las cuales posteriormente se promediaron. Los R_s y PM se determinaron por la interpolación de los volúmenes de elución a la curva de calibración realizada previamente con las 6 proteínas con R_s y PM conocidos. La determinación del estado oligomérico se realizó dividiendo el PM experimental sobre el PM teórico y el número resultante redondeado fue la cantidad de monómeros asociados de la enzima.

Determinación de la actividad enzimática y los parámetros cinéticos

La medición de la actividad catalítica para seguir los pasos de purificación se determinó mediante un ensayo acoplado en la dirección G3P hacia DHAP (Gómez-Puyou *et al.*, 1995). En la transformación de G3P a DHAP por parte de la TIM, se acopla la reducción del DHAP hacia glicerol-3-fosfato catalizada por la α -glicerol 3-fosfato deshidrogenasa (α GDH). La segunda reacción implica la oxidación de una molécula de NADH hacia NAD⁺, este último no absorbe a 340 nm por lo que se observa una disminución en la absorbancia. Así, por cada G3P transformado, se oxida un NADH. El Δ DO min⁻¹ que se obtiene es posteriormente transformado a μ mol de DHAP formado por minuto y miligramo de TIM utilizada en el ensayo usando un coeficiente de absortividad molar a 340 nm (ϵ_{340}) de 6 220 M⁻¹ cm⁻¹ para el NADH. En la figura 16 se muestra los pasos de la reacción acoplada:



Figura 16. Reacción acoplada en dirección G3P a DHAP. La reacción es monitoreada por el decaimiento en la absorbancia del NADH.

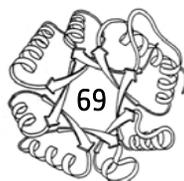


La reacción se llevó a cabo en una mezcla de reacción con: 1 mM de G3P, 0.2 mM de NADH y 0.9 U de α GDH en amortiguador F (100 mM de trietanolamina, 10 mM de EDTA, 1 mM DTT pH: 8.0). La actividad se monitoreó espectrofotométricamente, siguiendo la disminución de la absorbancia del NADH a 340 nm, a una temperatura de 25 °C. Las mediciones se realizaron en un espectrofotómetro Beckman DU7500 termorregulado (Beckman Coulter ®). La concentración de proteína que se utilizó para los ensayos fue de 5 ng mL⁻¹. Los resultados se expresaron como actividad específica con unidades de $\mu\text{mol min}^{-1} \text{mg}^{-1}$.

Para la medición de los parámetros cinéticos, se utilizaron las mismas condiciones anteriores pero variando la concentración de sustrato (G3P) en el intervalo de 0.1 a 3 mM.

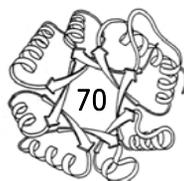
Determinación de la afinidad al PGH

Durante la catálisis se sabe que la reacción de la enzima es acompañada por un intermediario enediolato, lo cual ha sido dilucidado usando análogos del estado de transición como el PGH y el PGA. Los experimentos de afinidad al PGH se realizaron mediante Calorimetría de Titulación Isotérmica en un calorímetro Microcal VP-ITC (Malvern Instruments ®) a una temperatura de 25 °C y en colaboración con la M. en C. Jessica Díaz. Las muestras se equilibraron en el amortiguador F (pH: 7.4) mediante diálisis y la solución fue desgasificada antes de llenar la celda. En la celda del calorímetro se colocaron cada una de las TIMs en concentraciones de 0.15 mM. En la jeringa, se colocó el PGH a concentraciones entre 3 y 6 mM. La titulación consistió de una inyección preliminar de 1 μL de la solución de PGH seguida de 45 inyecciones más (3 μL , 5 μL y 10 μL). Para determinar el calor de dilución del PGH, se realizó la titulación sólo con el amortiguador F en la celda y bajo los mismos parámetros de inyección. Para el cálculo del ΔC_p de unión, los experimentos se hicieron bajo las mismas condiciones pero a 5 temperaturas distintas: 15, 20, 25, 30 y 35 °C.



Disociación por dilución

Para determinar la constante de disociación aparente de las enzimas K_{dis} se realizaron experimentos de estabilidad a la dilución a 15, 20, 25, 30 y 35 °C. Las enzimas se incubaron a 15 diferentes concentraciones de proteína en el intervalo de 500 a 0.007 $\mu\text{g mL}^{-1}$ (2×10^{-4} – 2.8×10^{-10} M, aproximadamente). Tras 24 horas de incubación a cada una de las temperaturas, se midió la actividad catalítica de cada dilución a una concentración final de TIM de 5 ng mL^{-1} en la celda de reacción. Las condiciones para la medición fueron: 0.2 mM de NADH, 2 mM de G3P y 0.9 U de αGDH en amortiguador F. Con la pendiente asociada al proceso, se calculó la actividad catalítica y se normalizó al valor máximo de actividad de acuerdo a cada enzima ensayada, graficando ésta contra la concentración de proteína utilizada en la incubación. Para comprobar la reversibilidad del proceso, se midió la actividad catalítica de las TIMs bacterianas a una concentración en la cual estamos seguros de que es un dímero activo ($100 \mu\text{g mL}^{-1}$). Posteriormente, se diluyó la proteína hasta una concentración de 100ng mL^{-1} en la cual estamos seguros de que la enzima ya no es activa debido a la disociación de la enzima. Después de confirmar la inactivación de las muestras, se concentró la proteína hasta llegar de nuevo a una concentración en la cual previamente sabíamos que era un dímero activo. Así, si comparamos la actividad al inicio del experimento y después de concentrar la proteína diluida, podemos saber si la proteína se volvió a asociar y recuperó su actividad, por lo que el proceso sería reversible completamente. Una falta de recuperación de la actividad sugiere la irreversibilidad del proceso de dilución. Los experimentos se realizaron a una concentración de TIM diluida de 100ng mL^{-1} en un amortiguador E volumen de 1.5 L, con el fin de que, al concentrar 3000 veces la proteína ($\approx 500 \mu\text{L}$ de volumen final), se llegue a una concentración de $100 \mu\text{g mL}^{-1}$.



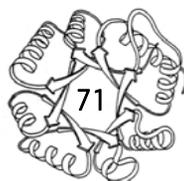
Desplegamiento proteico mediante urea

Los experimentos se realizaron a una concentración de proteína de 0.2 mg mL^{-1} en el amortiguador G con temperatura constante de $25 \text{ }^\circ\text{C}$. Las muestras se desnaturalizaron a distintas concentraciones de urea (0 a 6 M) y se midió el cambio de la señal de la IF. Las mediciones se realizaron en un espectrofluorómetro PC1 (ISS ®) acoplado a un sistema Peltier (PolyScience ®) para controlar la temperatura. Las muestras se excitaron a 295 nm y se midió la señal de fluorescencia de 300 a 450 nm. Para cada muestra se realizaron 5 repeticiones y los valores fueron promediados y posteriormente se restó la señal del amortiguador en cada condición. A la par, se realizaron experimentos de renaturalización de las proteínas para ver la reversibilidad del proceso. Para este tipo de experimentos se incubaron las muestras a las mismas concentraciones de urea utilizadas en los experimentos de despliegamiento, pero en esta ocasión se partió de proteína desnaturalizada a 6 M de urea e incubada previamente durante 12 horas a $25 \text{ }^\circ\text{C}$. Tanto las muestras desnaturalizadas como renaturalizadas se incubaron durante 3 días y se realizaron mediciones intermedias a tiempos de 1, 2, 3, 6, 12, 24, 48 y 72 horas para determinar el tiempo de equilibrio.

Desplegamiento proteico mediante temperatura

Dicroísmo Circular

La desnaturalización por temperatura de las TIMs bacterianas seguida por DC se realizó en un espectropolarímetro Chirascan™ (Applied Photophysics ®) acoplado a un sistema Peltier (PolyScience ®) para el control de temperatura y se determinó siguiendo el cambio en la señal de elipticidad a 195 y 222 nm como función de la temperatura. El proceso se monitoreó tanto al calentar la proteína como al enfriar la misma desde $20 \text{ }^\circ\text{C}$ hasta $90 \text{ }^\circ\text{C}$ con la misma tasa de calentamiento. La



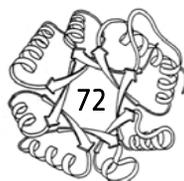
concentración de proteína fue de 0.4 mg mL^{-1} y se probaron velocidades de barrido de 30, 60, 90, 120, 150 y 180 K h^{-1} en un amortiguador G. Antes y después de cada desnaturalización se obtuvo un espectro de DC para observar el cambio en la señal provocado por el proceso de desnaturalización (180-260 nm y en las mismas condiciones explicadas previamente).

Calorimetría Diferencial de Barrido

El proceso de desnaturalización térmica en las TIMs bacterianas también se siguió por DSC, utilizando el calorímetro capilar VP-Capillary DSC (Microcal ®) del Laboratorio de Biofísicoquímica de la Facultad de Química, UNAM en colaboración con el Dr. Miguel Costas Basín. Primero se probó si el proceso de desplegamiento térmico era reversible, para lo cual se calentó la proteína hasta el final de la transición a una velocidad de barrido de 90 K h^{-1} (previamente se hizo una corrida hasta 90 °C para determinar la temperatura de transición de cada proteína). Después de la primera traza, se enfrió la proteína dejando un tiempo de equilibrio de 20 minutos a 20 °C y posteriormente se recolectaron una segunda y tercera trazas para determinar la reversibilidad del proceso. La concentración de proteína fue de 0.4 mg mL^{-1} en amortiguador G.

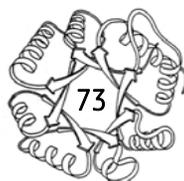
Para los experimentos de variación de la velocidad de barrido, se siguió el proceso de desnaturalización térmica de $20\text{-}90 \text{ °C}$ con 6 velocidades de barrido (30, 60, 90, 120, 150 y 180 K h^{-1}) y una concentración de proteína de 0.4 mg mL^{-1} en amortiguador G. En algunas proteínas (las que presentaron un desplegamiento térmico irreversible) se realizaron trazas calorimétricas en presencia de concentraciones variables de urea (0-2.0 M). Las corridas se realizaron a una velocidad de barrido de 90 K h^{-1} y la concentración de proteína fue de 0.4 mg mL^{-1} en amortiguador G.

Para las proteínas que presentaron un desplegamiento térmico reversible, se realizaron experimentos en el intervalo de $20\text{-}90 \text{ °C}$ a una misma velocidad de barrido (90 K h^{-1}) y variando la concentración de proteína de $0.2\text{-}2.0 \text{ mg mL}^{-1}$ en amortiguador G.



Análisis de secuencia y estructurales

El alineamiento de secuencias se realizó con MAFFT (Kato y Standley, 2013) y visualmente se modificó con Jalview (Waterhouse *et al.*, 2009). El porcentaje de identidad de las secuencias se calculó con el servidor SIAS (Reche, 2013). Se seleccionaron 19 estructuras diméricas para una comparación estructural; para todas estas estructuras se conocía la naturaleza de su desplegamiento térmico (incluyendo las estudiadas en este trabajo). Las características de estas estructuras tridimensionales se resumen en la tabla 6. Los alineamientos estructurales se realizaron utilizando UCSF Chimera (Pettersen *et al.*, 2004) y PyMOL v1.5.0.4 (Schrodinger, LLC). Los valores de RMSD se calcularon con una matriz de distancia utilizando GRASP (Nicholls *et al.*, 1991). La asignación y composición de los elementos de estructura secundaria se determinó con STRIDE (Heinig y Frishman, 2004). Los factores B se determinaron a partir de los datos cristalográficos reportados en cada archivo PDB. Los volúmenes y cavidades moleculares se calcularon con el software 3V (Voss y Gerstein, 2010) y el software MOLE v2.13.9.6 (Berka *et al.*, 2012) usando una ruta de exploración aleatoria y una sonda de radio de 1.5 Å. Las cavidades que tenían más del 70% de la superficie hidrofóbica se clasificaron como cavidades hidrofóbicas. La carga de los túneles y poros se calcularon como la suma de los residuos cargados (Arg, Lys, His: +1; Asp, Glu: -1).





Organismo	Clave	PDB ID	Resolución (Å)	Grupo espacial	Unidad asimétrica	Conformación de las asas del sitio activo		Referencia
						Subunidad A	Subunidad B	
<i>Clostridium perfringens</i>	<i>Cp</i> TIM	4Y8F	1.54	C 1 2 1	Monómero	Abierto	Abierto	Este trabajo
<i>Deinococcus radiodurans</i>	<i>Dr</i> TIM	4Y90	2.10	R 3 2	Tetramero	Abierto	Abierto	Este trabajo
<i>Entamoeba histolytica</i>	<i>Eh</i> TIM	1M6J	1.50	P 2 ₁ 2 ₁ 2	Dímero	Abierto	Abierto	Rodríguez-Romero <i>et al.</i> , 2002
<i>Escherichia coli</i>	<i>Ec</i> TIM	4K6A	1.80	P 2 ₁ 2 ₁ 2 ₁	Dímero	Abierto	Abierto	Misanov <i>et al.</i> , 2013
<i>Gallus gallus</i>	<i>Gg</i> TIM	1TPH	1.80	P 2 ₁ 2 ₁ 2 ₁	Dímero	Cerrado	Cerrado	Zhang <i>et al.</i> , 1994
<i>Gemmata obscuriglobus</i>	<i>Go</i> TIM	4Y96	1.58	P 6 ₅ 2 2	Dímero	Cerrado	Abierto	Este trabajo
<i>Geobacillus stearothermophilus</i>	<i>Gs</i> TIM	1BTM	2.80	P 2 ₁ 2 ₁ 2	Dímero	Cerrado	Cerrado	Delboni <i>et al.</i> , 1995
<i>Giardia lamblia</i>	<i>Gl</i> TIM	2DP3	2.10	I 2 2 2	Monómero	Cerrado	Abierto	Reyes-Vivas <i>et al.</i> , 2007
<i>Homo sapiens</i>	<i>Hs</i> TIM	2JK2	1.70	P 2 ₁ 2 ₁ 2 ₁	Dímero	Abierto	Abierto	Rodríguez-Almazán <i>et al.</i> , 2008
<i>Leishmania mexicana</i>	<i>Lm</i> TIM	1AMK	1.83	C 1 2 1	Monómero	Cerrado	Cerrado	Williams <i>et al.</i> , 1999
<i>Moritella marina</i>	<i>Mm</i> TIM	1AW2	2.65	P 1 2 ₁ 1	Dímero	Abierto	Abierto	Alvarez <i>et al.</i> , 1998
<i>Mycobacterium tuberculosis</i>	<i>Mt</i> TIM	3TA6	1.41	C 1 2 1	Dímero	Abierto	Abierto	Connor <i>et al.</i> , 2011
<i>Oryctolagus cuniculus</i>	<i>Oc</i> TIM	1R2R	1.50	P 2 ₁ 2 ₁ 2 ₁	Dímero	Abierto	Abierto	Aparicio <i>et al.</i> , 2003
<i>Plasmodium falciparum</i>	<i>Pf</i> TIM	1YDV	2.20	C 1 2 1	Dímero	Abierto	Abierto	Velanker <i>et al.</i> , 1997
<i>Saccharomyces cerevisiae</i>	<i>Yf</i> TIM	1NF0	1.60	P 2 ₁ 2 ₁ 2 ₁	Dímero	Abierto	Cerrado	Jogl <i>et al.</i> , 2003
<i>Streptomyces coelicolor</i>	<i>Sc</i> TIM	4Y9A	2.30	P 4 ₃	Dímero	Abierto	Abierto	Este trabajo
<i>Trichomonas vaginalis</i>	<i>Tv</i> TIM	3QSR	2.05	P 2 2 ₁ 2 ₁	Monómero	Abierto	Abierto	Lara-González <i>et al.</i> , 2014
<i>Trypanosoma brucei</i>	<i>Tb</i> TIM	5TIM	1.83	P 2 ₁ 2 ₁ 2 ₁	Dímero	Abierto	Cerrado	Wierenga <i>et al.</i> , 1991
<i>Trypanosoma cruzi</i>	<i>Tc</i> TIM	1TCD	1.83	P 2 ₁ 2 ₁ 2 ₁	Dímero	Cerrado	Abierto	Maldonado <i>et al.</i> , 1998

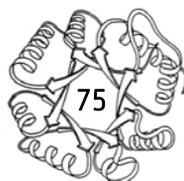
Tabla 6. Archivos PDB dimericos utilizados para la comparación estructural entre RevTIMs e IrrevTIMs.

El área de superficie accesible (ASA) y la diferencia en el área de superficie accesible al plegar y ensamblar la proteína se calcularon con el software PDBePISA v1.51 (Krissinel y Henrick, 2007), Travel Depth (Coleman y Sharp, 2006) y NACCESS (Hubbard y Thornton, 1993) usando un radio de sonda de 1.5 Å. Los cambios en ASA de la disociación de los monómeros se calcularon mediante un modelo de disociación de cuerpo rígido. Para el estado desplegado, la accesibilidad se calculó utilizando un péptido extendido de Gly-X-Gly con el programa VADAR (Willard *et al.*, 2003). Con el programa PDBePISA v1.51 (Krissinel y Henrick, 2007) se estimaron: 1) la energía libre de formación del oligómero, 2) la ganancia energética otorgada por la solvatación al formarse el dímero, 3) el área de la intercara, 4) el área hidrofóbica de la intercara, 5) los puentes de hidrógeno y salinos a través de la intercara y 6) el ASA antes y después de la formación del oligómero asumiendo un proceso de cuerpo rígido. Los enlaces de hidrógeno entre los átomos de las proteínas se calcularon utilizando las rutinas HBPLUS (McDonald y Thornton, 1994) y VADAR con sus parámetros por defecto para distancias y ángulos. Cuando dos átomos de carga opuesta se observaron dentro de 4 Å fueron asignados a un puente salino y cuantificados para todas las proteínas.

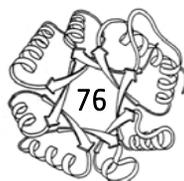
Parte 2: Barriles TIM diseñados *de novo*

Diseño de las proteínas

Los barriles TIM *de novo* se diseñaron utilizando distintos algoritmos de Rosetta, una suite de protocolos para modelo y diseño molecular (Rohl *et al.*, 2004; <https://www.rosettacommons.org/>). Todas las proteínas fueron diseñadas usando a sTIM11-SS como molde (variante de la sTIM11 que no posee cisteínas) con el fin de incrementar la estabilidad de la proteína molde y así utilizar alguno de los diseños como andamiaje para trabajos futuros de diseño de funciones moleculares específicas. El script utilizado para realizar el diseño de los barriles TIM sigue y ejecuta los pasos ordenados por



el algoritmo indicado en la figura 17. De forma general, el algoritmo selecciona primero la simetría con la que diseñará a las proteínas. Una vez que sabe si será una proteína con simetría de orden 2, 4 o asimétrica, elige la cantidad de residuos que mutará (dependiendo si es un cuarto, la mitad o toda la proteína). Después, ejecuta un algoritmo de *simulated annealing* o recocido simulado, el cual es un algoritmo de búsqueda meta-heurística con el fin de encontrar una buena aproximación al valor óptimo de una función en un espacio de búsqueda grande. Para el caso del diseño, lo que busca este paso es realizar una minimización de energía de la proteína que será el molde de los diseños. Posteriormente, realiza un diseño rápido por capas para reempacar el núcleo hidrofóbico de la proteína (en una o varias cavidades seleccionadas), minimizar las restricciones de cadena principal y cadenas laterales y, posteriormente, comparar cada uno de los diseños con el diseño original, checando para esto el valor de RMSD entre ambas proteínas (para el caso de las ReTIMs, el corte fue $<0.7 \text{ \AA}$). Después, depura y filtra los resultados para quedarse con aquellos diseños a los que, con las mutaciones sugeridas, se les logró incrementar el empaquetamiento y mantener la topología de referencia (en este caso, un barril TIM). La selección siguiente se realiza con una puntuación de energía, la cual consiste en elegir a los diseños con el valor más negativo de energía y que tengan el RMSD más bajo posible (ubicados abajo a la izquierda en la gráfica construida de puntuación de energía vs. RMSD). Los diseños que pasan el filtro entran a un paso posterior de *forward folding*, que se encarga de predecir de forma *ab initio* la estructura tridimensional de los diseños para saber si la proteína sugerida se plegará de la forma esperada. Finalmente, se analizan los diseños en un visualizador de proteínas y se seleccionan los candidatos a caracterizar experimentalmente con base a criterios de preferencia personal, ya que la cantidad final de secuencias propuestas generalmente es grande (mayor a 200) y no es posible probar todos ellos.



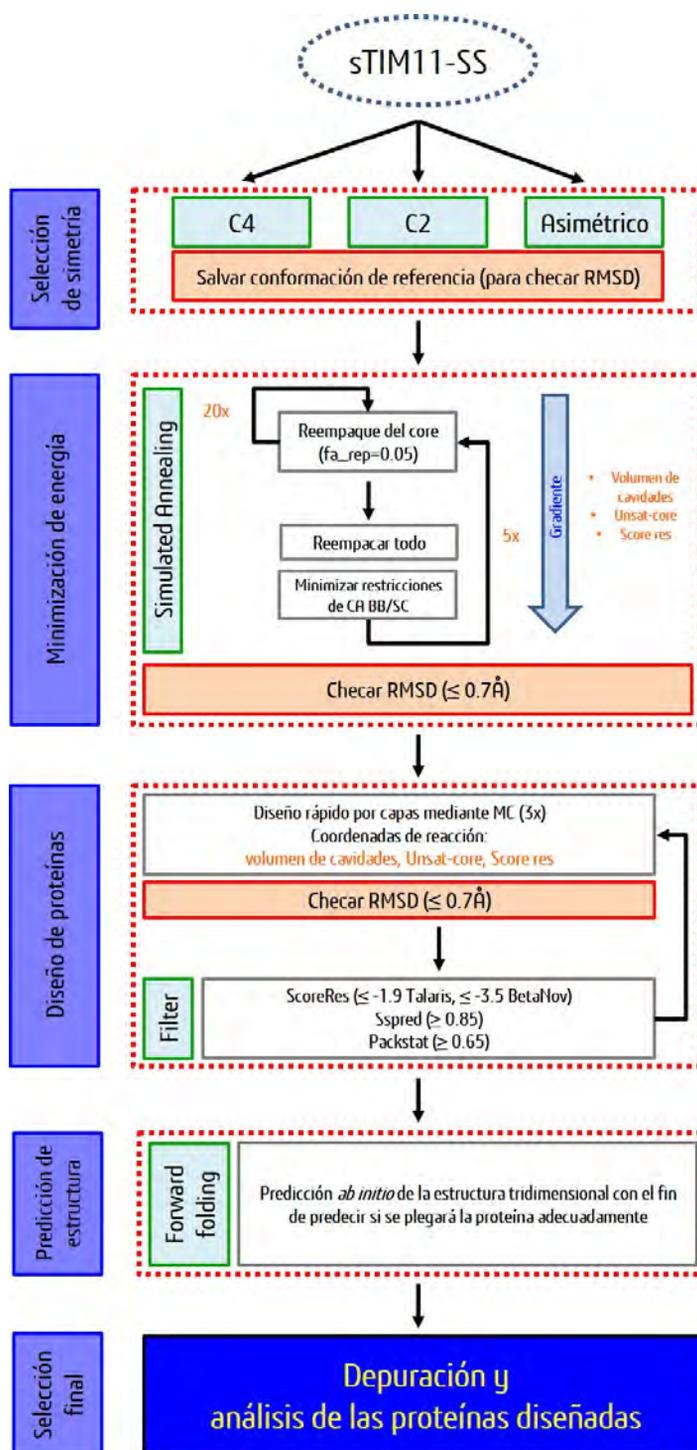
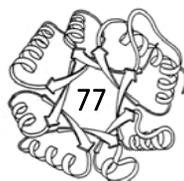
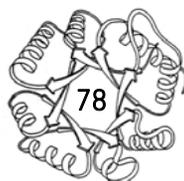


Figura 17. Algoritmo utilizado para el diseño de los barriles TIM. Algoritmo proporcionado por Daniel Silva del laboratorio de David Baker, Universidad de Washington.



Tras cada ciclo del algoritmo mostrado anteriormente, los resultados finales se analizaron individualmente y los modelos se visualizaron para así elegir los candidatos más prometedores en cada ronda de diseños y realizar su caracterización fisicoquímica. Metodológicamente, los diseños de las proteínas se dividieron en 5 grupos principales, todos siguiendo una simetría de orden 4 (es decir, 4 repeticiones de la misma secuencia $\beta\alpha$ - $\beta\alpha$):

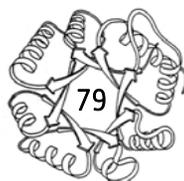
1. Diseños con mutaciones únicamente en el núcleo interno del barril. Esto se realizó porque en la estructura tridimensional de sTIM11-SS se observa una cavidad grande en el centro del barril y por lo tanto se deseó llenar este hueco con el fin de aumentar la estabilidad al incrementar el empaquetamiento de la proteína.
2. Diseños con mutaciones en el núcleo periférico A. Dicha cavidad se observa en la intercara de la hélice 1 y la hebra beta 8 y fue un blanco de los diseños al seguir la misma filosofía de rellenar el espacio, incrementar el empaquetamiento de la proteína y aumentar su estabilidad.
3. Diseños con mutaciones en el núcleo periférico B. Esta cavidad es un espacio grande que se observa entre la intercara de la primera hélice alfa y la primera hebra beta. El principio de diseño es el mismo que en los dos casos anteriores.
4. Diseños con mutaciones en sitios específicos. Se exploraron algunas mutaciones puntuales que el programa sugirió durante el proceso de diseño.
5. Diseños con mezclas de los grupos anteriores. Se mezclaron los mejores diseños de cada grupo y se armó una colección de proteínas diseñadas con el fin de ver la contribución de cada región en la estabilidad de los barriles TIM.



Sobreexpresión de la proteína

Los genes de los barriles TIM diseñados *de novo* fueron insertados en un vector de expresión pET-29b (Novagen[®]) mediante alguna de las dos formas siguientes: 1) utilizando gBlocks™ (Integrated DNA Technologies[®]) con el protocolo proporcionado por el fabricante o 2) subclonando los genes con los sitios de restricción *NdeI* en el extremo 5' y *XhoI* en el extremo 3' mediante una síntesis comercial en GenScript[®]. Las características del vector de expresión utilizado son: 1) regulación por *lacI* lo que permite que pueda ser inducido por un inductor gratuito como el Isopropil- β -D-1-tiogalactopiranosido (IPTG), provocando la transcripción por la ARN polimerasa del fago T7; 2) selección de las células transformantes mediante la resistencia al antibiótico kanamicina; 3) inserción de una etiqueta de histidinas en la región C-terminal de la proteína de interés; y 4) presencia de una secuencia de reconocimiento para trombina (proteasa) en el extremo N-terminal de la región de expresión, lo que permite retirar posteriormente la etiqueta de histidinas si se desea. La inserción de la secuencia de histidinas facilita la detección y posterior purificación de la proteína sobreexpresada (ver más adelante).

Los genes subclonados se transformaron en la cepa de expresión *E. coli* BL21(DE3)p*LysS* (Thermo Fisher Scientific[®]) mediante la técnica de choque térmico. Para analizar la sobreexpresión de las proteínas, inicialmente se realizaron pruebas ajustadas a un volumen pequeño con el fin de detectar si la proteína estaba siendo sobreexpresada. Las cepas se ensayaron con protocolos de sobreexpresión estándares que consistieron en inocular 10 mL de medio LB (suplementado con kanamicina a 30 mg mL⁻¹) con una colonia de bacteria transformada, crecida previamente en medio sólido, e incubar toda la noche a 37 °C con agitación a 250 rpm. A partir de este cultivo, se inocularon 100 μ L a otro tubo de 10 mL de medio LB (suplementado con kanamicina a 30 mg mL⁻¹) y se creció a 37 °C y 250 rpm hasta alcanzar una DO₆₀₀ de ~0.6. En seguida, el cultivo se indujo con 0.8 mM de IPTG y se incubó durante 4 horas a 37 °C y 250 rpm. Se cosechó 1 mL del cultivo mediante centrifugación a 13000 rpm por 15 minutos y las células fueron resuspendidas en líquido para

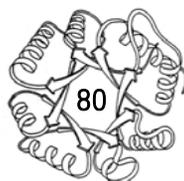


sonicación (Tris base 100 mM, KCl 150 mM, EDTA 1 mM, pH: 7.7) para posteriormente romperse por sonicación en frío con un procesador ultrasónico Cole-Parmer® (3 ciclos de 10 segundos a 30% de amplitud con intervalos de descanso de 10 segundos cada uno). Al terminar la sonicación, las células se centrifugaron a 15000 rpm por 15 minutos a 4 °C para recuperar la proteína ubicada en el sobrenadante. Las pruebas se analizaron en geles de poliacrilamida al 12% en condiciones desnaturalizantes y fueron teñidos con una solución al 0.1% de azul de coomassie G-250 (diluido en 50% de metanol y 10% de ácido acético), cargando tanto el sobrenadante como la pastilla después del centrifugado para observar si la proteína se encontraba en la parte soluble o insoluble del extracto.

Purificación de la proteína

La purificación de las enzimas se realizó en volúmenes grandes mediante el protocolo siguiente: a partir de una colonia transformada se inocularon 10 mL de medio LB (suplementado con kanamicina a 30 mg mL⁻¹) y se incubó toda la noche a 37 °C con agitación a 250 rpm. A partir de este cultivo, se inoculó 1 mL a dos matraces de 1 L que contenían 500 mL de medio LB (suplementado con kanamicina a 30 mg mL⁻¹), respectivamente, y se incubó a 37 °C y 250 rpm hasta alcanzar una DO₆₀₀ de ~0.6-0.8. Posteriormente, los cultivos se indujeron con 0.8 mM de IPTG y se incubaron durante 16 horas a 37 °C y 250 rpm. Las células se cosecharon por centrifugación a 5000 rpm por 20 minutos, y se resuspendieron en el amortiguador A. Posteriormente, las células se rompieron por sonicación en frío con un procesador ultrasónico Cole-Parmer® (16 ciclos de 45 segundos a 30% de amplitud con intervalos de descanso de 45 segundos cada uno) y se centrifugaron a 15000 rpm por 40 minutos a 4 °C para recuperar la proteína ubicada en el sobrenadante.

El sobrenadante recuperado, previamente filtrado con filtros de 0.45 µm, se pasó a una columna de afinidad HisTrap™ HP de 5 mL (GE Healthcare Life Sciences®) acoplada a un sistema

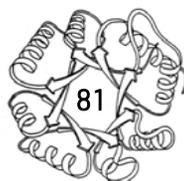


FPLC (GE Healthcare Life Sciences[®]) y equilibrada previamente con amortiguador A. La columna se lavó con 60 mL de amortiguador A y la proteína se eluyó con un gradiente lineal de 100 mL de amortiguador B (gradiente de 0.5 mM de imidazol por cada mililitro de elución). Las fracciones que presentaron una señal de UV a 280 nm mayor a 200 mUA, se colectaron y concentraron para el paso posterior. La proteína colectada se dializó contra el amortiguador G más NaCl 150 mM durante toda la noche y se introdujo a una columna de exclusión molecular Superdex 75 grado preparativo (GE Healthcare Life Sciences[®]) acoplada a un sistema FPLC (GE Healthcare Life Sciences[®]) y equilibrada previamente con amortiguador G. Se utilizó una elución isocrática con el mismo amortiguador presente en la proteína. Las fracciones que presentaron una pureza y concentración de proteína alta, se mezclaron y almacenaron a 4 °C para la utilización en experimentos posteriores. En cada paso de la purificación se obtuvieron alícuotas del proceso para cuantificar la cantidad de proteína, se construyeron las tablas de purificación correspondientes y se identificó el rendimiento final de la purificación.

Mediciones espectroscópicas

Dicroísmo Circular

Los ensayos de DC se realizaron bajo las mismas condiciones en que se estudiaron los barriles TIM naturales.



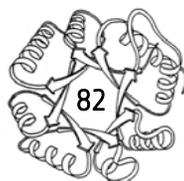
Desplegamiento proteico mediante hidrocloreuro de guanidinio

Los experimentos se realizaron a una concentración de proteína de 0.2 mg mL^{-1} en el amortiguador G con temperatura constante siempre de $25 \text{ }^{\circ}\text{C}$. Las muestras se desnaturalizaron a distintas concentraciones de GdnHCl (0 a 4 M) y se midió el cambio de la señal de la IF. Las mediciones se realizaron en un espectrofluorómetro PC1 (ISS®) acoplado a un sistema Peltier (PolyScience®) para controlar la temperatura. Las muestras se excitaron a 295 nm y se midió la señal de fluorescencia de 300 a 400 nm. Para cada muestra se realizaron 5 repeticiones y los valores fueron promediados y posteriormente se restó la señal del amortiguador en cada condición. A la par, se realizaron experimentos de renaturalización de las proteínas, esto para ver la reversibilidad del proceso y, en caso de que existiera, entonces era posible determinar parámetros termodinámicos. Para este tipo de experimentos se incubaron las muestras a las mismas concentraciones de urea utilizadas en los experimentos de desplegamiento, pero en esta ocasión se partió de proteína desnaturalizada a 4 M de GdnHCl e incubada previamente durante 12 horas a $25 \text{ }^{\circ}\text{C}$. Tanto las muestras desnaturalizadas como renaturalizadas se incubaron durante 3 días y se realizaron mediciones intermedias a tiempos de 1, 2, 3, 6, 12, 24, 48 y 72 horas para determinar el tiempo de equilibrio.

Desplegamiento proteico mediante temperatura

Dicroísmo Circular

La desnaturalización por temperatura de los barriles TIM diseñados seguida por DC se realizó bajo las mismas condiciones en que se estudiaron los barriles TIM naturales.



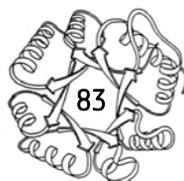
Calorimetría Diferencial de Barrido

El proceso de desnaturalización térmica también se siguió por DSC, utilizando la misma metodología utilizada para los barriles TIM naturales, a excepción de que en algunas proteínas se alcanzó una mayor temperatura de desnaturalización.

Resolución de la estructura tridimensional

La determinación de la estructura tridimensional se realizó mediante la técnica de cristalografía de proteínas con difracción de rayos X en el laboratorio de la Dra. Birte Höcker de la Universidad de Bayreuth, Alemania. Como primer paso, se analizaron las condiciones de cristalización de otras TIMs depositadas en el PDB. Se colocaron pruebas de cristalización para los diseños más interesantes y exitosos a una concentración de 10 mg mL⁻¹ y utilizando 7 distintos kits de Qiagen®: JCSG Core I, JCSG Core II, JCSG Core III, JCSG Core IV, JCSG +, Classics I y Classics II. En total se colocaron 672 pruebas de cristalización por proteína y fueron puestas en incubación a 20 °C en una placa de 96 pozos con la técnica de gota sedente y todas ellas son inspeccionadas cada 3 días para ver la posible formación de cristales. Para todas las condiciones de cristalización, la proteína previamente se dializó en amortiguador G y se centrifugó la muestra a 15 000 rpm durante 30 minutos.

Para las condiciones exitosas donde se observó presencia de cristales, éstos fueron congelados en presencia de una solución crioprotectora compuesta por la misma condición de cristalización más glicerol al 30%. Los cristales se enviaron al Sincrotrón de Suiza (Swiss Light Source) del Instituto Paul Scherrer. Para los cristales que presentaron buena calidad de difracción, se colectó el set completo de datos. La colecta se realizó en la línea PXII del sincrotrón suizo utilizando una longitud de onda de 1.0 Å y con un detector PILATUS 6M (25 Hz) (Dectris®). Después de la colecta de datos, se realizó la integración, reducción y escalamiento de los datos utilizando la misma metodología utilizada para los TIMs bacterianas. La estructura fue resuelta mediante



reemplazo molecular y mediante el programa Phaser (McCoy *et al.*, 2007) de la suite de Phenix (Adams *et al.*, 2010) utilizando el modelo de ReTIM18 arrojado por Rosetta durante el proceso de diseño de la proteína (TFZ: 92.3 y LLG: 5391.3). El afinamiento, construcción y validación se hizo de la misma forma que para los barriles TIM naturales.

Parte 3: Análisis y ajuste de los datos

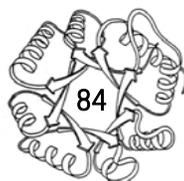
Parámetros cinéticos. Los datos experimentales (velocidades iniciales, V_0), procedentes de las mediciones de actividad catalítica a diferentes concentraciones de sustrato, se ajustaron a una regresión no lineal con un modelo cinético de Michaelis-Menten para obtener los parámetros cinéticos K_m y V_{max} (ecuación 4):

$$V_0 = \frac{V_{max} [S]}{K_m + [S]} \quad (4)$$

Para el cálculo de la eficiencia catalítica, se modificó el valor de la K_m obtenida considerando que sólo el 4% del G3P en solución tiene el sustrato aldehído no hidratado, única forma que la TIM es capaz de catalizar (relación 1:25) (Trentham *et al.*, 1969).

Afinidad al PGH. La constante de unión (K_b), el cambio de la entalpía de unión (ΔH_b) y la estequiometría (n) se determinaron por un ajuste no lineal de los datos normalizados a un modelo de unión con sitios idénticos e independientes (ecuación 5):

$$Q = \frac{nM_t \Delta H_b V_0}{2} \left[1 + \frac{X_t}{nM_t} + \frac{1}{nK_b M_t} - \sqrt{\left(1 + \frac{X_t}{nM_t} + \frac{1}{nK_b M_t} \right)^2 - \frac{4X_t}{nM_t}} \right] \quad (5)$$



donde Q es el calor total de intercambio hasta la adición i , X_i y M_i son las concentraciones de ligando y proteínas totales y V_0 es el volumen de la celda de reacción.

Para el análisis de datos se utilizaron los programas Microcal Origin (Microcal ITC tutorial guide v7.0) y AFFINImeter (<http://www.affinimeter.com/>). Los resultados obtenidos por ambas herramientas son comparables y no presentaron discrepancias significativas.

Desplegamiento proteico mediante dilución. La actividad catalítica normalizada se graficó en contra de la concentración de proteína utilizada en la incubación. La fracción de enzima dimérica (f_N) se expresó en función de la concentración de proteína (P_c) (en moles de dímero) a una temperatura (25 °C) por medio de la ecuación 6 (Cisneros-Armas, 2003):

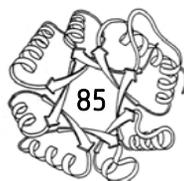
$$f_N(T) = \frac{K_{dis}(T) + 4P_c - \sqrt{K_{dis}(T)^2 + 8K_{dis}(T)P_c}}{4P_c} \quad (6)$$

Como se observó la disociación por dilución a temperatura constante (15-35 °C), la P_c fue la variable independiente y se dejó la K_{dis} (constante de disociación) como parámetro de ajuste. Esto permitió calcular el ΔG_{dis} a cada temperatura por medio de la ecuación 7 (Nájera *et al.*, 2003):

$$\Delta G_{dis(T)} = -RT \ln K_{dis(T)} \quad (7)$$

Desplegamiento proteico mediante urea. Los datos de los barriles TIM naturales se ajustaron a un modelo de dos estados reversible $N_2 \rightleftharpoons 2D$ (utilizando el máximo de emisión de la proteína nativa: ≈ 320 para todas las TIMs) (ecuación 8):

$$f_D = \frac{-1 + \sqrt{1 + \left(8P_c \cdot \exp\left(-\frac{\Delta G + (m \cdot u)}{RT} \right) \right)}}{4P_t \cdot \exp\left(-\frac{\Delta G + (m \cdot u)}{RT} \right)} \quad (8)$$



donde ΔG es la energía libre de desnaturalización en ausencia de urea, m es la pendiente de la transición, P_c es la proteína total en el ensayo expresada en moles de dímero, R es la constante de los gases y T es la temperatura expresada en K.

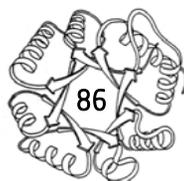
Para el desplegamiento químico de *NpTIM*, los datos experimentales se ajustaron a un modelo de 3 estados con intermediario monomérico: $N_2 \rightleftharpoons 2D' \rightleftharpoons 2D$, donde D' es un intermediario monomérico parcialmente plegado y M es el monómero desplegado (ecuación 9):

$$f_D = \frac{- \left(\frac{1}{\exp\left(-\frac{\Delta G_{pleg} + (m_{pleg} * u)}{RT}\right)} \right) + \sqrt{\left(\frac{1}{\exp\left(-\frac{\Delta G_{pleg} + (m_{pleg} * u)}{RT}\right)} \right)^2 + \left(8P_c * \exp\left(-\frac{\Delta G_{asoc} + (m_{asoc} * u)}{RT}\right)} \right)}{4P_c * \exp\left(-\frac{\Delta G_{asoc} + (m_{asoc} * u)}{RT}\right)} \quad (9)$$

donde ΔG_{pleg} es la energía libre de del proceso de plegamiento de los monómeros en ausencia de urea, m_{pleg} es la pendiente de la transición del plegamiento, ΔG_{asoc} es la energía libre de del proceso de asociación de los monómeros plegados en ausencia de urea, m_{asoc} es la pendiente de la transición en esa parte de la curva, P_c es la proteína total en el ensayo expresada en moles de dímero, R es la constante de los gases y T es la temperatura expresada en K.

Para el caso de los barriles TIM diseñados *de novo*, los datos se ajustaron a una ecuación de dos estados clásica.

Desnaturalización térmica seguida por DSC para las transiciones irreversibles y cálculo de los valores de m cinética en urea y m al equilibrio. La forma de las transiciones calorimétricas se describe adecuadamente por el modelo de dos estados irreversible: $N \rightarrow F$ (Sanchez-Ruiz, 1992; Lyubarev y Kurganov, 1998), donde N es la proteína dimérica nativa y F es el estado final incapaz de replegarse de nuevo a una forma nativa. La conversión cinética de N a F es descrita por una constante de primer



orden (k) que cambia con la temperatura de acuerdo a la ecuación de Arrhenius (Arrhenius, 1889) (ecuación 10):

$$k = \exp \left[-\frac{E_A}{R} \left(\frac{1}{T} - \frac{1}{T^*} \right) \right] \quad (10)$$

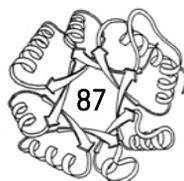
donde T^* es la temperatura a la cual $k=1 \text{ min}^{-1}$ y E_A es la energía de activación del estado nativo al estado de transición. La capacidad calorífica aparente (corregida por la línea base química) en el modelo de dos estados irreversible está dada por (ecuaciones 11 y 12):

$$C_p^{APP} = \frac{\Delta H E_A}{RT_m^2} \exp(x) * \exp[-\exp(x)] \quad (11)$$

$$x = \frac{E_A}{RT_m^2} (T - T_m) \quad (12)$$

donde T_m es la temperatura correspondiente al máximo de la transición, T es la temperatura en cada punto de la endoterma y ΔH es la entalpía de desplegamiento (considerada como constante). Las energías de activación fueron obtenidas siguiendo tres procedimientos. En el primero, las trazas de DSC fueron ajustadas utilizando las ecuaciones 11 y 12 para cada velocidad de barrido empleada, obteniendo valores individuales de E_A (uno para cada velocidad de barrido). En el segundo, la E_A fue obtenida de la pendiente de los gráficos de Arrhenius ($\ln k$ contra $1/T$), empleando todas las velocidades de barrido. Finalmente, la E_A también puede ser obtenida de una prueba de consistencia, evaluando el efecto de la velocidad de barrido (v) en la T_m (Costas *et al.*, 2009) (ecuación 13):

$$\ln \left(\frac{v}{T_m^2} \right) = \text{constante} - \frac{E_A}{RT_m} \quad (13)$$

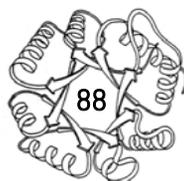


Los valores de la m cinética en urea (m^\ddagger) fueron usados como una medida de la exposición de la proteína (y, consecuentemente, el grado de desplegamiento) al solvente en el estado de transición (Rodríguez-Larrea *et al.*, 2006). Ellos fueron obtenidos usando la ecuación 14:

$$m^\ddagger = -\frac{E_A}{T_m} \left(\frac{dT_m}{d[\text{urea}]} \right) - RT_m \left(\frac{d \ln \left(\frac{E_A}{RT_m} \right)}{d[\text{urea}]} \right) \quad (14)$$

Los valores de la m de urea al equilibrio (m_{eq}) fueron calculados usando una correlación entre la m_{eq} y el ΔASA calculado con los datos estructurales (Myers *et al.*, 1995). Esta correlación fue obtenida a través de una correlación para 45 proteínas en donde se graficó el cambio del área accesible al solvente (ΔASA) en el desplegamiento con la m_{eq} calculada ($\text{cal mol}^{-1} \text{ M}^{-1}$); la relación lineal estuvo dada por $m_{eq} = 368 + (0.11 * \Delta ASA)$, donde el valor de ΔASA utilizado fue un promedio del cálculo para otras TIMs reportadas previamente ($393 \pm 555 \text{ \AA}^2$). El grado fraccional de exposición al solvente en el estado de transición se calculó del cociente entre m^\ddagger/m_{eq} . El número de residuos desplegados en el E_T se obtuvo con la relación: $N_{res} * (m^\ddagger/m_{eq})$, N_{res} es el número de residuos de la proteína.

Desnaturalización térmica seguida por DC y DSC para las transiciones reversibles. La entalpía del estado nativo al estado desplegado de van't Hoff, ΔH_{vH} , de los barriles TIM naturales se obtuvo a partir de un ajuste de la dependencia de la constante de equilibrio con respecto a la temperatura considerando un modelo de dos estados reversible ($N_2 \rightleftharpoons 2D$), donde N_2 representa al estado dimerico nativo y D indica el monómero desnaturalizado. La constante de equilibrio fue calculada con: $K_{eq} = (2f_D^2 P_c)/(1-f_D)$, usando la fracción de proteína desnaturalizada (f_D) obtenida con los datos de DC (ecuación 15):



$$f_D = \frac{(y_N + m_N T) + (y_D + m_D T) \exp \left[\frac{\Delta H_{vH}}{R} \left(\frac{1}{T} - \frac{1}{T_m} \right) \right]}{1 + \exp \left[\frac{\Delta H_{vH}}{R} \left(\frac{1}{T} - \frac{1}{T_m} \right) \right]} \quad (15)$$

donde $y_N + m_N T$ y $y_D + m_D T$ son las ecuaciones del ajuste de las zonas pre y post transición a una recta (y = ordenada al origen, m = pendiente y T = temperatura en K) y T_m es la temperatura a la cual se tiene el 50% de los estados nativos y desplegados, respectivamente. El valor de ΔG en cada punto se calculó por la ecuación 16:

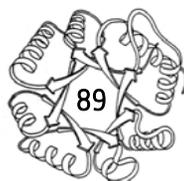
$$\Delta G_{(T)} = -RT \ln K_{eq(T)} \quad (16)$$

Para el caso de los barriles TIM diseñados *de novo*, se utilizaron las mismas ecuaciones pero la constante de equilibrio fue calculada utilizando un modelo de dos estados sencillo $N \rightleftharpoons D$: $K_{eq} = f_D / (1 - f_D)$, usando la fracción de proteína desnaturalizada (f_D) obtenida con los datos de DC.

Las endotermas de DSC fueron ajustadas a un modelo de dos estados usando (Microcal DSC tutorial guide v5.0) (ecuación 17):

$$C_p(T) = B_0 + B_1 T + f(T) \Delta C_p + \frac{\Delta H_{cal}(T)}{RT_m^2} \left[\frac{1 - f(T)}{1 - n + \frac{n}{f(T)}} \right] \quad (17)$$

donde B_0 y B_1 son constantes de la pretransición, n es el número de subunidades presentes en la macromolécula (1 para los barriles TIM diseñados *de novo* y 2 para los barriles TIM naturales) y $f(T)$ es la fracción de proteína en el estado monomérico. Con esta ecuación se obtienen los parámetros ΔH_{cal} , ΔC_p , and T_m . La entalpía de van't Hoff fue calculada mediante $\Delta H_{vH} =$



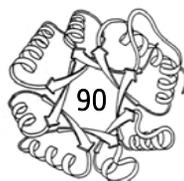
$4RT_{1/2}^2 C_{p1/2} / \Delta H_{cal}$ donde $T_{1/2}$ es la temperatura donde la fracción de proteína nativa y desnaturalizada es igual al 50% y $C_{p1/2}$ es la capacidad calorífica experimental evaluada a esa temperatura. De esta forma, el modelo ajustado para los barriles TIM naturales fue un modelo de dos estados con disociación para una proteína dimérica ($N_2 \rightleftharpoons 2D$) y para el caso de los barriles TIM diseñados *de novo*, los datos se ajustaron a un modelo de dos estados simple $N \rightleftharpoons D$.

El cambio de entalpía ($\Delta H'$) fue también calculado de la dependencia de la T_m con respecto a la concentración de proteína de acuerdo a la ecuación de Takahashi-Sturtevant (Takahashi-Sturtevant, 1981) (ecuación 18):

$$\ln P_c = b - \frac{\Delta H'}{RT_m} \quad (18)$$

donde b es una constante y T_m es la temperatura correspondiente a una proporción de 50% de los estados nativos y desplegados.

Las capacidades caloríficas experimentales para cada una de las 10 concentraciones de proteína ensayadas en los experimentos de DSC fueron convertidas en capacidades caloríficas absolutas utilizando el programa Microcal DSC (Microcal DSC tutorial guide v5.0). La capacidad calorífica absoluta para los estados nativo (C_{pN}) y desplegado (C_{pD}) fueron ajustados a una línea recta y el cambio en la capacidad calorífica en el desplegamiento fue evaluado a la T_m mediante $\Delta C_p = C_{pD} - C_{pN}$ (y posteriormente promediados todos los valores para cada concentración de proteína). El ΔC_p también fue obtenido como un parámetro ajustado usando el modelo de dos estados con disociación para una proteína dimérica (ecuación 17). Finalmente, los valores de ΔC_p predichos fueron obtenidos mediante las ecuaciones paramétricas correspondientes, correlaciones basadas en la secuencia aminoacídica y los cambios en la capacidad calorífica (ecuación 2; Milardi *et al.*, 1997; Robertson y Murphy, 1997) y correlaciones que involucran el ΔASA estructural y los cambios en la capacidad



calorífica que sucede en el proceso de desplegamiento (ecuación 3; Murphy y Freire, 1992; Spolar y Record, 1994; Makhatadze y Privalov, 1995; Myers *et al.*, 1995; Madan y Sharp, 2001).

Curva de estabilidad. Las curvas de estabilidad, $\Delta G(T)$, fueron calculadas usando los parámetros ΔH (ΔH_{ca}), ΔC_p y T_m obtenidos de la ecuación 17 y la ecuación de Gibbs-Helmholtz (ecuación 19):

$$\Delta G(T) = \Delta H \left(1 - \frac{T}{T_m} \right) - \Delta C_p \left(T_m - T + T \ln \left(\frac{T}{T_m} \right) \right) \quad (19)$$

El máximo de estabilidad ocurre a la temperatura T_s donde $d\Delta G/dT = 0 = \Delta S$ y $\Delta H_{T_s} = \Delta G(T_s)$.



RESULTADOS Y DISCUSIÓN

Parte 1: El paisaje conformacional de barriles TIM naturales

Al inicio de esta sección con el fin de dar claridad y continuidad de los resultados, se presentarán algunos datos que se obtuvieron durante la maestría y otros que en su momento eran preliminares y que durante el doctorado se mejoró la calidad de los resultados; todos estos abarcan parte de los primeros 5 incisos de las siguientes secciones, el resto son producto de la investigación doctoral.

Sobreexpresión de las BacTIMs

Para observar la sobreexpresión de las proteínas, inicialmente se realizaron pruebas en volumen pequeño. Tras la realización de éstas, se encontró que 5 de las 6 proteínas se sobreexpresaron de manera correcta y su peso corresponde al esperado de acuerdo a la secuencia aminoacídica (figura 18).

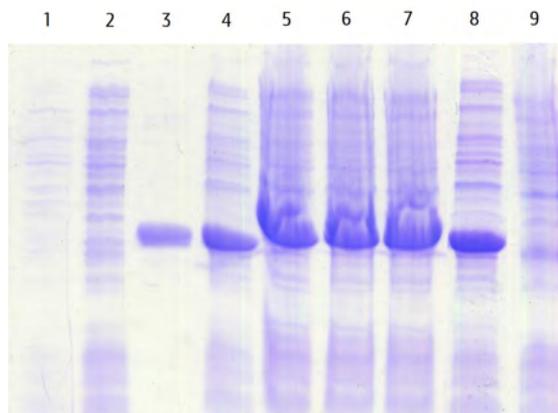
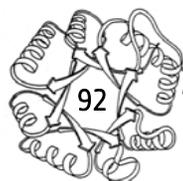


Figura 18. Gel de poliacrilamida SDS-PAGE al 12% con las pruebas de sobreexpresión para las BacTIMs. **1)** Control sin transformar, **2)** control sin inducir, **3)** TIM de *Saccharomyces cerevisiae* (*yfTIM*), **4)** sobreexpresión de *DrTIM*, **5)** sobreexpresión de *NpTIM*, **6)** sobreexpresión de *GoTIM*, **7)** sobreexpresión de *CpTIM*, **8)** sobreexpresión de *ScTIM*, **9)** extracto total después de inducir la cepa que contiene el gen codificante de *AtTIM* donde se observa la falta de sobreexpresión.



Como se puede ver, no fue posible sobreexpresar a *AvTIM* en estas condiciones (figura 18), por lo que se diseñaron diversos protocolos de sobreexpresión. Se realizaron un total de 137 pruebas de sobreexpresión modificando la temperatura de crecimiento, la concentración de IPTG para inducir, el tiempo de inducción, la densidad óptica del cultivo al momento de la inducción y el uso 8 cepas diferentes de *E. coli* para la sobreexpresión. En ninguno de los casos se obtuvo una sobreexpresión visible de la proteína. Debido a la falta de éxito en la sobreexpresión de *AvTIM*, se decidió cambiar el gen a otro vector de expresión. La elección fue el pET-22a (Novagen®) que tiene los mismos sitios de multiclonación que el vector usado previamente, por lo cual se utilizaron las mismas enzimas de restricción para liberar el fragmento, la única diferencia es que el vector presente, no coloca una etiqueta de histidinas como lo hace el pET-28b(+). Se mandaron a secuenciar ambas construcciones y tras el resultado positivo de la secuencia y sabiendo que el gen de interés se encontraba en la posición correcta (en los dos vectores), se probó la sobreexpresión de la enzima, la cual se logró sobreexpresar aunque ésta se depositaba en cuerpos de inclusión (figura 19). Al saber esto, se prosiguió a tratar de solubilizar la proteína, para esto, se probaron varios métodos de desnaturalización-renaturalización sin tener éxito en ninguno de ellos. Por estas razones, se decidió dejar a un lado la purificación de esta enzima y dedicar el trabajo a las otras cinco proteínas que se podían purificar en buena cantidad y de forma homogénea.

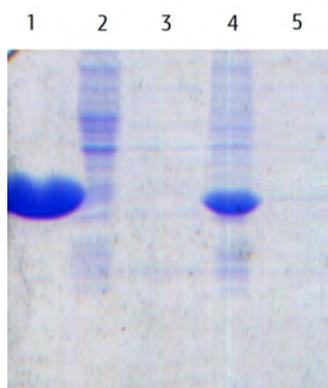
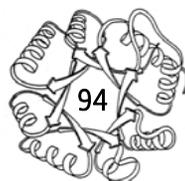


Figura 19. Gel de poliacrilamida SDS-PAGE al 12% con las pruebas de sobreexpresión para *AvTIM*. 1) Marcador de peso molecular (*D**T**M*), 2) sobrenadante del control sin inducir, 3) pastilla del control sin inducir, 4) pastilla de la sobreexpresión de *AvTIM*, 5) sobrenadante de la sobreexpresión de *AvTIM*.

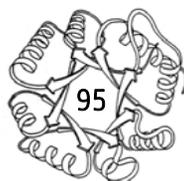
El comportamiento descrito para *AvTIM* nos indica que probablemente la enzima es tóxica para el crecimiento celular, ya que se observó en las diversas pruebas que el crecimiento al parecer se detiene y decae considerablemente comparado con controles sin inducir y con sistemas que producen TIM sobreexpresadas. Como posible explicación para el comportamiento de esta TIM, se realizó el análisis del índice de inestabilidad para las secuencias de todas las TIMs bacterianas estudiadas. El análisis se hizo mediante el método de composición de dipéptidos (Guruprasad *et al.*, 1990), obteniendo una matriz de 400 dimensiones y cuya sumatoria nos da un valor que es indicativo del índice de inestabilidad. Dicho índice de inestabilidad se basa en una valoración totalmente paramétrica, resultado de una colección de enzimas estables e inestables comprobadas experimentalmente, con las que construye un índice basado en la secuencia aminoacídica y se califica posteriormente a la proteína de interés, con base en qué tan semejante o no es a una proteína clasificada como estable, o bien qué tanto se parece a una inestable; valores menores a 40 indican una probabilidad alta para que la proteína sea estable, mientras que una magnitud por arriba de 40 sugiere que la proteína es inestable. Se encontró que de las 6 TIMs bacterianas elegidas, sólo *AvTIM* tuvo un índice superior a 40, lo que indica que la proteína es clasificada como inestable. En el resto de las TIMs bacterianas se obtuvieron índices de inestabilidad bajos (*DvTIM*: 26.5, *NpTIM*: 33.3, *GoTIM*: 34.1, *CpTIM*: 21.2 y *ScTIM*: 15.3). Además, curiosamente si se realiza el análisis de la secuencia clonada en los dos vectores distintos donde se subclonó el gen (pET-28b(+)) y pET-22a), se encuentra que para la secuencia en pET-28b(+) arroja un índice de inestabilidad de 47.6, mientras que para la secuencia en pET-22a, el índice es de 41.3, tal vez esta diferencia haga que en el primer vector no hayamos podido sobreexpresar la enzima contrario al segundo en el cual sí obtuvimos sobreexpresión aunque la enzima forma cuerpos de inclusión. Por último, también se realizó una búsqueda de proteínas parecidas a la secuencia de *AvTIM* en la base de datos BLAST (Madden *et al.*, 1996) con el fin de encontrar semejanzas con otras proteínas de *E. coli* que probablemente estuvieran generando el comportamiento anómalo observado para esta TIM. Como se esperaba, los puntajes más altos de identidad fueron para Triosafosfato Isomerasas, sin embargo, en la lista de



resultados existe una segunda proteína con un valor alto de identidad (66%), la Exodesoxiribonucleasa V (RecB), encargada de hidrolizar cadenas dobles de ADN foráneo como mecanismo de defensa para la bacteria. Se sabe que cuando el complejo de exodesoxiribonucleasa eleva su concentración en la célula, llega a un punto crítico donde el crecimiento se detiene y la célula muere debido a que se activan mecanismos de muerte celular que previenen la captación extrema de ADN foráneo (Seiki *et al.*, 2008; Pavankumar *et al.*, 2010). Con este panorama, se puede hipotetizar que el parecido de *AvTIM* a RecB pudiera generar un fenómeno de toxicidad celular, ya que tal vez la sobreexpresión de *AvTIM* provoque una disminución del crecimiento celular y su posterior muerte mediante mecanismos semejantes a los que activa RecB. En general, se cree que el comportamiento anómalo de *AvTIM* pudiera estar relacionado al menos con estas dos causas, razón por la cual se imposibilitó la purificación de la enzima. Sin embargo, se necesitan más experimentos para comprobar o descartar las hipótesis mencionadas; por ejemplo, se puede utilizar otro sistema de expresión que no contenga estas proteínas con las que tiene alta identidad *AvTIM*.

Purificación de las BacTIMs

Los pasos de purificación de las 5 enzimas se siguieron por cromatografía acoplada a FPLC. Un ejemplo particular se muestra en la figura 20. Los resultados de cada paso de purificación analizados en geles de SDS-PAGE se señalan en la figura 21. Además, también se muestra un ejemplo de la tabla de purificación que se construyó en todo el proceso realizado (tabla 7), en todos los casos se observó un comportamiento similar; el rendimiento fue de aproximadamente 60 miligramos de proteína purificada por litro de cultivo bacteriano.



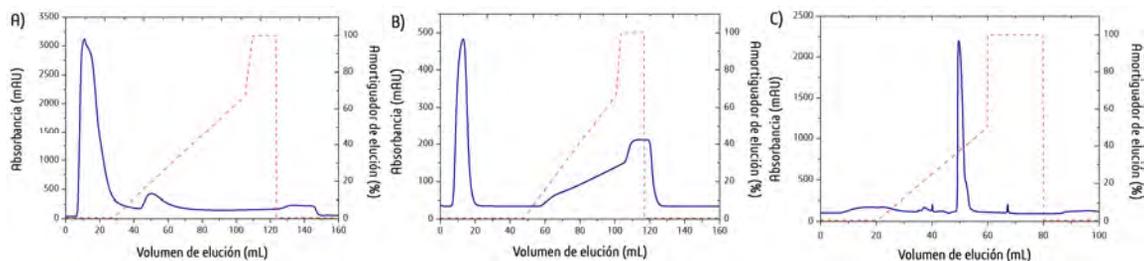


Figura 20. Cromatogramas de la purificación de *DTIM*. **A)** Cromatograma del primer paso de purificación utilizando una columna de afinidad. El pico a los 45-65 mL muestra la elución de la proteína deseada con etiqueta de histidinas. Amortiguador de elución: amortiguador B. **B)** Cromatograma del segundo paso de purificación para *DTIM* utilizando una columna de afinidad. Se muestra el cromatograma al pasar la proteína después de realizar el corte con el kit de trombina. El pico a los 9-20 mL muestra la elución de la proteína deseada sin etiqueta de histidinas. Amortiguador de elución: amortiguador B. **C)** Cromatograma del tercer paso de purificación para *DTIM* utilizando una columna de intercambio aniónico. El pico a los 48-52 mL muestra la elución de la proteína deseada. Amortiguador de elución: amortiguador E.

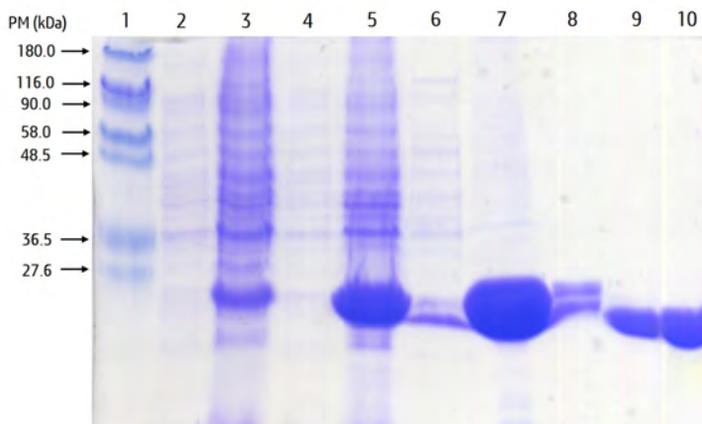


Figura 21. Gel de poliacrilamida SDS-PAGE al 12% con los pasos de la purificación de *DTIM*. **1)** Marcador de peso molecular, **2)** control sin inducir, **3)** extracto total, **4)** pastilla después de sonicar, **5)** sobrenadante después de sonicar (antes de meter a la columna de afinidad), **6)** frente de elución, **7)** TIM con etiqueta de histidinas, **8)** TIM después de cortar con trombina (antes de meter a la columna de afinidad), **9)** TIM sin etiqueta de histidinas (antes de meter a la columna de intercambio aniónico), **10)** TIM pura.

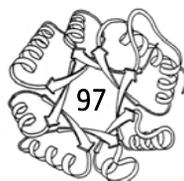
Tabla 7. Tabla de purificación para *DtTIM*.

Paso de purificación	Volumen (mL)	Concentración proteína (mg mL ⁻¹)	Proteína total (mg)	Actividad total (μmol min ⁻¹)	Actividad específica (μmol min ⁻¹ mg ⁻¹)	Número de veces purificada (n)	Rendimiento (%)
Antes de la primera columna de afinidad (Sobrenadante)	9.6	16.71	160.38	112650	702	1.00	100.00
Después de la columna de afinidad (con HisTag)	7.5	6.61	49.54	106513	2150	3.06	94.55
Después del corte con trombina	12.0	2.87	34.41	97914	2845	4.05	86.92
Después de la segunda columna de afinidad (sin HisTag)	10.5	2.35	24.68	96454	3909	5.57	85.62
Después de la columna Mono Q	2.0	5.35	10.70	95854	8955	12.75	85.09

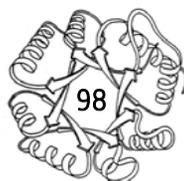
Como criterio de pureza para todas las proteínas estudiadas se realizó la electroforesis en gel de poliacrilamida al 12% sobrecargado de proteína, con el fin de ver algún contaminante en la muestra concentrada. En todas ellas se observó una única banda de proteína, lo que indicó que nuestras muestras estaban puras en un porcentaje mayor al 95%. Además, en el cromatograma del último paso de purificación (columna Mono Q, figura 20c) se puede observar que la proteína eluye en una sola fracción, lo que nos indica indirectamente también que la proteína está pura. También, en la resolución tridimensional de la estructura se pudo comprobar la pureza de la proteína, ya que se pudieron cristalizar las proteínas adecuadamente (ver más adelante).

Nivel estructural secundario y terciario: propiedades espectroscópicas de las BacTIMs

En los espectros de DC de las BacTIMs se distinguen los mínimos característicos para proteínas que poseen hélices α y hebras β , con mínimos en 220 y 210 nm y un máximo en 195 nm (figura 22a). La deconvolución de los espectros de DC realizada con los programas CDNN (Böhm *et al.*, 1992), CDPro (Sreerama y Woody *et al.*, 2004) y DichroWeb (Whitmore y Wallace, 2008) indica porcentajes



de estructura secundaria similares entre todas las BacTIMs, lo que confirma su contenido de estructura secundaria predicho para un barril TIM así como su parecido topológico entre todas las BacTIMs (tabla 8). Los espectros de fluorescencia, tanto en los datos crudos como en su forma normalizada, muestra que las cinco TIMs bacterianas tienen una λ_{\max} de emisión entre 325-330 nm (figura 22b), lo cual nos indica que sus triptófanos están en un ambiente hidrofóbico y ocultos del solvente, lo que implica que la proteína está correctamente plegada. Sin embargo, al comparar los valores máximos de emisión, podemos ver que hay diferencias considerables. Por ejemplo, el máximo de intensidad de fluorescencia de *Sc*TIM es 9 veces mayor que el de *Gø*TIM, a pesar de que esta última tiene un triptófano de más. Al analizar *in silico* un modelo de estas enzimas, los 2 triptófanos conservados están en las asas 1 y 6. Con esto, podemos especular que las diferencias en la señal pueden deberse al apagamiento de alguno de los triptófanos de *Gø*TIM en el estado nativo. De hecho, para otras TIMs (*Gt*TIM y *Tb*TIM), se ha observado que el apagamiento de triptófanos es debido a dos mecanismos principales: 1) interacciones con otros aminoácidos aromáticos debido a la proximidad y orientación de los grupos indol y 2) transferencia de energía de residuos cargados que rodean los triptófanos (Chávez-Cárdenas *et al.*, 2002; Hernández-Alcántara *et al.*, 2008); probablemente en *Gø*TIM pudiera estar sucediendo alguno de estos dos mecanismos.



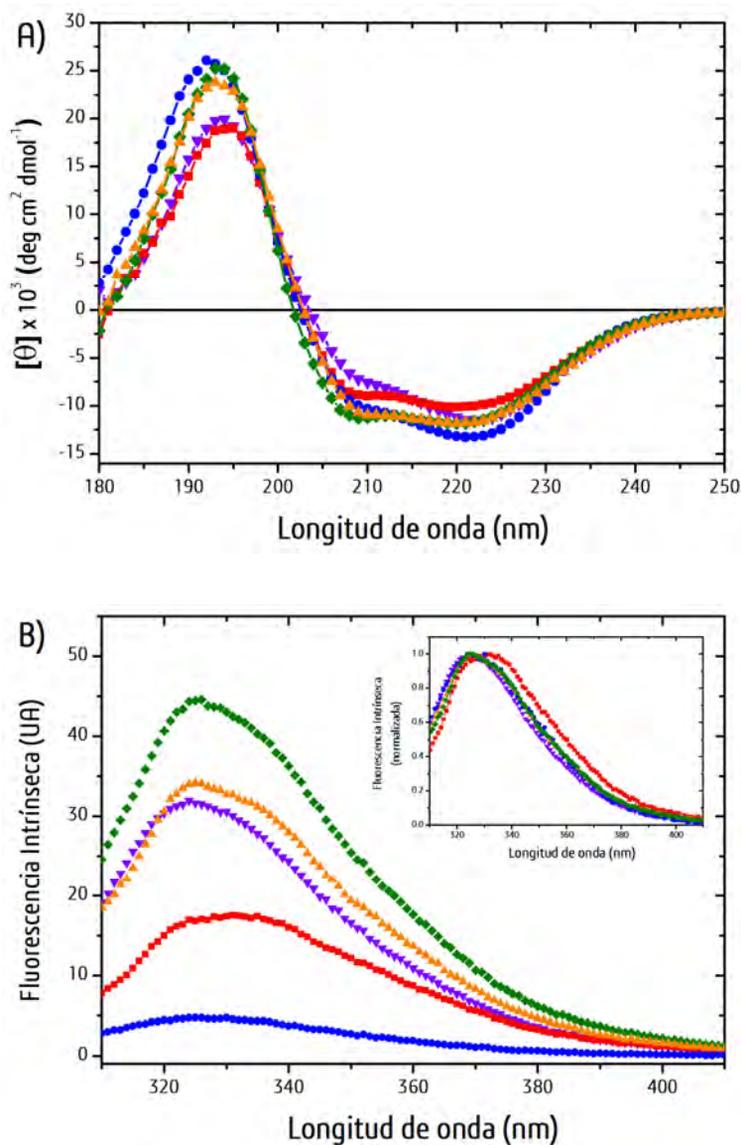


Figura 22. Propiedades espectroscópicas de las BacTIMs. A) Espectro nativo de Dicroísmo Circular. **B)** Espectros nativos de fluorescencia intrínseca (datos crudos). El inserto muestra los datos normalizados. En todos los casos la temperatura fue de 25 °C y en amortiguador G. Para los espectros de DC la concentración de proteína fue de 0.4 mg mL⁻¹ y para los espectros de FI fue de 0.2 mg mL⁻¹. *DpTIM* (--- ▽ ---), *NpTIM* (--- ■ ---), *GoTIM* (--- ● ---), *CpTIM* (--- ◆ ---), *ScoTIM* (--- ▲ ---).

Tabla 8. Contenido de estructura secundaria para las BacTIMs.

TIM	Contenido de estructura secundaria (%)				
	Tipo	Hélice α	Hebra β	Giro β	No repetitiva
D_r TIM	Deconvolución DC	33.3 ± 0.9	16.3 ± 0.7	16.2 ± 0.3	34.3 ± 1.1
	Estructura tridimensional	43.7	21.0	ND	35.3
N_p TIM	Deconvolución DC	36.1 ± 0.6	15.1 ± 0.5	16.3 ± 0.2	32.6 ± 0.8
	Estructura tridimensional	ND	ND	ND	ND
G_o TIM	Deconvolución DC	31.4 ± 1.7	16.8 ± 0.8	15.8 ± 0.5	35.1 ± 0.5
	Estructura tridimensional	44.3	19.9	ND	35.8
S_c TIM	Deconvolución DC	41.3 ± 1.0	13.3 ± 0.9	15.7 ± 0.4	29.7 ± 0.7
	Estructura tridimensional	48.8	19.6	ND	31.6
C_p TIM	Deconvolución DC	33.9 ± 1.9	16.5 ± 0.8	16.8 ± 0.4	32.8 ± 1.4
	Estructura tridimensional	42.9	21.2	ND	36.00

ND: No determinado, ya sea por falta de estructura cristalográfica o por falta de asignación de ese tipo de estructura secundaria en la estructura tridimensional.

Nivel estructural terciario: estructura tridimensional de las BacTIMs

Con la finalidad de establecer una relación energético-estructural entre los distintos niveles estructurales de las BacTIMs y los datos termodinámicos de estabilidad y plegamiento que se determinaron más adelante, se determinaron las estructuras tridimensionales de las BacTIMs mediante cristalización con difracción de rayos X. Las condiciones exitosas de cristalización para las enzimas se muestran en la tabla 9.

Las cuatro estructuras se depositaron en el PDB y en la tabla 10 se muestran sus estadísticas. Como podemos observar, todas las proteínas se determinaron de forma adecuada, por lo que el sistema de validación del PDB no nos arrojó errores graves, sólo algunos residuos fuera de las regiones permitidas del mapa de Ramachandran, pero éstos están contruidos de acuerdo con la densidad electrónica; uno de ellos es la lisina catalítica, la cual se ha visto que regularmente se encuentra fuera de las regiones favorecidas/permitidas del mapa de Ramachandran.

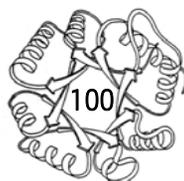


Tabla 9. Condiciones exitosas de cristalización para las BacTIMs.

TIM	Código PBD	Resolución (Å)	Condición de cristalización	Concentración de proteína (mg mL ⁻¹)
<i>Dr</i> TIM	4Y90	2.1	1.8 M Sulfato de Amonio, 0.1 M Citrato de sodio pH: 5.6, 0.2 M Tartrato de sodio	11.5
<i>Go</i> TIM	4Y96	1.58	0.2 M Tartrato de sodio-potasio, 0.1 M Citrato de sodio tribásico pH: 5.6, 2 M Sulfato de amonio	8.5
<i>Cp</i> TIM	4Y8F	1.54	0.2 M Acetato de Amonio, 0.1 M Acetato trihidratado de sodio pH: 4.6, 30% w/v de Polietilenglicol 4000	5.5
<i>Sc</i> TIM	4Y9A	2.3	0.1 M TRIS-HCl pH: 8.5, 2.0 M Sulfato de amonio	3.5

Los datos cristalográficos muestran un estado de asociación dimérico, en concordancia con los experimentos de cromatografía de exclusión molecular. La topología canónica de un barril (β/α)₈ es muy similar en todas las estructuras de las BacTIMs determinadas, con un RMSD menor a 1 Å entre ellas (figuras 23 y 24). La orientación de los monómeros con respecto a su contraparte en el dímero es también similar; de hecho, esta es una propiedad conservada a través de todas las TIMs en los distintos dominios de la vida. El llamado “lado catalítico” de los barriles TIMs está formado por ocho asas que conectan el carboxilo-terminal de las hebras- β con el amino-terminal de las α -hélices (asas β - α). En las BacTIMs, al igual que todas las otras estructuras determinadas, estas asas proveen los residuos catalíticos y la mayoría de las interacciones que estabilizan la intercara monómero-monómero. El otro lado del barril está formado por las asas estabilizantes que conectan el carboxilo-terminal de las α -hélices con el amino-terminal de las hebras- β (figura 24). Las propiedades estructurales y su relación con las propiedades termodinámicas se discutirá en contexto de la reversibilidad del proceso de desplegamiento en una sección más adelante cuando se retoman todos los análisis pertinentes.

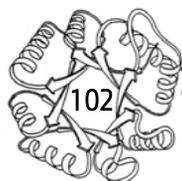


Tabla 10. Estadísticas tras la colecta, integración, reducción, escalamiento y afinamiento de los datos para las BacTIMs.

Colecta de datos ^a				
	<i>Go</i> TIM	<i>Dr</i> TIM	<i>Sc</i> TIM	<i>Cp</i> TIM
PDB ID	4Y96	4Y90	4Y9A	4Y8F
Límites de resolución (Å)	36.36-1.58 (1.64-1.58)	36.13-2.09 (2.17-2.09)	43.03-2.29 (2.37-2.29)	37.59-1.54 (1.59-1.54)
Grupo espacial	P 6 ₅ 2 2	R 3 2	P 4 ₃	C 1 2 1
Dimensiones de la celda unitaria				
a, b, c, (Å)	124.8, 127.8, 134.2	169.6, 169.6, 202.3	86.1, 86.1, 134.0	75.3, 49.6, 71.5
α, β, γ , (°)	90.0, 90.0, 120.0	90.0, 90.0, 120.0	90.0, 90.0, 90.0	90.0, 120.0, 90.0
Contenido en la unidad asimétrica	Dímero	Tetrámero	Dímero	Monómero
Reflexiones totales	779530 (107888)	261090 (36321)	128420 (19245)	142402 (17348)
Reflexiones únicas	81335 (8192)	64962 (6342)	43475 (4241)	33930 (3130)
Multiplicidad	9.6 (9.0)	4.0 (3.9)	3.0 (3.1)	4.2 (3.7)
Integridad (%)	99.3 (97.1)	99.7 (98.6)	99.7 (98.1)	99.2 (94.4)
I / sigma (mean)	29.0 (4.8)	11.0 (2.9)	10.0 (3.0)	14.6 (3.1)
R-merge (%)	4.1 (48.7)	9.7 (46.4)	7.1 (37.1)	6.2 (44.8)
CC1/2 (%)	99.7 (85.2)	98.3 (67.8)	98.1 (69.3)	98.9 (70.6)
Estadísticas del afinamiento				
	<i>Go</i> TIM	<i>Dr</i> TIM	<i>Sc</i> TIM	<i>Cp</i> TIM
R _{work} / R _{free} (%)	17.8 / 20.1	13.8 / 18.6	23.1 / 28.2	14.8 / 18.1
Factor B promedio (Å ²)	24.9	25.9	35.8	16.6
Proteína	23.1	24.8	35.9	14.1
Ligando (crioprotector: glicerol)	32.8	49.3	NA	31.5
Solvente	35.2	33.2	32.7	29.5
Número de átomos	4446	8069	7943	2362
Proteína	3777	7142	7690	1977
Ligando	14	87	NA	11
Solvente	639	832	250	372
Residuos de proteína	500	976	1018	251
RMS (enlaces) (Å)	0.007	0.011	0.009	0.010
RMS (ángulos) (°)	1.11	1.30	1.33	1.24
Ramachandran favorecidos (%)	98.43	97.74	96.15	98.47
Ramachandran permitidos (%)	1.18	1.85	3.56	1.15
Ramachandran outliers (%)	0.39	0.41	0.30	0.38
Clashscore	1.83	1.25	5.41	1.76

^a Entre paréntesis se muestran las estadísticas para la capa de más alta resolución.

NA: No aplica.



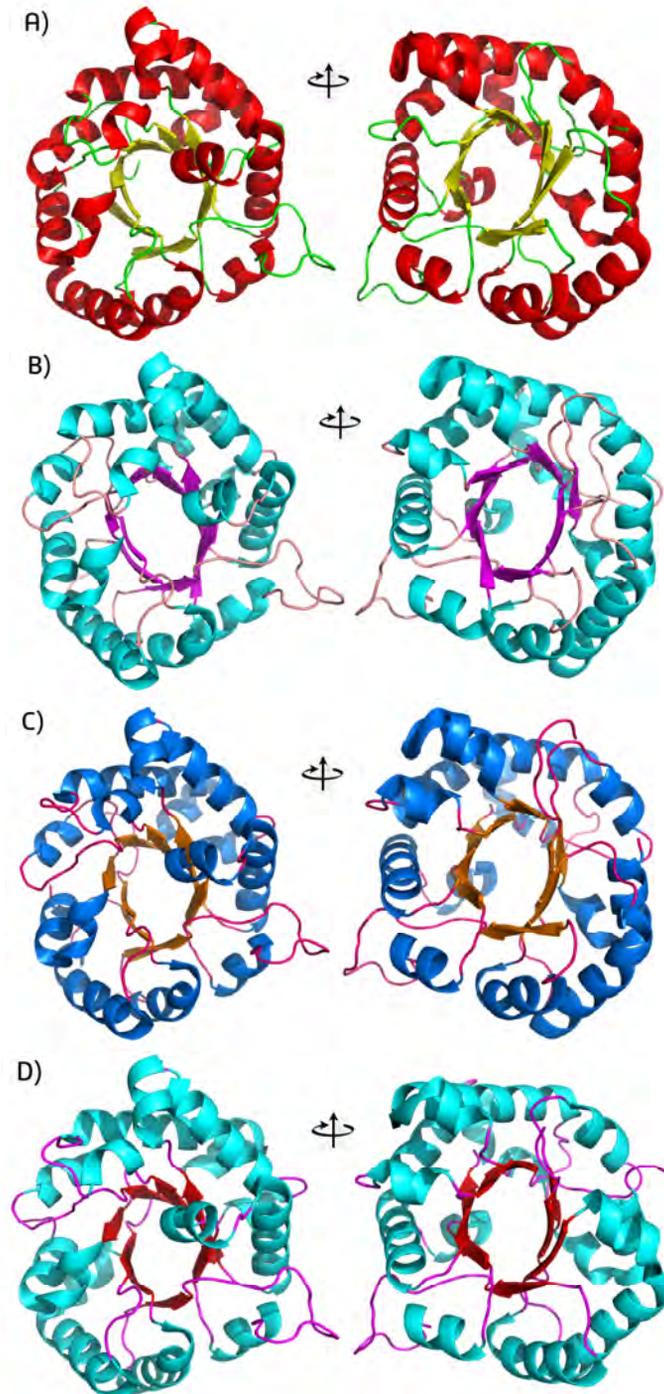


Figura 23. Estructuras tridimensionales de las BacTIMs. Se muestran los elementos de estructura secundaria remarcados con diferente color en cada caso. En la parte derecha se muestra una vista inferior del barril y en la parte izquierda una vista superior. **A)** *Dt*TIM (PDB ID: 4Y90), **B)** *Go*TIM (PDB ID: 4Y96), **C)** *Cp*TIM (PDB ID: 4Y8F), **D)** *Sc*TIM (PDB ID: 4Y9A).

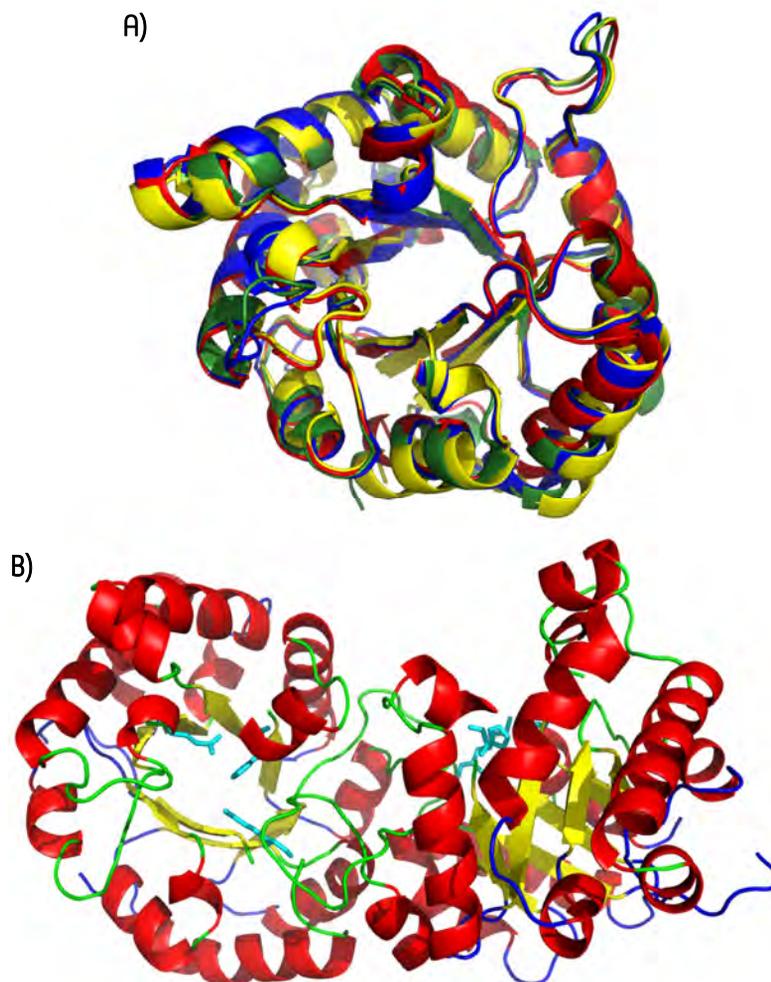
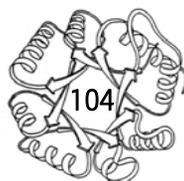


Figura 24. La estructura tridimensional de la TIM. A) Superposición estructural de un monómero de *D*-TIM (amarillo), *G*-TIM (rojo), *C*-TIM (azul) y *S*-TIM (verde). **B)** Representación de la estructura tridimensional del dímero de *C*-TIM. Las asas del lado catalítico son mostradas en verde mientras que las asas de estabilidad se muestran en azul. Los residuos catalíticos son mostrados en cian.

Nivel estructural cuaternario: estado oligomérico de las BacTIMs

Para conocer el estado oligomérico de las enzimas se realizaron experimentos de cromatografía de exclusión molecular. En la figura 25 se muestra como ejemplo el cromatograma de elución *D*-TIM a



3 diferentes concentraciones de proteína (las demás TIMs bacterianas tuvieron un comportamiento igual al mostrado en el ejemplo). A partir de la interpolación de las curvas construidas con las proteínas que sirvieron de calibración se determinó el radio de Stokes y los pesos moleculares experimentales de las BacTIMs. La determinación del estado oligomérico se realizó dividiendo el peso molecular experimental sobre el peso molecular teórico y el número resultante redondeado fue la cantidad de monómeros asociados de la enzima (tabla 11). Al igual que la mayoría de las TIMs mesófilas reportadas, se encontró que todas las BacTIMs se asocian en un estado dimérico. Se sabe que la dimerización para la TIM es un evento clave que permite el plegamiento completo de los monómeros hasta el estado nativo y el correcto posicionamiento del sitio activo (Schliebs *et al.*, 1997). Las causas por las cuales la TIM ha evolucionado hacia dímero, apuntan hacia razones funcionales (Nájera *et al.*, 2003), además de una cuestión de estabilidad (Mainfroid *et al.*, 1996a-b): se hipotetiza que para el caso de dímeros que se pliegan en dos etapas, la dimerización sucedió tempranamente en la evolución debido quizá a una baja estabilidad inicial de los monómeros (D'Alessio, 1999).

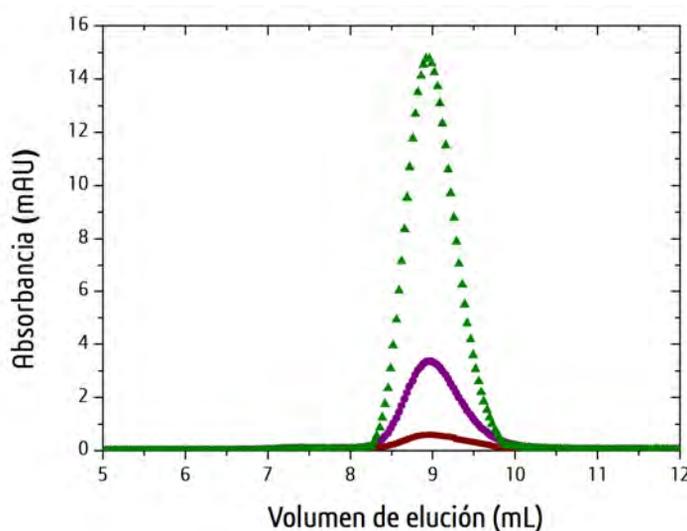


Figura 25. Cromatograma de elución para D/TIM. Se observa que el volumen de elución corresponde a un estado de asociación dimérico. Los resultados detallados de los experimentos se encuentran en la tabla 11. Temperatura: 25 °C y amortiguador D. 10 $\mu\text{g mL}^{-1}$ (--- ■ ---), 100 $\mu\text{g mL}^{-1}$ (--- ● ---) y 500 $\mu\text{g mL}^{-1}$ (--- ▲ ---).

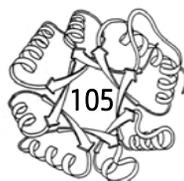
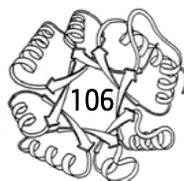


Tabla 11. Estado oligomérico de las TIMs bacterianas.

TIM	Peso molecular teórico (monómero) (Da)	Volumen de elución (mL)	Radio de Stokes (Å)	Peso molecular experimental (Da)	Estado oligomérico (monómeros calculados)
<i>Dr</i> TIM	25593	9.13	31.56	53872	Dímero (2.1)
<i>Np</i> TIM	27053	9.16	31.38	53041	Dímero (2.0)
<i>Go</i> TIM	27025	9.10	31.73	54636	Dímero (2.0)
<i>Cp</i> TIM	27339	9.04	32.03	55993	Dímero (2.0)
<i>Sco</i> TIM	28018	9.07	31.88	55314	Dímero (2.0)

Estudio de la función de la TIM: parámetros cinéticos de las BacTIM por un sustrato natural

Todo lo anteriormente explicado abarca el análisis experimental de los cuatro niveles estructurales de las proteínas: estructura primaria, secundaria, terciaria y cuaternaria. A pesar de que estos niveles ya nos hablan de la naturaleza de la proteína en muy diversos aspectos, una parte esencial de la caracterización proteica es el cómo estos niveles estructurales se relacionan con la función de la proteína, y en el caso de una enzima, como lo es la TIM, con su actividad catalítica. Debido a esto, se realizaron experimentos para medir la actividad de la enzima y determinar los parámetros cinéticos. Los datos experimentales de 3 lotes diferentes de proteínas a los cuales se les calcularon los parámetros cinéticos, se muestran en la figura 26 (curvas de cinética enzimática) y tabla 12 (resultados de los parámetros cinéticos). Los resultados reflejan una firma característica de las TIMs: enzima controlada por difusión (eficiencia catalítica $\approx 10^9 \text{ M}^{-1} \text{ s}^{-1}$). Esta propiedad se aplica cuando una enzima cataliza una reacción química en forma tan eficiente, que el paso limitante de la velocidad es el de difusión; ya sea el de sustrato hacia el sitio activo, o el del producto hacia el medio. En el caso de la TIM su actividad está cerca del límite de la difusión, por lo que es un ejemplo de catalizador perfecto, tal y como se esperaba que fuese desde un inicio. Los valores de K_m de las TIMs analizadas son muy parecidas a algunas TIMs eucariotes reportadas hasta el momento y, aunque



para *Np*TIM y *Go*TIM tienen una V_{\max} menor que las otras tres TIMs, presentan una K_m más afín, por lo que existe una compensación en la enzima y la eficiencia catalítica no se ve severamente alterada.

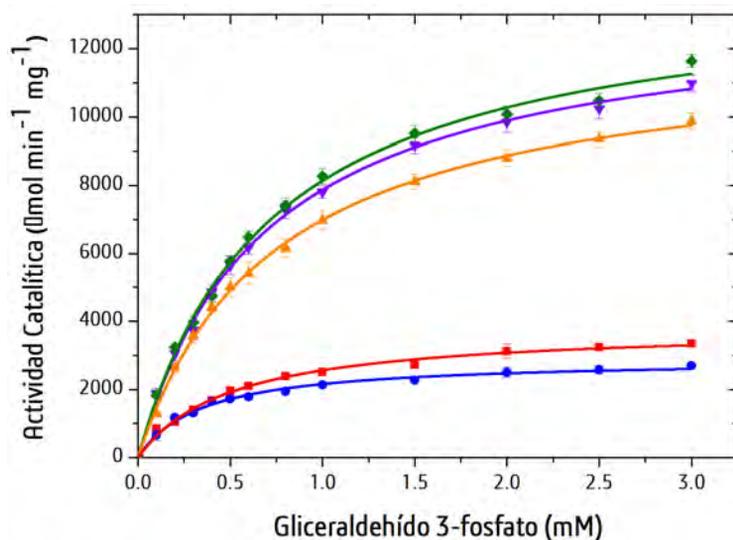
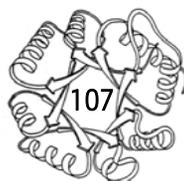


Figura 26. Cinética enzimática de las BacTIMs. Los valores se determinaron a 25 °C, pH 8.0 y diferentes concentraciones de G3P (0.1-3.0 mM); los datos se ajustaron a una regresión no lineal con una cinética tipo Michaelis-Menten. Las barras de incertidumbre corresponden a los 3 diferentes lotes de proteína en los que se determinaron los parámetros cinéticos. *Dr*TIM (--- ▾ ---), *Np*TIM (--- ■ ---), *Go*TIM (--- ● ---), *Cp*TIM (--- ◆ ---), *Sc*TIM (--- ▲ ---).

Tabla 12. Parámetros cinéticos de las TIMs bacterianas.

TIM	K_m (mM)	V_{\max} ($\mu\text{mol min}^{-1} \text{mg}^{-1}$)	k_{cat} (min^{-1})	k_{cat} (s^{-1})	Eficiencia Catalítica ($\text{M}^{-1} \text{s}^{-1}$) [†]
<i>Dr</i> TIM	0.69 ± 0.03	13313 ± 181	$(3.41 \pm 0.05) \times 10^5$	5678 ± 84	$(2.06 \pm 0.18) \times 10^8$
<i>Np</i> TIM	0.51 ± 0.04	3851 ± 89	$(1.05 \pm 0.03) \times 10^5$	1736 ± 42	$(8.63 \pm 0.29) \times 10^7$
<i>Go</i> TIM	0.34 ± 0.02	2901 ± 57	$(7.84 \pm 0.02) \times 10^4$	1306 ± 27	$(9.51 \pm 0.14) \times 10^7$
<i>Cp</i> TIM	0.72 ± 0.04	13930 ± 272	$(3.81 \pm 0.08) \times 10^5$	6348 ± 127	$(2.24 \pm 0.10) \times 10^8$
<i>Sc</i> TIM	0.74 ± 0.02	12137 ± 132	$(3.41 \pm 0.04) \times 10^5$	5668 ± 62	$(1.93 \pm 0.06) \times 10^8$

[†] Realizando la corrección en la K_m al considerar que sólo el 4% del G3P en solución tiene el sustrato aldehído no hidratado, única forma que la TIM es capaz de catalizar (relación 1:25) (Trentham *et al.*, 1969).

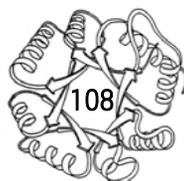


Estudio de la función de la TIM: afinidad de las BacTIMs a un inhibidor

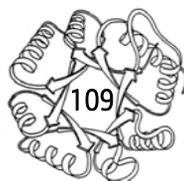
Como parte de una caracterización completa de la estructura-función de la proteína, se efectuaron estudios relacionados con el sitio activo de la enzima. Como se determinaron los parámetros cinéticos de las enzimas, se decidió estudiar el sitio activo utilizando algún inhibidor de la proteína. Actualmente se sabe que el mecanismo catalítico de la enzima involucra un intermediario enediolato, lo cual ha sido dilucidado usando análogos del estado de transición como el fosfoglicolohidroxamato (PGH) y el ácido 2-fosfoglicólico (PGA). Estudios previos han demostrado que la afinidad para el 2-PG es variable, pudiéndose encontrar en un intervalo de 89 mM para *Mm*TIM y de 0.0074 mM para *Hs*TIM. Sin embargo, poco se conoce sobre la afinidad al PGH y no se sabe si existe también esta diversidad en la afinidad al inhibidor. Debido a esto, se eligió este inhibidor para realizar los estudios.

Los experimentos a 25 °C mostraron una estequiometría de 1:1 y los datos se ajustan bien al modelo de un sitio de unión independiente por monómero (figura 27). Al comparar las constantes de afinidad y entalpías de unión, se encontró que son diferentes entre especies (tabla 13). Además, en el caso de *Np*TIM, incluso la firma termodinámica cambia: el ΔS se vuelve positivo y el ΔH disminuye, aunque se realiza una compensación termodinámica y, al final, el ΔG es similar entre todas las proteínas y la unión es siempre controlada entálpicamente (figura 28). Posiblemente, al determinar la estructura tridimensional se logre encontrar cuál es la causa de que la firma termodinámica de unión a PGH de *Np*TIM sea tan distinta a las de otras BacTIMs. Algunos factores que pueden estar interviniendo son el ambiente de los aminoácidos que participan en la unión, las moléculas de agua involucradas en el proceso o el área que se oculta/expone en el proceso de unión del inhibidor; todas estas causas se relacionan directamente con los parámetros termodinámicos que describen la unión de una proteína con su ligando.

Como nos interesa realizar una relación energético-estructural del proceso de unión del inhibidor al sitio activo y ver las diferencias entre distintos organismos, se realizaron experimentos de ITC a diferentes temperaturas con el fin de obtener un ΔC_p y así poder relacionar los cambios en



el área expuesta al solvente durante el proceso de unión del PGH al sitio activo con las estructuras cristalográficas determinadas (ver más adelante). Los experimentos se realizaron en las mismas condiciones que los llevados a cabo a 25 °C pero variando la temperatura (figura 29a). Por problemas de síntesis de PGH y estabilidad del equipo, hasta el momento la determinación del ΔC_p se realizó únicamente para *D*rTIM, sin embargo como perspectiva se propone realizarlo con todas las proteínas. En el caso de *D*rTIM, las isothermas se ajustaron bien al modelo propuesto anteriormente y de ellos se obtuvieron los parámetros de unión. De estos, se regraficó el ΔH con respecto a la temperatura ensayada y de esta gráfica se obtuvo un valor de ΔC_p ($-0.093 \text{ kcal mol}^{-1} \text{ K}^{-1}$) (figura 29b). El valor obtenido está en completa coincidencia con el determinado en estudios previos para *y*rTIM realizados con titulación fluorimétrica y Calorimetría de Titulación Multitérmica: $-0.105 \text{ kcal mol}^{-1} \text{ K}^{-1}$. (Serratos *et al.*, 2011) para la TIM de levadura. Se espera relacionar dicho valor con el que se calculará a través de las estructuras cristalográficas que se determinarán próximamente en presencia de PGH.



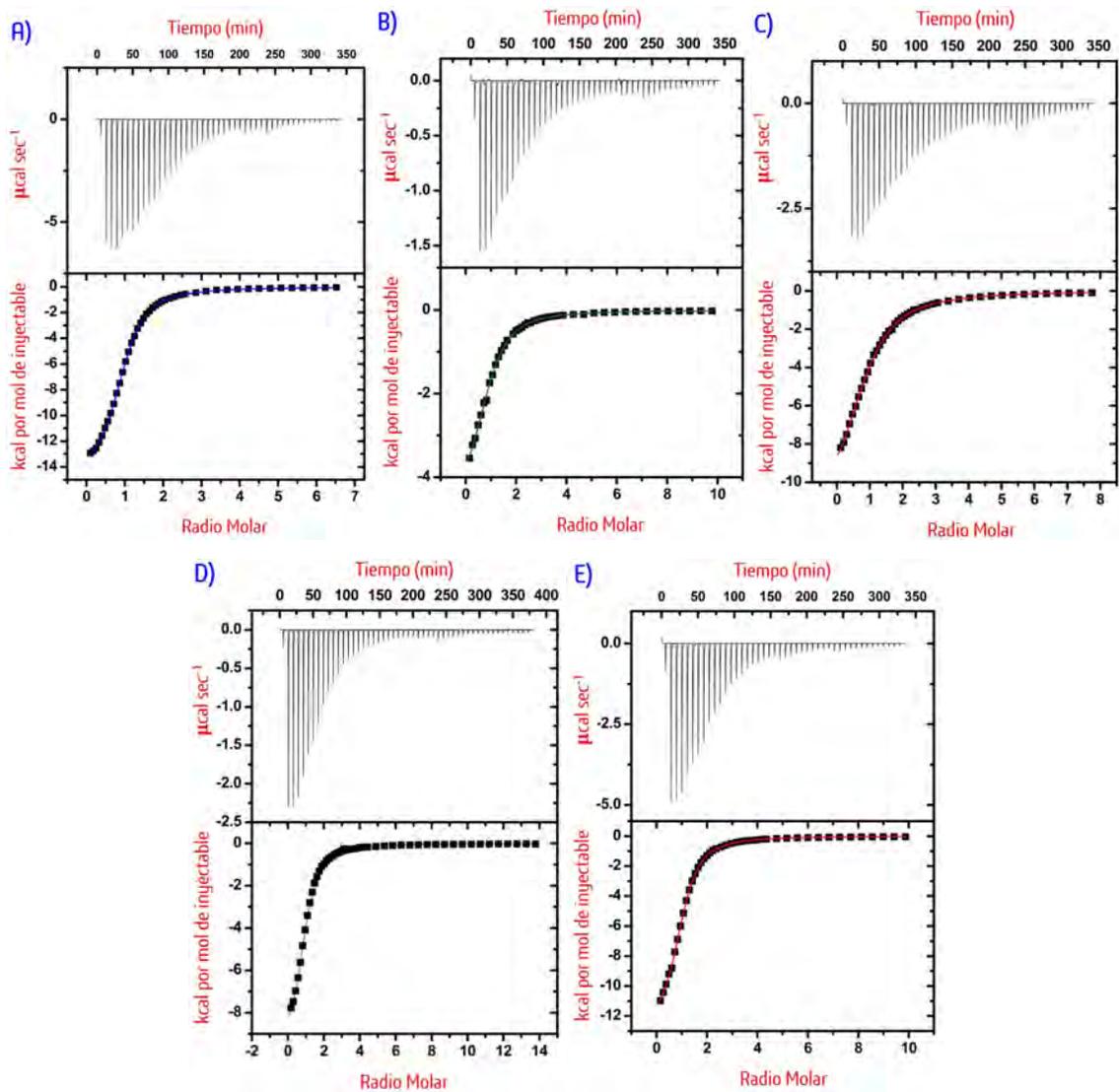


Figura 27. ITC de unión de PGH para las BacTIMs. Termograma de unión a 25 °C (panel superior). Isotherma a partir de la integración de los picos del termograma (panel inferior). Amortiguador F. A) *DpTIM*, B) *NpTIM*, C) *GoTIM*, D) *CpTIM*, E) *SctIM*.

Tabla 13. Determinación de los parámetros termodinámicos de unión TIM-PGH usando ITC.

TIM	N	K (M ⁻¹)	ΔG (kcal mol ⁻¹)	ΔH (kcal mol ⁻¹)	T ΔS (kcal mol ⁻¹)	Referencia
<i>Dr</i> TIM	1	5.61 x 10 ⁴	-6.5	-14.4	-7.9	Este estudio
<i>Np</i> TIM	1	3.96 x 10 ⁴	-6.3	-4.1	2.2	Este estudio
<i>Go</i> TIM	0.98	2.07 x 10 ⁴	-5.9	-10.9	-5.1	Este estudio
<i>Cp</i> TIM	1	7.65 x 10 ⁴	-6.7	-9.4	-2.8	Este estudio
<i>Sc</i> TIM	1.1	6.01 x 10 ⁴	-6.5	-13.1	-6.6	Este estudio
<i>yt</i> TIM	0.96	1.71 x 10 ⁵	-7.1	-14.6	-7.4	Serratos <i>et al.</i> , 2011
<i>Mt</i> TIM	1.24	1.30 x 10 ⁵	-7.0	-9.1	-16.1	Connor <i>et al.</i> , 2011

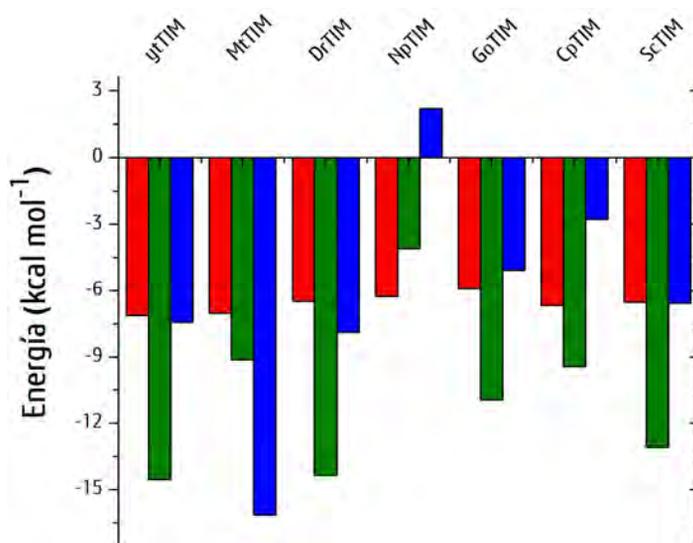
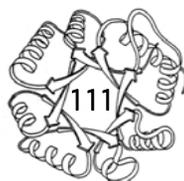


Figura 28. Firma termodinámica de las BactIM para la unión de PGH. Se presentan los parámetros termodinámicos determinados a 25 °C y en comparación a los reportados para *Yt*TIM y *Mt*TIM. (■) ΔG , (■) ΔH , (■) T ΔS .



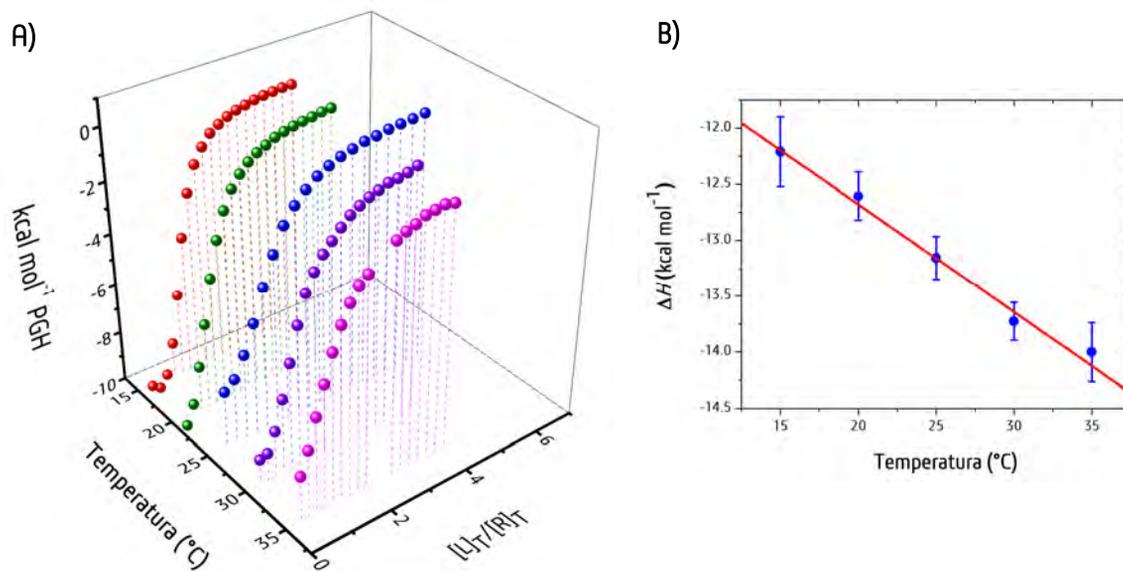
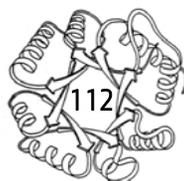


Figura 29. Cálculo del ΔC_p mediante ITC para *D7TIM*. **A)** Isothermas de unión de *D7TIM* a PGH. Se muestran las isothermas a (---●---) 15, (---●---) 20, (---●---) 25, (---●---) 30 y (---●---) 35 °C. Los símbolos representan los calores medidos durante cada adición normalizada por la cantidad de inhibidor añadido. Amortiguador F. **B)** Regráfico de los valores de la entalpía con respecto a la temperatura. El valor del ΔC_p se obtiene a través de la pendiente de la variación lineal del ΔH con respecto a la temperatura (línea punteada roja; $-0.093 \text{ kcal mol}^{-1} \text{ K}^{-1}$). Los símbolos representan las entalpías de unión calculadas a partir de la respectiva isoterma mostrada en el panel A junto con el error del ajuste para cada isoterma.

El desplegamiento de las BacTIMs: desplegamiento mediado por dilución

Se sabe que el desplegamiento de la TIM está asociado a la disociación de los monómeros de la enzima, es decir, que cuando se disocia la proteína, los monómeros se despliegan. Además, debido a que los monómeros son inactivos, la pérdida del estado dimérico puede ser seguida mediante la actividad enzimática después de que la enzima es incubada a diferentes concentraciones de proteína; con estos datos experimentales se puede calcular una constante de disociación aparente (K_{dis}) y ésta es informativa sobre el desplegamiento de la enzima mediado por dilución de la enzima. Para estudiar el desplegamiento de las BacTIMs y calcular la constante de disociación derivada de este proceso, primero se analizó si el proceso era reversible o no, para así saber si la constante calculada era aparente o era una magnitud real. Para comprobar la reversibilidad del proceso, se midió la



actividad catalítica de las TIMs bacterianas a una concentración en la cual estamos seguros de que es un dímero activo. Posteriormente, se diluye la proteína hasta una concentración de 100 ng mL^{-1} en la cual estamos seguros de que la enzima ya no es activa debido a que ésta ya se disoció y perdió la actividad. Al asegurarnos sobre la pérdida de actividad, se concentra la proteína hasta llegar de nuevo a una concentración en la cual previamente sabíamos que era un dímero activo. Así, si comparamos la actividad al inicio del experimento y después de concentrar la proteína diluida, podemos saber si la proteína se volvió a asociar y recuperó su actividad, por lo que el proceso sería reversible completamente. Una falta de recuperación de la actividad sugiere la irreversibilidad del proceso de dilución. La comparación de la cantidad de proteína y actividad catalítica recuperada al finalizar el experimento se puede observar en la figura 30. Como se puede ver, el proceso fue reversible $\approx 90\%$ para *Dt*TIM, *Cp*TIM y *Sc*TIM y es irreversible para *Np*TIM y *Go*TIM, por lo que se podrán considerar como constantes de disociación reales las de *Dt*TIM, *Cp*TIM y *Sc*TIM, mientras que las de *Np*TIM y *Go*TIM serán consideradas como constantes de disociación aparentes.

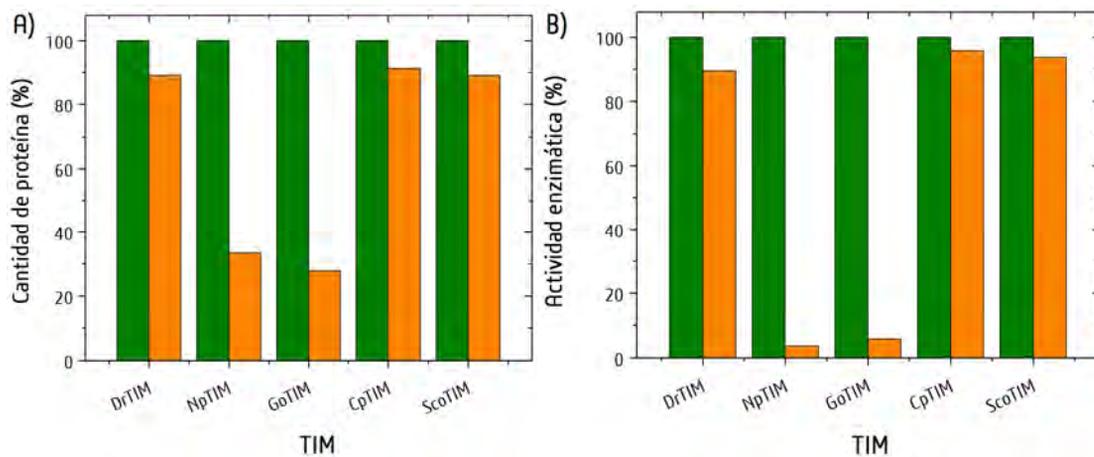


Figura 30. Experimentos de estabilidad a la dilución de las BacTIMs. A) Recuperación de la proteína inicial tras concentrar las distintas TIMs diluidas para los experimentos de estabilidad a la dilución. B) Recuperación de la actividad enzimática tras concentrar las distintas TIMs diluidas para los experimentos de estabilidad a la dilución. En ambas gráficas se indica el comparativo de la cantidad de proteína inicial en el experimento (■) y la cantidad de proteína recuperada tras concentrar las TIMs diluidas (■).

Para determinar la constante de disociación (K_{dis}) de las enzimas, se realizaron experimentos de estabilidad a la dilución a cinco temperaturas distintas: 15, 20, 25, 30 y 35 °C. Se graficó la actividad normalizada contra la concentración de proteína utilizada en la incubación de la cual procedió cada muestra. Con la K_{dis} calculada, se determinó el ΔG_{dis} a cada temperatura ensayada. Las curvas de estabilidad a la dilución para cada proteína a las 5 distintas temperaturas se observan en la figura 31 y los valores de las constantes de disociación y del ΔG_{dis} calculado a partir de esas constantes se pueden comparar en la tabla 14.

Se sabe que la intercara del dímero de la TIM está formada por interacciones no covalentes entre los monómeros; la constante de asociación de las dos subunidades está en el orden de 10^{-9} - 10^{-16} M (Lambeir *et al.*, 1987; Borchert *et al.*, 1993; Mainfroid *et al.*, 1996b; Landa *et al.*, 1997). Tras estos experimentos realizados, el comportamiento de las TIMs bacterianas en estudios de estabilidad a la dilución es muy semejante en términos energéticos. Para todas ellas, los valores de la K_{dis} son mayores a los reportados para otras TIMs silvestres (Mainfroid *et al.*, 1996b; Cisneros-Armas, 2003; López-Velázquez *et al.*, 2004), los cuáles están entre 10^{-9} y 10^{-16} M. A pesar de que la K_{dis} para las TIMs bacterianas fue mayor a la de otras TIMs silvestres, no consideramos que esta característica afecte en nuestros experimentos realizados, en específico en la determinación de los parámetros cinéticos (y otros experimentos dependientes de la actividad enzimática de las TIMs), ya que al realizar la medición de actividad enzimática siempre se observaron trazos lineales, sin signos de inactivación, por lo que aunque las TIMs bacterianas tengan una K_{dis} alta, la cinética de disociación es lenta, lo que permite hacer todos los experimentos sin modificar apreciablemente las poblaciones monoméricas y diméricas. Asimismo, la estabilidad del dímero aumenta con la ocupación de los sitios activos (Garza-Ramos *et al.*, 1992; Jiménez *et al.*, 2003), una razón más por la cual no se observa inactivación en las mediciones de actividad enzimática. Todos los experimentos de estabilidad de las BacTIMs se llevaron a cabo con concentraciones de proteína mayores a la K_{dis} , debido a esto, podemos tener la certeza de que en todos los experimentos se trabajó con dímeros activos. Más adelante se



discutirán y retomarán los valores de ΔG_{dis} determinados por los experimentos de dilución, esto cuando se comparen todos los valores de energía libre de Gibbs obtenidos (químico y temperatura).

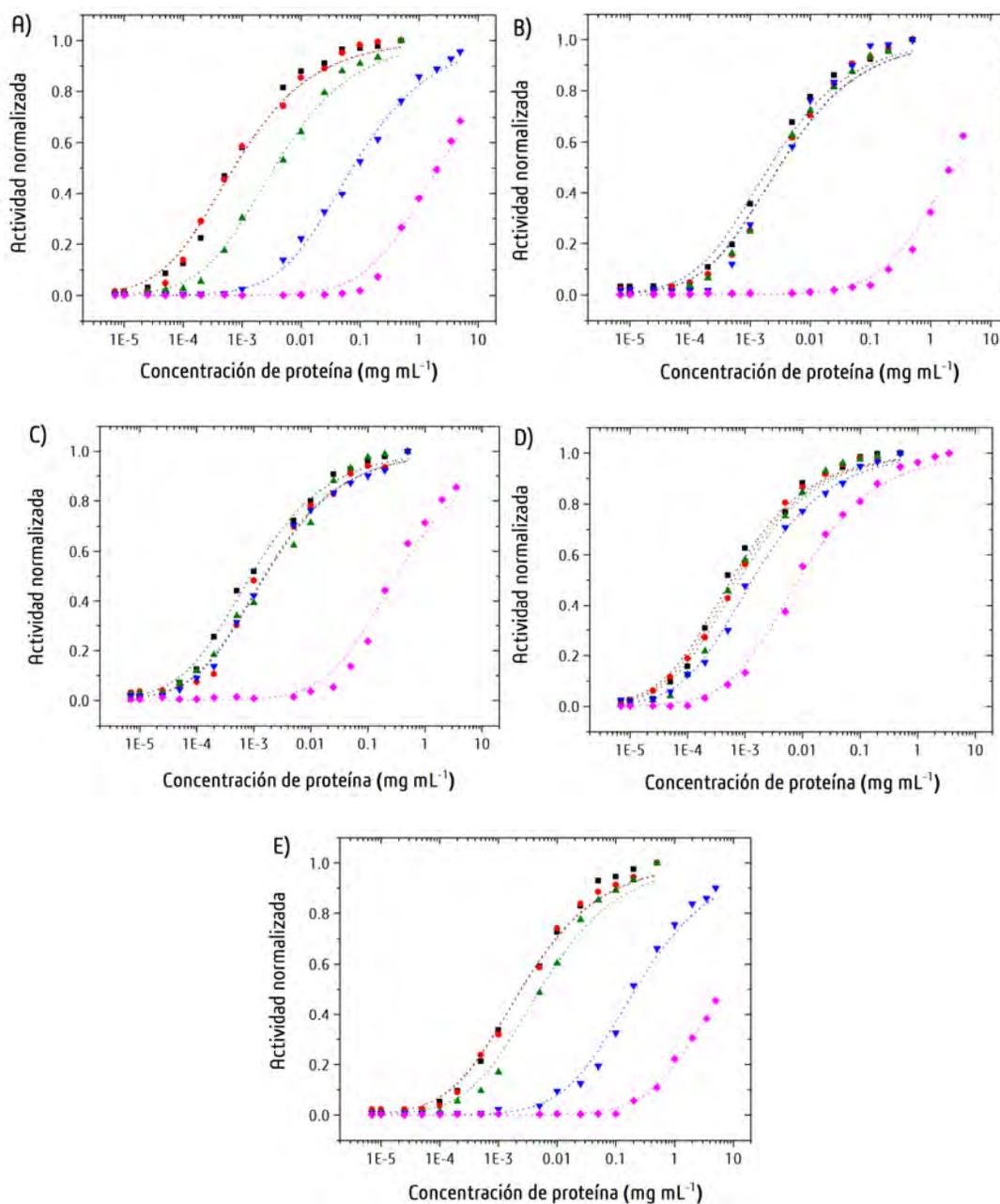


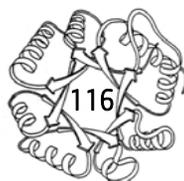
Figura 31. Despliegamiento mediado por dilución para las BacTIMs. Se indica la fracción de dímeros activos en función de la concentración de proteína a la que se hicieron. En símbolos se muestran los datos experimentales y en la línea punteada, el ajuste realizado con la ecuación 6. A) *Dt*TIM, B) *Go*TIM, C) *Np*TIM, D) *Cp*TIM y E) *Sc*TIM. 15 °C (-- ■ --), 20 °C (-- ● --), 25 °C (-- ▲ --), 30 °C (-- ▼ --), 35 °C (-- ◆ --).

Tabla 14. Parámetros termodinámicos derivados de los experimentos de dilución para las BacTIMs.

Temperatura (°C)	<i>Dr</i> TIM	<i>Np</i> TIM	<i>Go</i> TIM	<i>Cp</i> TIM	<i>Sc</i> TIM
K_{dis} (nM)					
15	0.28 ± 0.01	0.33 ± 0.02	0.77 ± 0.04	0.21 ± 0.01	0.86 ± 0.06
20	2.85 ± 0.14	5.47 ± 0.34	10.92 ± 0.98	2.41 ± 0.16	8.89 ± 0.74
25	13.56 ± 1.01	5.66 ± 0.79	10.99 ± 0.87	2.89 ± 0.17	18.27 ± 0.98
30	288.13 ± 14.16	5.88 ± 0.62	11.17 ± 1.12	4.90 ± 0.36	765.91 ± 38.62
35	8204.66 ± 533.25	1169.24 ± 98.51	9338.02 ± 687.96	30.98 ± 1.23	23192.61 ± 1352.21
ΔG_{dis} (kcal mol ⁻¹)					
15	9.96 ± 0.65	9.87 ± 0.51	9.38 ± 0.42	10.13 ± 0.49	9.32 ± 0.28
20	10.12 ± 0.63	9.74 ± 0.11	9.34 ± 0.89	10.22 ± 0.20	9.46 ± 0.37
25	9.37 ± 0.25	9.89 ± 0.61	9.49 ± 0.22	10.29 ± 0.69	9.19 ± 0.93
30	7.69 ± 0.15	10.03 ± 0.18	9.64 ± 0.34	10.14 ± 0.77	7.10 ± 0.52
35	5.76 ± 0.12	6.95 ± 0.32	5.68 ± 0.74	9.18 ± 0.39	5.12 ± 0.49

El desplegamiento de las BacTIMs: desplegamiento mediado por urea

Con el fin de comparar la estabilidad de las BacTIMs con el resto de las TIMs estudiadas previamente, además de determinar la vía de plegamiento con un agente químico desnaturizante, se realizaron los experimentos de desnaturización utilizando a la urea como agente caotrópico. En cuatro de ellas (*Dr*TIM, *Np*TIM, *Go*TIM y *Sc*TIM) se observó que el proceso de desnaturización es reversible, ya que el espectro de fluorescencia se recupera al renaturalizar la proteína y es casi idéntico al de la proteína nativa (figura 32a-d). No se pudieron realizar los experimentos a profundidad para *Cp*TIM, ya que, en todas las condiciones ensayadas, no se observó un cambio en la fluorescencia (figura 32e), debido a esto se descartó la proteína para su caracterización utilizando este agente químico; se propone seguir este desplegamiento DC.



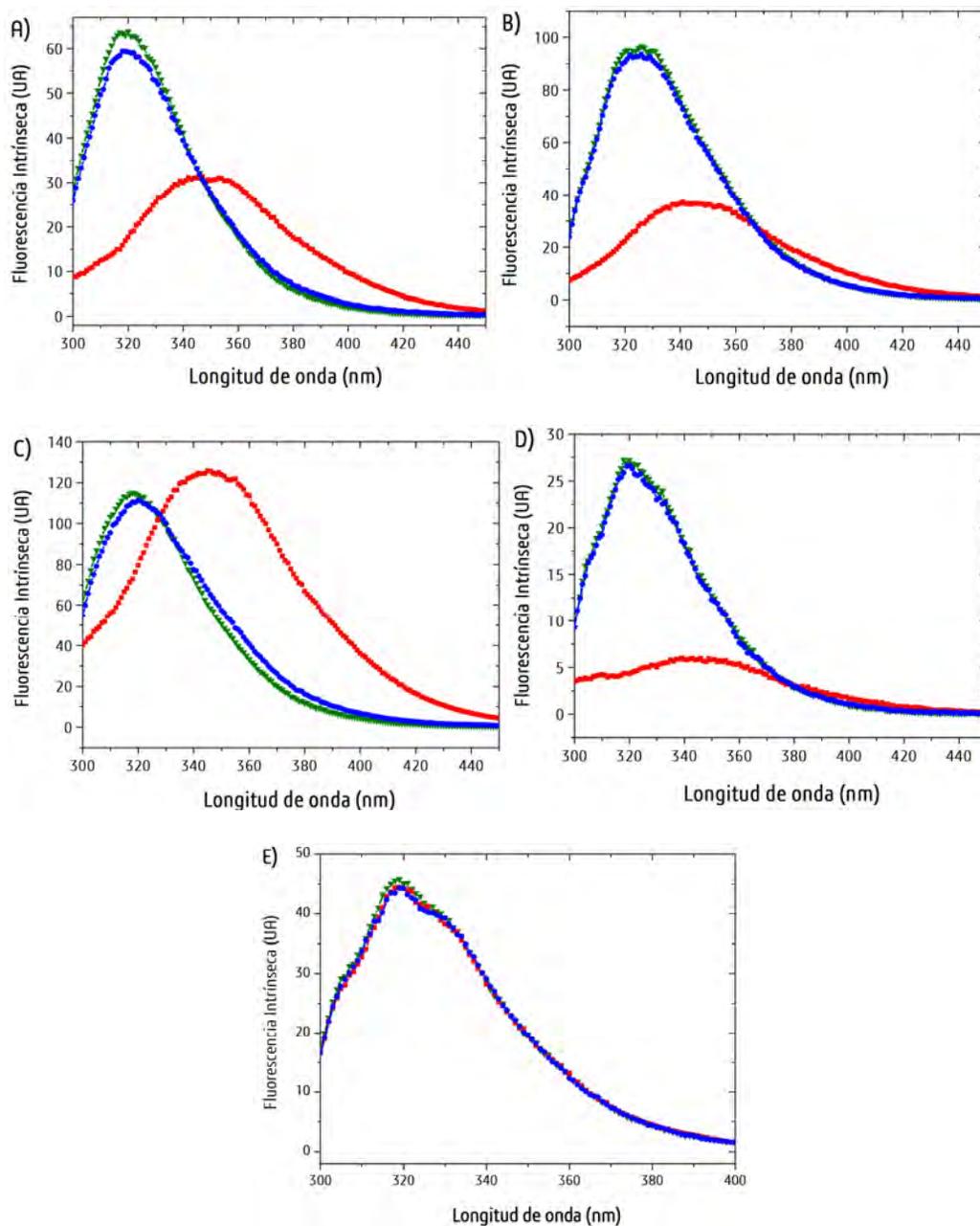
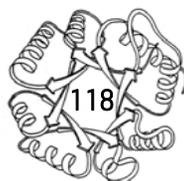


Figura 32. Experimento de reversibilidad para el desplegamiento químico de las BacTIMs. Se muestra la prueba de reversibilidad siguiendo las propiedades espectroscópicas de la proteína nativa (triángulos verdes), la proteína desnaturalizada (cuadros rojos) y la proteína renaturalizada a casi 0.3 M de urea (círculos azules). **A)** D α TIM, **B)** N ρ TIM, **C)** G σ TIM, **D)** S ζ TIM y **E)** C ρ TIM. En todos los experimentos la cantidad de proteína utilizada fue de 0.8 μ M en amortiguador G y la temperatura utilizada fue de 25 $^{\circ}$ C.

Con las pruebas de reversibilidad exitosas, se determinó el tiempo de equilibrio del desplegamiento de las BacTIMs; en todos los casos, 24 horas fueron suficientes para no observar un cambio con el tiempo en la señal espectroscópica y ese fue el tiempo que se utilizó en los experimentos posteriores. Enseguida, se realizaron las curvas de desplegamiento para ajustar los datos al modelo respectivo que describiera la vía de plegamiento observada. Las curvas de desplegamiento mediado por urea se muestran en la figura 33. En el caso de *D α* TIM, *Sc*TIM y *Go*TIM se observó que el desplegamiento es un proceso de dos estados sin intermediarios estables al equilibrio en las condiciones estudiadas; para *Np*TIM se observa la presencia presuntiva de un intermediario monomérico (que se determinó tras variar la cantidad de proteína en el experimento, lo que provoca un movimiento a la derecha de la segunda primera transición). De estos experimentos y ajustes, se obtuvieron los parámetros termodinámicos reportados en la tabla 15. Más adelante se discutirán y retomarán los valores de ΔG determinados por los experimentos de desplegamiento químico.



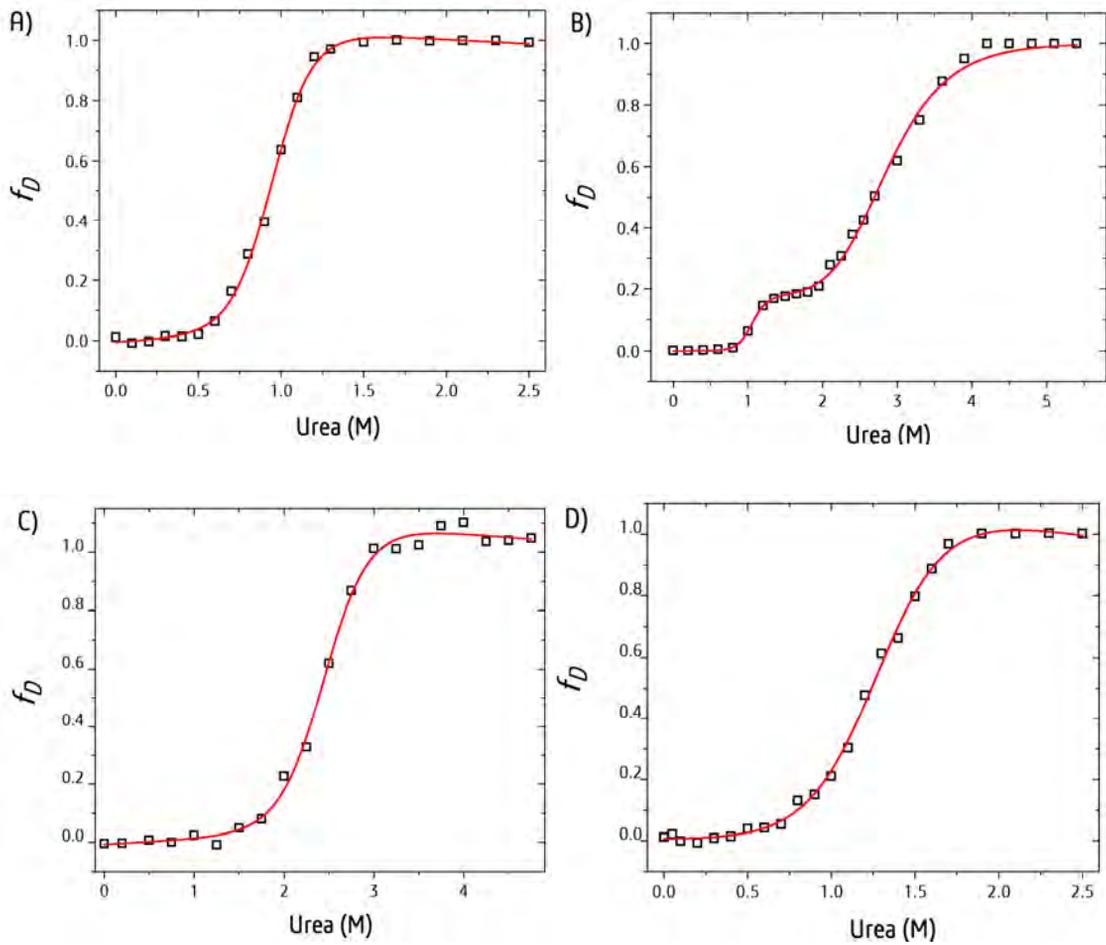
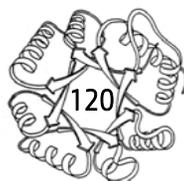


Figura 33. Curvas de desplegamiento químico de las BacTIMs. Se muestra la fracción de proteína desnaturizada obtenida de los experimentos de desnaturización por urea seguidos por fluorescencia intrínseca a 320 nm. La línea continua es el ajuste a un modelo de dos estados ($N_2 \rightleftharpoons 2D$) para **A)** D7TIM, **C)** G7TIM y **D)** S7TIM y a un modelo de tres estados con intermediario monomérico ($N_2 \rightleftharpoons 2D' \rightleftharpoons 2D$) para **B)** NpTIM. En todos los experimentos la cantidad de proteína utilizada fue de 0.8 μ M en amortiguador G y la temperatura utilizada fue de 25 $^{\circ}$ C.

Tabla 15. Parámetros termodinámicos del desplegamiento químico de las BacTIMs.

TIM		ΔG (kcal mol ⁻¹)	C_m (M)	m (kcal mol ⁻¹ M ⁻¹)
<i>Dr</i> TIM		-9.18 ± 0.51	0.93 ± 0.04	-9.71 ± 0.52
<i>Go</i> TIM		-23.48 ± 1.15	2.39 ± 0.08	-4.79 ± 0.46
<i>Sc</i> TIM		-7.39 ± 0.45	1.23 ± 0.02	-5.78 ± 0.35
<i>Np</i> TIM	Plegamiento	-3.66 ± 0.28	1.22 ± 0.09	-1.28 ± 0.40
	Asociación	-15.55 ± 1.06	3.11 ± 0.12	-7.82 ± 1.09
	Total	-19.21 ± 1.14		

Se sabe que la vía de plegamiento no está conservada evolutivamente en las TIMs, presentando una diversidad en sus mecanismos. En el caso de *Dr*TIM, *Go*TIM y *Sc*TIM, se encontró ser un proceso de dos estados, al igual que ha sido descrito para *Gs*TIM, *Oc*TIM, *Gg*TIM, *Lm*TIM y *Hs*TIM. Para *Np*TIM se encontró que su vía de plegamiento es de tres estados con un intermediario monomérico, igual que había sido descrito para la de *yt*TIM y *Et*TIM. Se sabe que la primera etapa del plegamiento de la TIM involucra la formación de intermediarios monoméricos, tal como se observa en la vía de plegamiento de *Np*TIM al equilibrio. Se han reportado este tipo de intermediarios para la desnaturalización con GdnHCl de *Tc*TIM (Chávez-Cárdenas *et al.*, 2005), *Tb*TIM (Chávez-Cárdenas *et al.*, 2002; Guzmán-Luna y Garza-Ramos, 2012), *yt*TIM (Morgan *et al.*, 2000; Nájera *et al.*, 2003), *Et*TIM (Vázquez-Pérez y Fernández-Velasco, 2007) y *Tv*TIM (Lara-González *et al.*, 2014). Algunas de las características comunes a todos ellos son que conservan estructura residual, son inactivos y expandidos. Posterior a la formación de los monómeros ocurre la asociación y los rearrreglos estructurales necesarios para la consolidación del estado nativo. Es interesante que la diversidad de vías de plegamiento se conserva en enzimas derivadas de organismos bacterianos, lo cual nos habla que probablemente la diversidad de hábitats pueda estar relacionado con la vía de plegamiento que presenta una TIM en específico o que esta diversidad también estuviera presente en proteínas ancestrales pertenecientes a un origen común de barriles TIM. Sin embargo, estudios



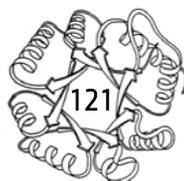
de reconstrucción ancestral pudieran ayudarnos a responder algunas de estas hipótesis sobre la evolución de las vías de plegamiento.

El desplegamiento de las BacTIMs: desplegamiento mediado por temperatura

Para la descripción de los resultados se agruparán las proteínas en dos grandes bloques: 1) aquellas cuyo desplegamiento térmico fue irreversible (IrrevTIMs) y 2) aquellas que mostraron un desplegamiento térmico reversible (RevTIMs).

BacTIMs con desplegamiento térmico irreversible: NpTIM y GoTIM

Los experimentos de desplegamiento térmico se siguieron por DC y DSC. En los experimentos de DC no se recupera la señal espectroscópica después de calentar y enfriar la muestra (inserto figura 34a-b); asimismo, la rampa de replegamiento no da ninguna señal que se parezca a la obtenida en el proceso de desplegamiento (figura 34a-b). Además, en DSC no se observó transición alguna después de un segundo ciclo de calentamiento (figura 34c-d). Al finalizar el experimento, en todos los casos se observó agregación de la proteína en la celda de DC. Para ambas proteínas no hubo recuperación de la actividad enzimática después de calentar la muestra. Todos estos resultados son indicativos de que después del primer ciclo de calentamiento, las enzimas no se repliegan al regresar a temperatura ambiente, esto es, el proceso de desnaturalización térmica es irreversible.



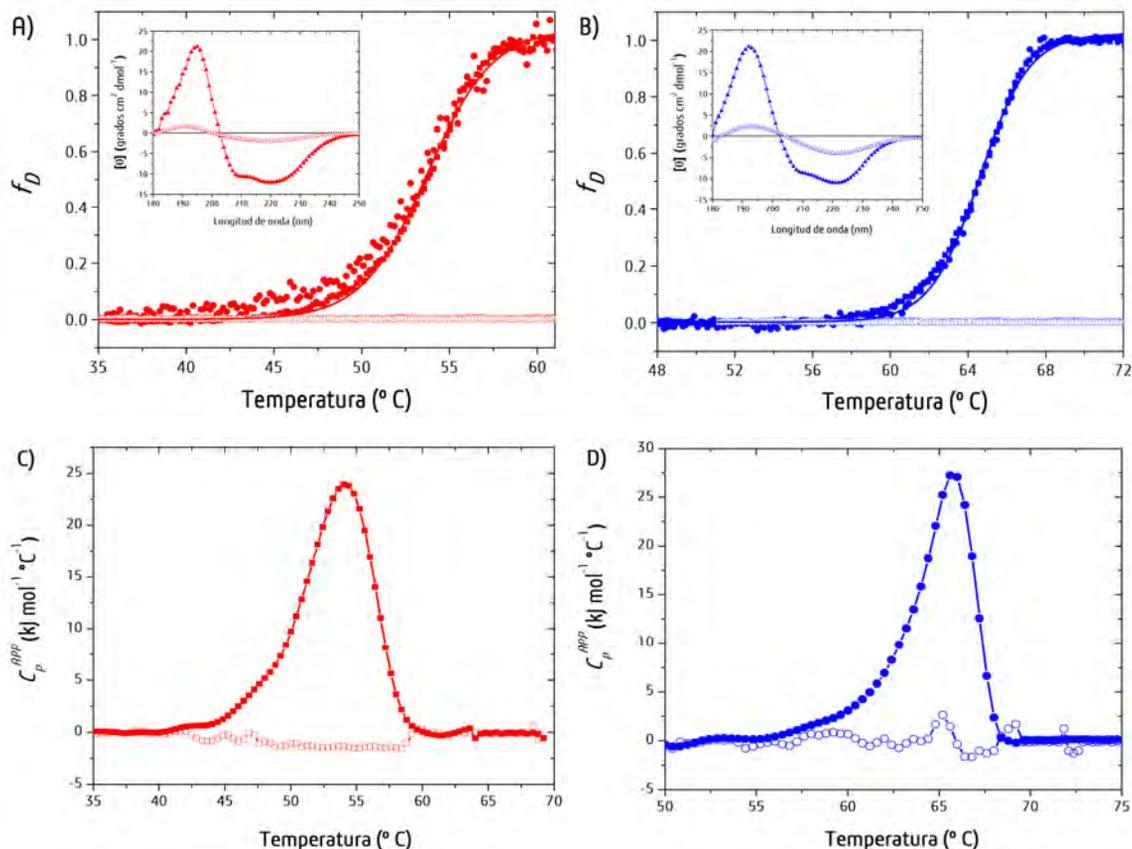
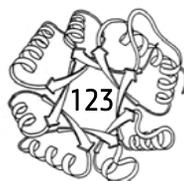


Figura 34. Despliegamiento térmico de las IrrevTIMs. Se muestran los experimentos seguidos por DC para **A)** *NpTIM* y **B)** *GøTIM*. Se muestran los datos para 195 nm (cuadro) y 222 nm (círculos) y del proceso de despliegamiento (símbolos cerrados) y repliegamiento (símbolos abiertos); las líneas muestran el ajuste de los datos a una ecuación de Boltzmann con la que se estima un valor de T_m . En los insertos se muestran los espectros de DC nativo a 25 °C (símbolos llenos) y los espectros de DC después de haber calentado y enfriado la muestra a 25 °C (símbolos vacíos). En los paneles inferiores se muestran los experimentos de DSC para **C)** *NpTIM* y **D)** *GøTIM*. Se indica la primera endoterma (símbolos llenos) y la segunda endoterma después de calentar y enfriar la enzima (símbolos vacíos). En todos los experimentos la concentración de proteína utilizada fue de 0.4 mg mL⁻¹ en amortiguador G y la velocidad de barrido fue 90 K min⁻¹.

Debido a la irreversibilidad en el despliegamiento, las trazas calorimétricas fueron bien ajustadas a un modelo de dos estados irreversible ($N \rightarrow F$, modelo de Lumry-Eyring) y se observó que la T_m se incrementó 5 K cuando la velocidad de barrido fue incrementada de 30 a 180 K h⁻¹, indicando que el proceso está bajo control cinético, un comportamiento comúnmente observado en las TIMs previamente estudiadas (figura 35a). Las siguientes evidencias indican que el despliegamiento inducido por temperatura de estas proteínas puede ser descrito por el modelo

irreversible de dos estados: 1) a todas las velocidades de barrido, los termogramas fueron bien ajustados, dando valores de E_A similares, 2) las gráficas de Arrhenius fueron lineales (figura 35b), y 3) la prueba de consistencia de datos fue exitosa, es decir, se cumplió la ecuación 13 (figura 35c). Estas evidencias también implican que el estado de transición cinéticamente relevante es dimérico. La E_A promedio obtenida de los tres métodos son 86 ± 6 kcal mol⁻¹ para $NpTIM$ y 171 ± 12 kcal mol⁻¹ para $GoTIM$ (tabla 16). Estos valores son similares a los valores más bajos y más altos reportados para la E_A de las TIMs de eucariontes (78-185 kcal mol⁻¹, tabla 16). Una consecuencia de que la E_A de $GoTIM$ sea del doble de $NpTIM$ es que la tasa de desnaturalización en condiciones fisiológicas para $GoTIM$ es muchos órdenes de magnitud más lenta que para $NpTIM$, lo que indica una estabilidad cinética mucho mayor para la primera. Ha sido propuesto que la estabilidad cinética es un parámetro clave para comprender el comportamiento de las proteínas, ya que juega un papel central en cómo la evolución ha logrado un balance entre función y estabilidad en escalas de tiempo biológicamente relevantes. Al respecto, también se ha planteado que la estabilidad cinética está íntimamente correlacionada con la flexibilidad térmica (Quezada *et al.*, 2017), lo que sugiere que podría estudiarse este aspecto en las IrrevTIMs para ver si esta correlación también aplica en dos enzimas procedentes de un mismo dominio pero con estabilidades cinéticas muy diferentes.



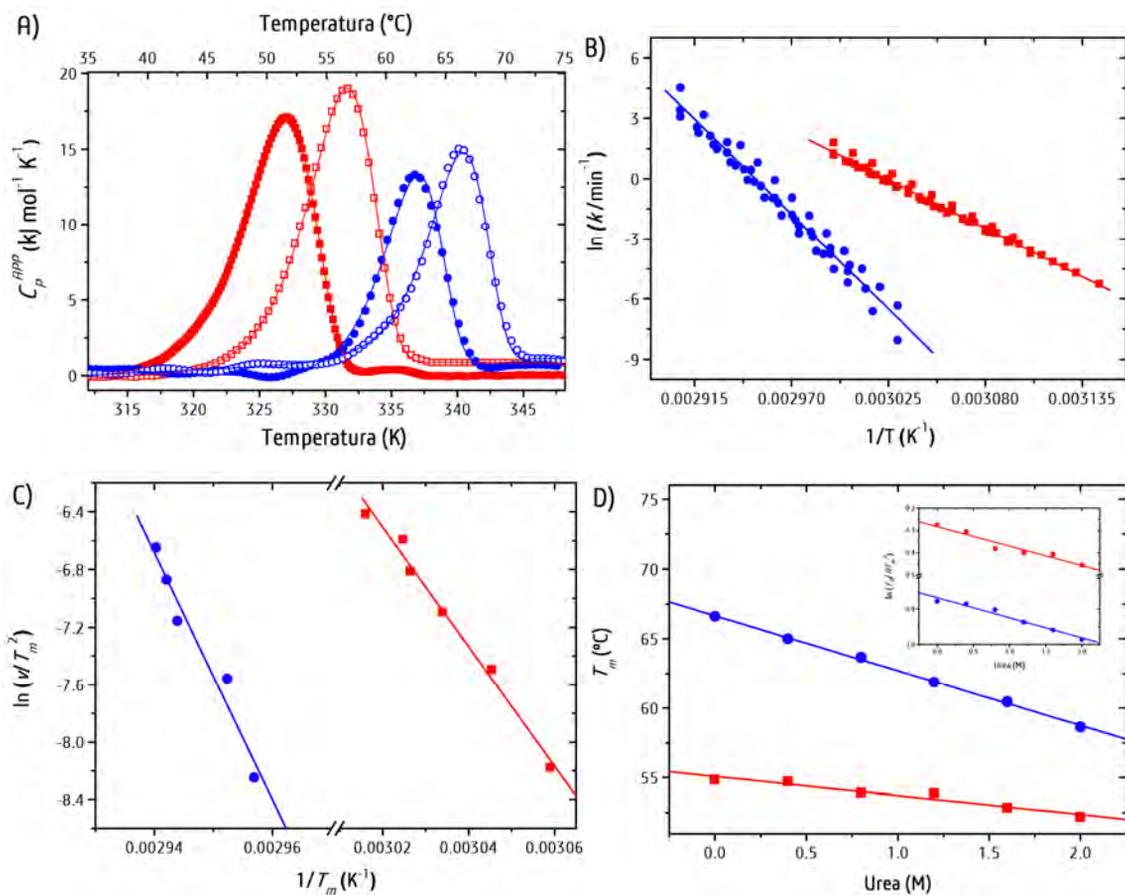


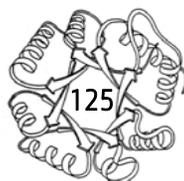
Figura 35. Análisis termodinámico para el desplegamiento térmico de las IrrevTIMs. **A)** Termogramas de DSC a 30 y 180 K h⁻¹ (símbolos cerrados y abiertos, respectivamente). Las líneas muestran el mejor ajuste a un modelo de dos estados irreversible. Otras velocidades de barrido (60, 90, 120 y 150 K h⁻¹) fueron también realizadas, pero no son mostrados por claridad. **B)** Gráfico de Arrhenius de los datos obtenidos a 6 velocidades de barrido distintas. Las líneas muestran el mejor ajuste a la ecuación de Arrhenius (ecuación 10). **C)** Efecto de la velocidad de barrido sobre la T_m para el desplegamiento de las IrrevTIMs. El inserto ilustra el cálculo del cambio de la energía de activación con respecto a la urea que está involucrado en la determinación de los valores de la m cinética. **D)** Efecto de la concentración de urea en la T_m . En todos los experimentos la concentración de proteína utilizada fue de 0.4 mg mL⁻¹ en amortiguador G para (--- ■ ---) NpTIM y (--- ● ---) GpTIM.

Tabla 16. Parámetros termodinámicos del desplegamiento térmico de las IrrevTIMs.

Parámetro termodinámico	IrrevTIMs		TIMs eucariontes				
	$N\rho$ TIM	Go TIM	G /TIM ^a	Hs TIM ^a	Lm TIM ^b	Tb TIM ^b	Tc TIM ^b
Energía de activación (kcal mol ⁻¹) del promedio de los ajustes individuales provenientes de las endotermas de DSC	88 ± 11	171 ± 15	NR	NR	75 ± 1	95 ± 1	193 ± 2
Energía de activación (kcal mol ⁻¹) del gráfico de Arrhenius	88 ± 4	171 ± 6	88 ± 2	86 ± 1	75 ± 1	95 ± 1	189 ± 18
Energía de activación (kcal mol ⁻¹) de la prueba de consistencia llevada a cabo a varias velocidades de barrido	82 ± 5	171 ± 21	NR	NR	83 ± 9	100 ± 13	171 ± 33
Energía de activación promedio (kcal mol ⁻¹)	86 ± 6	171 ± 11	NR	NR	78 ± 5	97 ± 7	185 ± 18
m^\ddagger (kcal mol ⁻¹ M ⁻¹)	0.399	1.942	0.942	0.662	0.351	0.404	1.988
m_{eq} (kcal mol ⁻¹ M ⁻¹)	6.681 ^c	6.558	7.174	7.192	7.361	7.253	7.331
m^\ddagger / m_{eq}	0.06	0.29	0.13	0.09	0.05	0.06	0.27
Número de residuos desplegados en el estado de transición ^d	30	148	67	45	24	28	136

Se ha observado una relación entre los valores m y los cambios en el ASA (Myers *et al.*, 1995). Así, los valores de m^\ddagger obtenido por los experimentos con urea pueden transformarse a grados fraccionales de exposición al solvente en el E_T mediante el cociente m^\ddagger / m_{eq} . El grado fraccionario de exposición de las proteínas al disolvente en el estado de transición (m^\ddagger / m_{eq}) se calculó a partir de experimentos de DSC utilizando diferentes concentraciones de urea (figura 35d). Si se compara este grado de exposición al solvente entre $N\rho$ TIM y Go TIM, se puede observar que el grado de exposición (0.06 y 0.29) son bajos, por lo cual el E_T tiene características semejantes al estado nativo en lo que refiere a la exposición al solvente, a pesar de esto, Go TIM muestra una mayor exposición al solvente que $N\rho$ TIM (tabla 16).

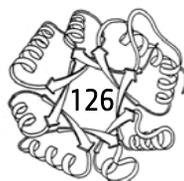
La E_A puede ser vista como la diferencia de la entalpía entre el estado de transición y el estado nativo ($E_A = \Delta H^\ddagger$). Los valores de m^\ddagger pueden tomarse como un estimado del grado de desplegamiento (o de ΔASA) en la conversión del estado nativo al estado de transición; el hecho de que se obtengan valores más altos de E_A con valores de m^\ddagger bajos, indica que el balance energético en el estado de transición es diferente que en los estados nativos. Si calculamos la entalpía en el E_T multiplicando el grado de exposición al solvente por la entalpía calculada del proceso de desplegamiento (tabla 16)



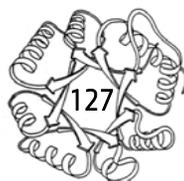
($NpTIM$: $0.06 \times 117 \text{ kcal}^{-1} \text{ mol}^{-1} = 7.02 \text{ kcal}^{-1} \text{ mol}^{-1}$; $GoTIM$: $0.29 \times 66 \text{ kcal}^{-1} \text{ mol}^{-1} = 16.82 \text{ kcal}^{-1} \text{ mol}^{-1}$) podemos ver que estos valores comparados con la E_R ($NpTIM$: $86 \pm 6 \text{ kcal}^{-1} \text{ mol}^{-1}$ y $GoTIM$: $171 \pm 11 \text{ kcal}^{-1} \text{ mol}^{-1}$) son mucho más pequeños. Para explicar esta diferencia tan grande, hay que considerar que claramente, el valor de la E_R experimental no puede ser explicado únicamente en términos del grado de desplegamiento en el E_T , ya que existen contribuciones energéticas adicionales (no presentes en el estado nativo) que necesitan involucrarse para la energética del estado de transición. Para otras TIMs reportadas, se ha visto que la contribución de la barrera de solvatación asociada con interacciones parcialmente desfavorecidas del agua en el estado de transición puede ser una causa de esta diferencia (Costas *et al.*, 2009). Se sabe que las barreras energéticas para una proteína en su proceso de plegamiento-desplegamiento son principalmente determinadas por el tamaño y la topología. A pesar de esto, las diferencias en la función biológica pueden requerir proteínas con el mismo tamaño y topología pero diferente barrera energética, tal como se observa con estas IrrevTIMs. Al parecer los efectos de solvatación pueden ser mecanismos eficientes para la modulación de estas barreras energéticas ya que evolutivamente pudieran estar controlando las estabildades cinéticas diferenciales en proteínas pertenecientes a una misma familia. Por lo anterior, se hipotetiza que esta variación en los efectos de solvatación pudiera estar sucediendo en $NpTIM$ y $GoTIM$; se propone hacer un análisis a profundidad con el fin de comprobar o descartar esta posibilidad.

BacTIMs con desplegamiento térmico reversible: DrTIM, CpTIM y ScTIM

Al igual que en las otras TIMs, el desplegamiento térmico se siguió por DC y DSC. Para los experimentos espectroscópicos se observa una recuperación mayor al 80% de la señal de DC (inserto figura 36a-c) y una coincidencia de las trazas de ida y regreso en las rampas de temperatura (figura 36a-c). De igual modo, se encontró una recuperación total de las endotermas de DSC después de un segundo ciclo de desnaturalización y ambos termogramas se ajustaron correctamente a un modelo



de dos estados con disociación ($N_2 \rightleftharpoons 2D$) (figura 37a-c). Tras calentar las muestras de proteína se recuperó la actividad enzimática en un 82% para *D*TIM, 81% para *Cp*TIM y 88% para *Sc*TIM. Además, al variar la concentración de proteína, se observa claramente una dependencia de la T_m con esta variación, tal cual se espera para una proteína dimérica cuyo desplegamiento está al equilibrio (inserto figura 37a-c). También, no se observa una dependencia de la T_m con respecto a la velocidad de barrido, contrario a lo observado para las IrrevTIMs bacterianas (figura 38). En conjunto, todos estos resultados son indicativos de que, para estas tres proteínas, el desplegamiento térmico es un proceso reversible y está al equilibrio, resultado nunca antes observado para una TIM. En las TIMs estudiadas hasta antes del presente trabajo, el desplegamiento térmico es un proceso irreversible, fuera del equilibrio o ambos (con presencia de histéresis), es decir, un proceso controlado cinéticamente, lo cual ha impedido una caracterización energético-estructural detallada. Debido a la histéresis y al alto grado de irreversibilidad del proceso, el uso de la temperatura como agente desnaturizante se había limitado a la estimación de la T_m de algunas TIMs y sus mutantes (Mainfroid *et al.*, 1996b; Schliebs *et al.*, 1996; Beaucamp *et al.*, 1997; Schliebs *et al.*, 1997; Alvarez *et al.*, 1998; Alvarez *et al.*, 1999; Gopal *et al.*, 1999; Sun y Sampson, 1999; Williams *et al.*, 1999; Benítez-Cardoza *et al.*, 2001; Jiménez *et al.*, 2003; Mathur *et al.*, 2006; Peimbert *et al.*, 2008a; Téllez *et al.*, 2008; Costas *et al.*, 2009; Lara-González *et al.*, 2014) y a la determinación de algunos parámetros termodinámicos dependiendo de la TIM reportada (Mainfroid *et al.*, 1996b; Benítez-Cardoza *et al.*, 2001; Téllez *et al.*, 2008; Costas *et al.*, 2009; Lara-González *et al.*, 2014).



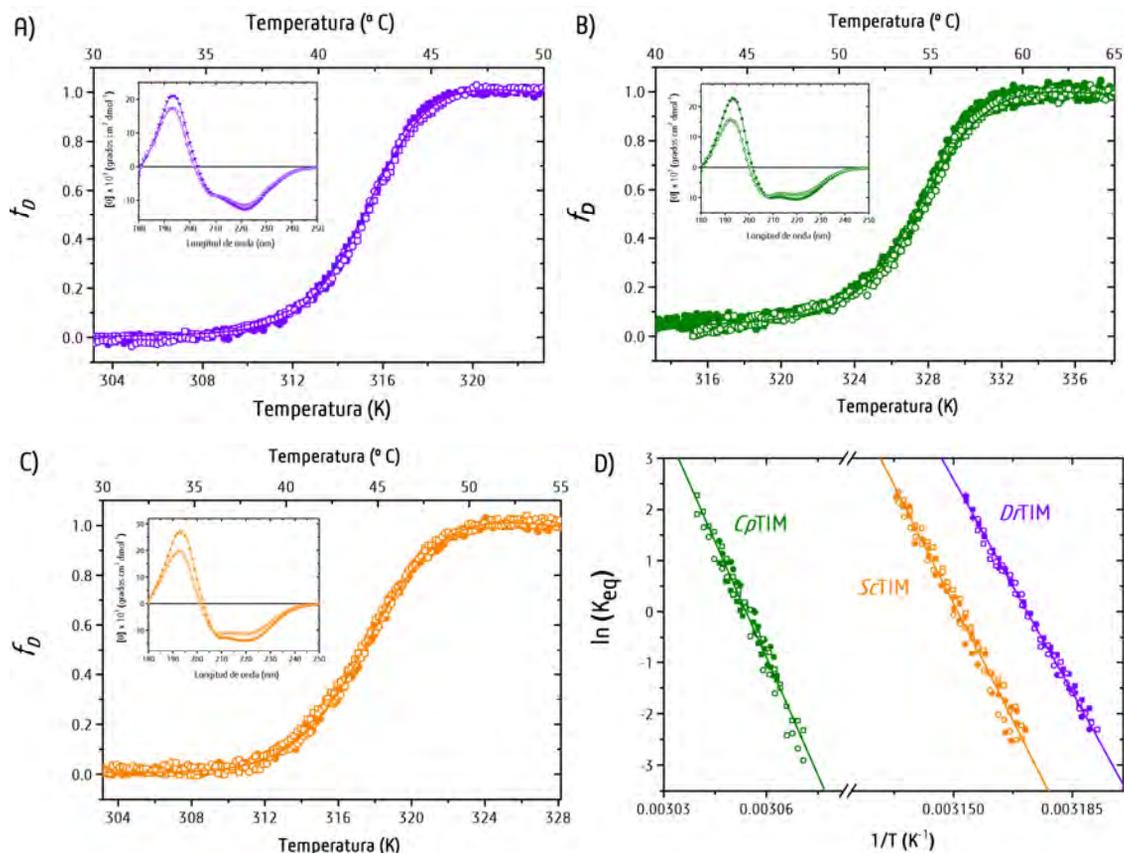


Figura 36. Despliegamiento térmico de las RevTIMs seguido por CD. Se siguieron los datos de despliegamiento/repliegamiento (símbolos cerrados y abiertos, respectivamente) para **A) DrTIM**, **B) CpTIM** y **C) ScTIM** a 220 nm (cuadrados) y 195 nm (círculos) a 90 K min⁻¹. Las líneas muestran el mejor ajuste a un modelo de disociación de dímeros de dos estados ($N_2 \rightleftharpoons 2D$). En los insertos se muestran los espectros de DC nativo a 25 °C (símbolos llenos) y los espectros de DC después de haber calentado y enfriado la muestra a 25 °C (símbolos vacíos). **D)** Gráficos de van't Hoff para los datos de despliegamiento y repliegamiento (símbolos cerrados y abiertos, respectivamente) seguidos a 220 nm (cuadrados) y 195 nm (círculos). Las líneas son los mejores ajustes a la ecuación de van't Hoff. Los valores de ΔH_{vH} derivados de estos análisis se indican en la tabla 16. En todos los paneles, la concentración de proteína fue de 0.4 mg mL⁻¹ y en amortiguador G.

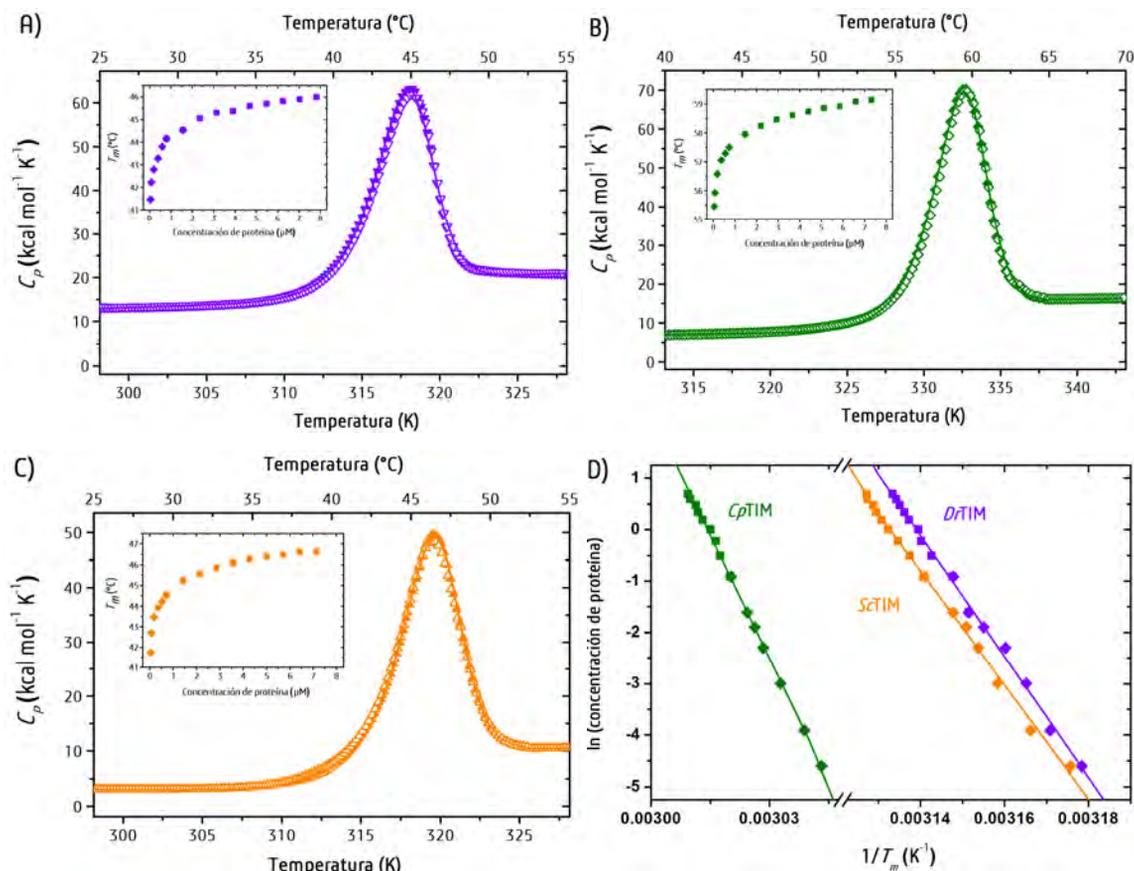


Figura 37. Despliegamiento térmico de las RevTIMs seguido por DSC. La primera y segunda endotermas (símbolos cerrados y abiertos, respectivamente) se muestran para **A)** *DpTIM*, **B)** *CpTIM* y **C)** *ScTIM*. Las curvas continuas muestran el mejor ajuste a un modelo de disociación de dímeros de dos estados ($N_2 \rightleftharpoons 2D$). En todos los paneles, la velocidad de barrido fue de 90 K min⁻¹ y la concentración de proteína fue de 0.4 mg mL⁻¹, a excepción de los insertos donde se muestran resultados variando la concentración de proteína. El inserto indica los valores de T_m obtenidos por CD (rombos) y DSC (cuadrados) en función de la concentración de TIM. **D)** Diagrama de Takahashi-Sturtevant para datos de CD y DSC (diamantes y cuadrados, respectivamente). Las líneas muestran los mejores ajustes de los datos a la ecuación 18.

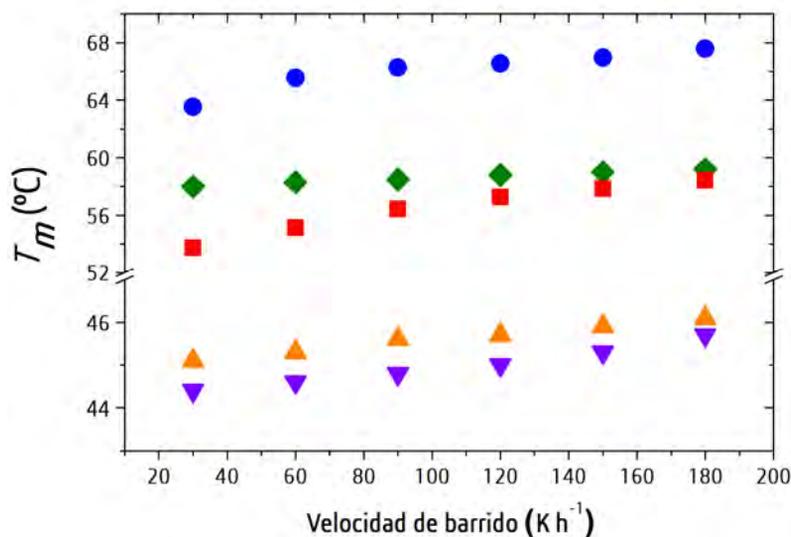
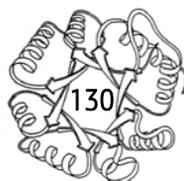


Figura 38. Dependencia de la T_m con respecto a la velocidad de barrido para las BacTIMs. En todos los casos la concentración de proteína utilizada fue de 0.4 mg mL^{-1} .

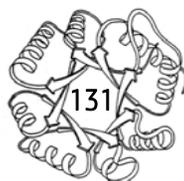
El valor obtenido de T_m para cada proteína (tablas 17 y 21) nos dejó con una amplia sorpresa, ya que el valor de $46 \text{ }^\circ\text{C}$ para $D\bar{T}IM$, es de los más bajos reportados hasta ahora para las TIMs, sólo por arriba de la TIM de *Moritella marina* ($41 \text{ }^\circ\text{C}$). Este comportamiento es destacable, debido a que *D. radiodurans* es un organismo termotolerante y nuestro valor de T_m está por debajo del intervalo de las TIMs provenientes de organismos termófilos ($\approx 80 \text{ }^\circ\text{C}$). Al respecto, se han reportado algunos mecanismos que utiliza *D. radiodurans* para resistir a la temperatura y radiación, entre los que destacan: presencia de grandes concentraciones de cosolutos como la trehalosa, iones como el manganeso, alteración de la membrana celular con incorporación de mayor número de proteínas, mayor presencia de chaperonas y proteínas para la reparación de ADN (Battista, 1997; Daly *et al.*, 2004; Daly, 2009). *ScoTIM* y *NpTIM* presentan valores dentro del intervalo de los reportados para la TIM de organismos mesófilos ($\approx 55 \text{ }^\circ\text{C}$). En cambio, para *GoTIM* y *CpTIM* presentaron los valores más altos de T_m , por arriba de las TIMs de organismos mesófilos y muy cercanas a las T_m de organismos termófilos. Esto sugiere que tal vez dichos organismos se enfrentan a altas temperaturas en su hábitat natural, por lo que deben de desarrollar una mayor estabilidad térmica en sus estructuras



proteicas (Somero, 1995). Entre las estrategias hasta ahora asociadas a la termoestabilidad de la enzima, se encuentra el estado de oligomerización, ya que se ha visto que, en la mayoría de las TIMs termófilas hasta ahora estudiadas, el estado de asociación es tetramérico y no dimérico como lo observado en TIMs mesófilas (a excepción de *GsTIM* que es un dímero (Álvarez *et al.*, 1999) y de *GTIM* que varía su estado de oligomerización entre dímero/tetrámero de acuerdo con su ciclo de vida (López-Velázquez *et al.*, 2004; Reyes-Vivas *et al.*, 2007)). Otra estrategia puede ser la reversibilidad del proceso de desnaturalización térmica que se observa en algunas TIMs bacterianas, aunque se desconoce si este mecanismo se encuentra diversificado en otros fila bacterianos o en especies del mismo filo.

El desplegamiento térmico de *D π TIM*, *ScTIM* y *C ρ TIM* puede describirse bien mediante un proceso de dos estados con disociación; tanto las curvas de desplegamiento seguidas por DC (figura 36a-c) y DSC (figura 37a-c) estaban bien ajustadas a este modelo. Además, el cambio de entalpía del estado nativo al estado desplegado (ΔH_{vH}) calculado a partir del análisis de van't Hoff de los datos de DC (figura 36d) y los obtenidos a partir de experimentos DSC fueron similares (tabla 17). También, el gráfico de Takahashi-Sturtevant construido con los datos tanto de DC como de DSC fue lineal (figura 37d), lo que nos dice que en ambas técnicas vemos la misma naturaleza del proceso de desplegamiento.

Finalmente, para evaluar si el desplegamiento térmico de estas proteínas es de dos estados, se determinó con el criterio calorimétrico, mediante el cociente entre el ΔH_{cal} (el obtenido por el ajuste de las trazas a un modelo de dos estados con disociación) y el ΔH_{vH} determinado por el área bajo la curva y la ecuación mostrada en la sección de estrategia metodológica. Por tanto, la comparación de ambos cambios de entalpía no sólo nos dice si el modelo es o no de dos estados, sino que además podemos determinar si existen diferentes unidades cooperativas de plegamiento. El cociente de ambas entalpías obtenidas a la temperatura media de desnaturalización para una proteína oligomérica de n subunidades, está dado por la ecuación 20 (Freire, 1989):



$$\frac{\Delta H_{vH}}{\Delta H_{cal}} = \frac{2n}{(n+1)} \quad (20)$$

Así, para proteínas diméricas que siguen un mecanismo de una transición simple involucrando sólo dos especies (N y D), $\Delta H_{cal} / \Delta H_{vH} \approx 1.33$. En cambio, si el mecanismo es diferente, dado que ΔH_{vH} está derivado de una suposición incorrecta, se observa que $\Delta H_{cal} \neq \Delta H_{vH}$. A esto es a lo que se le llama el "criterio calorimétrico" de evaluar si la transición es o no de dos estados. Para el caso de las RevTIMs, la relación de las entalpías van't Hoff y calorimétricas obtenidas (tabla 17) estuvo en excelente concordancia con la esperada para un proceso acoplado de disociación/despliegamiento.

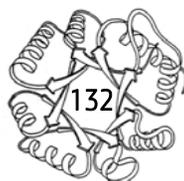
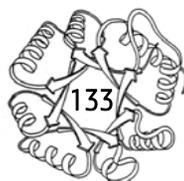


Tabla 17. Parámetros termodinámicos del desplegamiento térmico de las RevTIMs.

Parámetro fisicoquímico	D_r TIM	S_c TIM	C_p TIM
Temperatura media de desnaturalización (T_m) (°C)			
T_m^a	44.6 ± 0.1	45.3 ± 0.2	58.3 ± 0.1
Cambio de entalpía (ΔH) (kcal mol⁻¹)			
ΔH_{vH} por DC ^a	256 ± 7	282 ± 9	288 ± 11
ΔH_{vH} por DSC ^a	268 ± 10	300 ± 23	294 ± 16
ΔH_{cal} por DSC ^a	194 ± 8	225 ± 8	214 ± 12
Criterio calorimétrico ($\Delta H_{vH}/\Delta H_{cal}$)	1.38 ± 0.06	1.33 ± 0.09	1.37 ± 0.08
$\Delta H'$ del gráfico de Takahashi-Sturtevant	265 ± 11	275 ± 8	298 ± 15
Cambio en la capacidad calorífica (ΔC_p) (kcal mol⁻¹ K⁻¹)			
ΔC_p experimental ^a	6.6 ± 0.2	6.2 ± 0.8	6.8 ± 0.8
ΔC_p ajustado ^b	6.5 ± 0.2	6.6 ± 0.4	6.8 ± 0.4
ΔC_p promedio	6.6 ± 0.4	6.4 ± 0.6	6.8 ± 0.5
ΔC_p predicho basado en la secuencia aminoacídica			
ΔC_p Milardi <i>et al.</i> , 1997	7.7	8.5	8.4
ΔC_p Robertson y Murphy <i>et al.</i> , 1997	6.9	7.2	7.0
ΔC_p predicho basado en las relaciones estructurales de ΔASA			
ΔC_p Murphy y Freire <i>et al.</i> , 1992	13.8	13.8	14.2
ΔC_p Spolar y Record, 1994	10.5	10.5	10.8
ΔC_p Makhatadze y Privalov, 1995	17.0	17.0	17.5
ΔC_p Myers <i>et al.</i> , 1995	9.6	9.6	10.0
ΔC_p Robertson y Murphy <i>et al.</i> , 1997	7.9	7.9	8.4
ΔC_p Madan y Sharp, 2001	8.9	9.1	9.6

^a Los errores son la desviación estándar de los experimentos realizados a distintas concentraciones de proteína y velocidades de barrido.

^b ΔC_p determinado del ajuste a un modelo de dos estados con disociación tal cual se describe en el análisis de los datos.



Descripción fisicoquímica de la estabilidad de las RevTIMs: la curva de estabilidad

La reversibilidad y equilibrio en el proceso de desplegamiento permitió la estimación de los parámetros fisicoquímicos T_m , ΔC_p y ΔH de los experimentos de DSC (tabla 17). El ΔC_p fue obtenido de los termogramas de DSC a diferentes concentraciones de proteína, empleando un análisis posterior de dos formas distintas: 1) el C_p del estado nativo y desnaturalizado fue calculado de la diferencia de las capacidades caloríficas absolutas de dichos estados y 2) las endotermas fueron ajustadas a un modelo de dos estados con disociación y de dicho ajuste se obtiene un valor de ΔC_p experimental. Estos valores están en concordancia con los valores predichos por la secuencia aminoacídica; sin embargo, los valores son más bajos que los predichos por los cambios estructurales de ASA, posiblemente porque el estado desplegado no es un estado desnaturalizado en su totalidad; lo cual se ha observado en otras proteínas (Doyle *et al.*, 2013).

Usando la ecuación correspondiente y los parámetros termodinámicos determinados, las curvas de estabilidad para las RevTIMs fueron calculadas, es decir, la dependencia del ΔG como función de la temperatura (figura 39). Los valores de ΔG totales coinciden con los valores de ΔG calculados de la transición de los experimentos de DC (símbolos llenos de la figura 39). Los valores de ΔG_{dis} obtenidos de los experimentos de dilución llevados a cabo a diferentes temperaturas también están mostrados en la figura (símbolos vacíos de la figura 39). A partir de estas curvas, se obtuvieron los valores de la temperatura de estabilidad máxima y el valor de ésta para cada proteína además de su temperatura de desnaturalización en frío, valores que se muestran en la tabla 18.

Como se puede ver, las curvas de estabilidad de estas proteínas son distintas entre sí, lo que nos habla de la naturaleza propia de cada una de ellas. Se sabe que la forma de la curva de estabilidad puede ser modulada mediante cambios en la T_s (temperatura donde $\Delta S=0$), el ΔH_{T_s} (valor de entalpía en la T_s), y/o el ΔC_p . En este caso, la forma de la curva de estabilidad es principalmente debida a los cambios en la T_s y el ΔH_{T_s} , ya que el ΔC_p es casi idéntico entre las diferentes RevTIMs y contribuye de forma igual en las curvas de estabilidad respectivas. ΔH_{T_s} puede ser cambiado por mutaciones



puntuales y ha sido relacionado con un incremento en las interacciones estabilizantes en el estado nativo. T_s originalmente no fue considerado una fuente de estabilización, en embargo, recientemente ha sido reportado que cambios duales en ΔH_{TS} y T_s pueden incrementar la T_m (Hart *et al.*, 2014).

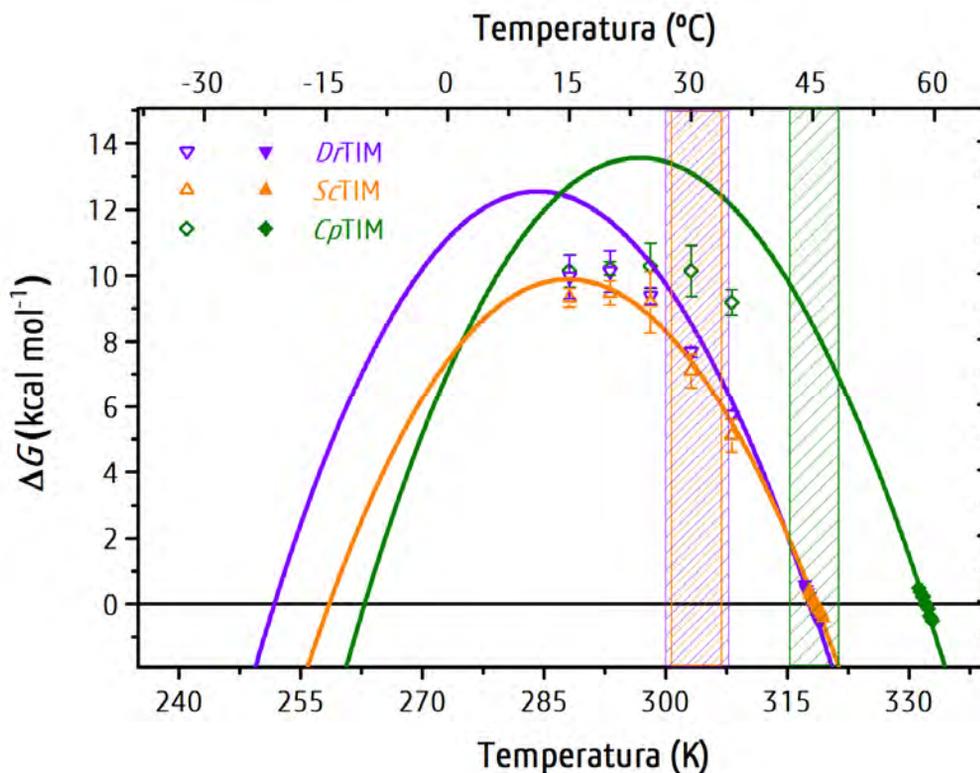


Figura 39. Curvas de estabilidad de barriles TIM naturales: *DpTIM*, *CpTIM* y *ScTIM*. Las líneas sólidas son las curvas de estabilidad construidas con los datos obtenidos de los experimentos de DSC y la ecuación de Gibbs-Helmholtz (ecuación 19). Los símbolos abiertos indican los valores experimentales de ΔG obtenidos de los experimentos de dilución y los símbolos llenos indican los valores experimentales de ΔG obtenidos de la transición de las rampas de desnaturalización de DC. La franja morada, naranja y verde indican los rangos de la temperatura óptima de crecimiento de *D. radiodurans*, *C. perfringens* y *S. coelicolor*, respectivamente.

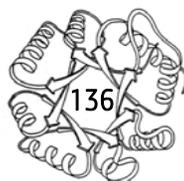
Tabla 18. Valores de estabilidad a partir de las curvas construidas para *DrTIM*, *CpTIM* y *ScoTIM*.

Parámetro	<i>DrTIM</i>	<i>CpTIM</i>	<i>ScoTIM</i>
Temperatura de máxima estabilidad (°C)	11.05	23.65	14.85
ΔG de máxima estabilidad (kcal mol ⁻¹)	12.55	13.57	9.89
ΔG a 25 °C (kcal mol ⁻¹)	10.39	13.55	8.78
Temperatura de desnaturalización en frío (°C)	-21.35	-10.25	-14.55

D. radiodurans, *S. coelicolor* y *C. perfringens* son mesófilos, sin embargo, cuando sus curvas de estabilidad son comparadas, existe una diferencia de 14 °C en la T_m y de 4 kcal mol⁻¹ en su estabilidad máxima. Lo anterior pudiera sugerir que, en caso de que la estabilidad sea un parámetro que está bajo presión de selección natural, esta presión ha sido diferente para las 3 proteínas, por lo que poseen valores de estabilidad y termoestabilidad completamente diferentes. Una comparación del intervalo de la temperatura óptima de crecimiento (TOC) (Murray, 1992; Bursy *et al.*, 2008; Li y McClane, 2006) con la curva de estabilidad indica que estas proteínas realizan su función a temperaturas más altas que la temperatura de máxima estabilidad (banda morada, naranja y verde en la figura 39). Es interesante notar que una TOC incrementada está acompañada de un incremento de la T_m , de tal forma que la estabilidad conformacional dentro de la TOC fluctúa en un rango muy reducido (7-9 kcal mol⁻¹).

En búsqueda de una posible relación entre las propiedades moleculares y la reversibilidad en el plegamiento

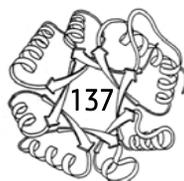
El origen molecular de la reversibilidad del plegamiento de las proteínas es desconocido. Por lo tanto, como un primero paso para avanzar en este conocimiento, es interesante establecer las relaciones razonables entre la reversibilidad y los distintos niveles moleculares y estructurales de las proteínas.



De esta forma, se exploraron las posibles relaciones en tres niveles distintos: 1) propiedades afines a la secuencia aminoacídica y las relaciones filogenéticas, 2) propiedades relacionadas con los estados estables (nativo, intermediarios, desnaturalizado) y 3) propiedades fisicoquímicas de la transición desplegamiento/replegamiento. Para fines de la comparación, se dividieron las TIMs en dos grandes grupos: aquellas que presentan un desplegamiento reversible (*D τ* TIM, *C ρ* TIM y *S τ* TIM) fueron llamadas RevTIMs y aquellas que su desplegamiento térmico es irreversible, que incluyó tanto a *N ρ* TIM y GoTIM como otras 17 TIMs que se habían estudiado previamente (tabla 6 y apéndice C).

Propiedades relacionadas con la secuencia y las relaciones filogenéticas.

Se compararon las secuencias aminoacídicas de los dos grupos de TIMs (RevTIMs vs. IrrevTIMs) y se encontró que el porcentaje de identidad varían entre 40 y 98%, sin embargo, no hay diferencias significativas entre los dos grupos, es decir, que exista un porcentaje de identidad mayor o menor entre las RevTIMs o de la misma forma pero con las IrrevTIM (tabla 19). Tampoco se observó relación con la abundancia particular de un grupo de aminoácidos en específico (tabla 20). De igual forma, no se encontró una región particular que pudiera diferenciar a las RevTIMs con respecto a las IrrevTIMs (figura 40). En cuanto a las propiedades globales de la secuencia, el grupo de las RevTIMs tiende a tener un punto isoeléctrico más bajo que las IrrevTIMs, grupos que se logran separar de forma estadísticamente significativa si se analizan con esta propiedad (tabla 21). Como posible relación con la reversibilidad, es sabido que la reversibilidad incrementa cuando el pH experimental está lejos del punto isoeléctrico de la proteína (Trivedi *et al.*, 1997; Shiraki *et al.*, 2002); esta característica se relacionará más adelante con otras dos propiedades numéricas de las secciones que se discutirán a continuación.





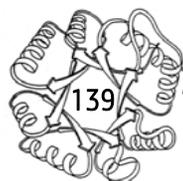
ND: el valor de RMSD no está determinado porque no existe estructura cristalográfica de *NpTIM*, *SsTIM* y *TsTIM*.

		Valores de RMSD																						
		IrrevTIMs															RevTIMs							
		<i>EcTIM</i>	<i>EHTIM</i>	<i>GgTIM</i>	<i>G/TIM</i>	<i>GsTIM</i>	<i>HsTIM</i>	<i>LmTIM</i>	<i>MmTIM</i>	<i>MtTIM</i>	<i>OcTIM</i>	<i>PfTIM</i>	<i>SsTIM</i>	<i>TbTIM</i>	<i>TcTIM</i>	<i>TsTIM</i>	<i>TvTIM</i>	<i>YtTIM</i>	<i>NpTIM</i>	<i>GoTIM</i>	<i>D/TIM</i>	<i>ScTIM</i>	<i>CpTIM</i>	
Porcentaje de identidad	IrrevTIMs	<i>EcTIM</i>	-----	1.51	1.30	1.53	1.48	1.39	1.13	1.24	1.10	1.50	1.27	ND	1.41	1.48	ND	1.24	1.08	ND	1.49	1.37	1.36	1.46
		<i>EHTIM</i>	41	-----	0.83	1.50	1.16	0.97	1.09	1.65	1.15	0.93	1.04	ND	0.82	1.07	ND	0.90	1.21	ND	1.51	1.20	0.87	1.07
		<i>GgTIM</i>	45	50	-----	1.01	0.92	0.40	0.65	1.20	1.05	0.34	0.85	ND	0.61	0.73	ND	0.62	0.68	ND	1.36	1.00	0.87	1.22
		<i>G/TIM</i>	40	43	44	-----	1.21	1.01	1.05	1.09	1.37	0.98	1.30	ND	1.10	0.88	ND	1.04	1.12	ND	0.86	1.49	1.45	1.52
		<i>GsTIM</i>	39	41	39	38	-----	0.99	1.49	1.51	1.20	0.94	1.30	ND	1.13	1.30	ND	1.11	1.24	ND	1.33	1.07	0.88	0.78
		<i>HsTIM</i>	44	49	90	45	38	-----	0.93	1.26	1.06	0.22	0.96	ND	0.69	0.79	ND	0.64	0.87	ND	1.13	1.29	0.91	1.17
		<i>LmTIM</i>	42	51	50	46	43	50	-----	1.16	1.13	0.91	0.84	ND	0.46	0.51	ND	1.03	0.92	ND	1.46	1.05	1.23	1.39
		<i>MmTIM</i>	65	41	40	35	41	40	38	-----	1.42	1.33	1.36	ND	1.40	1.17	ND	1.17	0.99	ND	1.31	1.39	1.44	1.26
		<i>MtTIM</i>	41	36	37	38	43	37	38	41	-----	1.11	0.93	ND	1.11	1.25	ND	1.15	1.10	ND	1.49	0.61	0.70	1.40
		<i>OcTIM</i>	44	49	89	45	38	98	50	40	37	-----	1.10	ND	0.76	0.87	ND	0.63	0.86	ND	1.14	1.13	0.92	1.11
		<i>PfTIM</i>	39	44	43	40	37	42	45	38	35	42	-----	ND	0.87	0.93	ND	1.07	1.01	ND	1.52	1.10	1.01	1.42
		<i>SsTIM</i>	43	48	87	44	37	93	49	39	35	93	42	-----	ND									
		<i>TbTIM</i>	42	43	51	44	40	52	69	38	38	51	43	39	-----	0.47	ND	0.89	0.98	ND	1.48	1.07	0.96	1.35
		<i>TcTIM</i>	42	45	50	46	41	52	68	39	39	51	43	50	74	-----	ND	0.91	0.88	ND	1.38	1.49	1.17	1.48
<i>TsTIM</i>	49	48	61	47	39	59	46	39	38	59	41	58	47	50	-----	ND								
<i>TvTIM</i>	40	44	48	41	43	47	43	40	39	47	38	46	42	42	48	-----	1.02	ND	1.15	1.01	0.89	0.95		
<i>YtTIM</i>	44	43	53	48	38	52	46	42	39	52	42	51	48	46	49	46	-----	ND	1.49	1.27	1.15	1.47		
<i>NpTIM</i>	39	38	37	33	47	36	36	43	40	36	33	35	34	32	34	38	35	-----	ND					
<i>GoTIM</i>	46	41	46	39	45	45	39	43	44	45	40	43	41	40	45	39	40	38	-----	1.25	1.43	1.44		
RevTIMs		<i>D/TIM</i>	38	34	38	36	41	38	37	39	48	37	33	37	39	38	39	37	40	40	-----	0.76	1.35	
		<i>ScTIM</i>	42	40	38	38	48	37	40	40	65	37	34	36	37	38	40	43	37	46	45	49	-----	0.94
		<i>CpTIM</i>	45	42	41	38	55	40	42	44	41	40	40	40	40	40	41	43	40	48	44	42	52	-----

Tabla 19. Matriz con el porcentaje de identidad y el promedio de los valores de RMSD para las IrrevTIMs y RevTIMs.

Tabla 20. Composición aminoacídica de las RevTIMs e IrrevTIMs.

Propiedad estructural	<i>Dr</i> TIM	<i>Sc</i> TIM	<i>Cp</i> TIM	Promedio RevTIMs	Promedio IrrevTIMs
Composición aminoacídica (%)					
Alanina (A)	17.0	13.6	12.4	14.3 ± 2.0	12.0 ± 3.1
Cisteína (C)	1.2	1.6	2.4	1.7 ± 0.5	1.5 ± 0.4
Ácido aspártico (D)	4.9	7.4	4.8	5.7 ± 1.2	4.4 ± 1.4
Ácido glutámico (E)	8.1	7.4	9.6	8.3 ± 0.9	7.2 ± 1.3
Fenilalanina (F)	0.4	1.6	2.8	1.6 ± 1.0	3.2 ± 0.9
Glicina (G)	10.9	8.9	7.2	9.0 ± 1.5	8.8 ± 1.4
Histidina (H)	2.0	3.1	2.0	2.4 ± 0.5	2.3 ± 0.6
Isoleucina (I)	3.2	5.0	8.0	5.4 ± 1.9	6.9 ± 1.2
Lisina (K)	3.6	7.0	7.6	6.1 ± 1.7	6.9 ± 1.8
Leucina (L)	10.1	8.5	4.8	7.8 ± 2.2	7.1 ± 1.3
Metionina (M)	1.6	1.9	3.2	2.2 ± 0.7	1.6 ± 0.7
Asparagina (N)	3.2	2.3	4.8	3.4 ± 1.0	4.0 ± 1.2
Prolina (P)	4.0	2.3	2.8	3.1 ± 0.7	3.5 ± 0.9
Glutamina (Q)	2.8	3.9	3.6	3.4 ± 0.4	4.0 ± 1.1
Arginina (R)	4.9	3.5	2.8	3.7 ± 0.9	3.6 ± 0.8
Serina (S)	4.9	3.9	2.4	3.7 ± 1.0	5.1 ± 1.1
Treonina (T)	4.5	4.3	7.2	5.3 ± 1.3	5.1 ± 0.9
Valina (V)	8.9	9.7	8.8	9.1 ± 0.4	9.2 ± 1.1
Triptófano (W)	1.2	0.8	0.8	0.9 ± 0.2	1.5 ± 0.5
Tirosina (Y)	2.4	3.5	2.4	2.8 ± 0.5	2.0 ± 0.6
Propiedades fisicoquímicas de los aminoácidos (%)					
Cargados (DEHKR)	23.5	28.3	26.7	26.2 ± 2.0	24.4 ± 1.9
Positivamente cargados (HKR)	10.5	13.6	12.4	12.1 ± 1.2	12.7 ± 1.3
Negativamente cargados (DE)	13.0	14.7	14.3	14.0 ± 0.8	11.6 ± 1.4
Alifáticos (AGILPV)	54.3	48.1	43.8	48.7 ± 4.3	47.5 ± 3.3
Aromáticos (FHWY)	6.1	8.9	8.0	7.7 ± 1.2	9.0 ± 0.8
Polares (DEKNQR)	27.5	31.4	33.1	30.7 ± 2.3	30.1 ± 2.4
Polares neutros (CNQSTY)	19.0	19.4	22.7	20.4 ± 1.7	21.8 ± 2.7
No polares neutros (AFGILMPVW)	57.5	52.3	50.6	53.5 ± 2.9	53.8 ± 2.9
Hidrofóbicos (CFILMVW)	26.7	29.1	30.7	28.8 ± 1.6	31.0 ± 1.4
Pequeños (ACGSTV)	47.4	41.9	40.2	43.2 ± 3.1	41.8 ± 3.0
Medianos (DEHILMN PQ)	40.1	41.9	43.4	41.8 ± 1.4	41.0 ± 2.9
Grandes (FKRWY)	12.6	16.3	16.3	15.1 ± 1.8	17.2 ± 2.1



.....EcTIM..... 1 . . . MRHPLVMGNWKLN . GSRHMVHELVSNLRK . . . ELAGV . . AGCAVA IAPPEMY IDMAKREA . . . EGS HIM 60
EhTIM..... 1 . MSARKFVVGGNWKC . GTLAS IETLTKGVAASVDAELA . . . KKVEV IVGVFF IY I PKVQQL I LAGEANGAN I L 68
GgTIM..... 1 MAP . RKFFVVGGNWKM . GDKKS LGEL IHTLNG . . . AKLS . . . ADTEVVCGAPS IYLDFAHQKL . . . DAK IG 60
GtTIM..... 1 . MPATRRPF I GGNFKCN . GS LDF I KSHVAA IAA . . . HK IP . . . DSDVYV IAPS AVHLSTAI AN . . . TSKQLR 61
GsTIM..... 1 . . . MRKP I IAGNWKMH . KTLAEAVQFVEDVKG . . . HYPPA . . . DEV I SVVCAPFLFLDRLVQAA . . . DGTDLK 60
HsTIM..... 1 MAPSRKFVVGGNWKM . GRKQSLGEL IGT LNA . . . AKVP . . . ADTEVVCAPPTAY IDFAHQKL . . . DPK IA 61
LmTIM..... 1 MSAKPQP I AAANWKC . GTTAS I EKL VQVFNE . . . HT IS . . . HDVQC VVAPT FVH I PLVQAKL . . . RNPKYV 62
MmTIM..... 1 . . . MRHPVVMGNWKLN . GSKEMVVD LNLGNA . . . ELEGV . . . TGDVAVAPPALFVDLAERTLT . . . EAGSA I I 62
MtTIM..... 1 . MSRKPL IAGNWKMN . LNHYEALVQK IAF . . . SLPDKYYDRVDVAV I PPFTDLRSVQTLV . . . DGDKLR 63
OcTIM..... 1 MAPSRKFVVGGNWKM . GRKKNLGEL ITT LNA . . . AKVP . . . ADTEVVCAPPTAY IDFAHQKL . . . DPK IA 61
PftTIM..... 1 . . . MARKYFVAANWKC . GTLES I KSL TNSFNN . . . LDFDP . . . SKLDV VVFPVSVHYDHTRKL . . . QSKFS 60
SsTIM..... 1 MAPARKFVVGGNWKM . GRKNNLGEL INTLNA . . . AKLP . . . ADTEVVCAPPTAY IDFAHQKL . . . DPK IA 61
TbTIM..... 1 . MSKPQP I AAANWKC . GSQQLSEL IDLFNS . . . TS IN . . . HDVQC VVAST FVHLMTKERL . . . SHPKFV 61
TcTIM..... 1 MASKPQP I AAANWKC . GSESLLVPL IETLNA . . . ATFD . . . HDVQC VVAPT FLH I PMPKARL . . . TNPKFQ 62
TtTIM..... 1 . MTRKLFVVGGNWKM . GSYSHINTFFDTLQK . . . ADTD . . . PNAD I V IGVFACYLKYAQDKA . . . PKG IK 59
TvTIM..... 1 . . . MRTFFVVGGNWKNPKTVEEA EKL I EMLNG . . . AKVE . . . GNVEVVVAAPF I FLPTLQQL . . . RKDWK 59
YtTIM..... 1 . . . MARTFFVVGGNFKLN . GSKQS I KE IVERLNT . . . AS IP . . . ENVEVV ICPPATYLDYSVSLV . . . KKPQVT 60
NpTIM..... 1 . . . MRK I V IAGNWKMF . KTQAE TQEF LQGFLP . . . HLEETP . QGREV I LCP PFTDL SVLSKTL . . . HGSL IQ 61
GoTIM..... 1 . MPT RKKFVAGNWKMN . TTLAEAKA LGA AVAK . . . GVTD . . . DRVT VAVFP PYPWLTAVGEVL . . . KGS PVA 61
DrTIM..... 1 . . . MQTLLALNWKMN . KPTTEARSWAEELTT . . . KYAPA . . . EGVDLAVLAPALDLSALAANL . . . PAG IA 58
ScTIM..... 1 . MTRTRPLMAGNWKMN . LNHLEA IAHVQK LAF . . . ALADKDYDAVEAVLAPFTDLRSVQTLV . . . DGDKLK 64
CpTIM..... 1 . . . MRTP I IAGNWKMH . YTIDEAVK LVEELKP . . . LVKD . . . AKCEVVVCPTFVCLDAVKKAV . . . EGTN IK 59

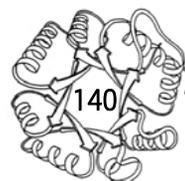
Consensus

R . . . G N W K N G . L . . . V V V . P P . . . L . . .
 M+PMRKFVVGGNWKMN-GTKAEA+ELVETLNA--AKI++-YADVEVVVAPPFVYLDFAQQKL+---DGPKI+

.....EcTIM..... 61 L . . GAQNVDLNLGSAFTGET SAAMLKD IGAQY I I I GHSERRTYHKE SDEL IAKKFAVLKEOGLTPVLCI GETE 131
EhTIM..... 69 V . SAENAW . TKSGAYTGEVHVGM LVDCCQVPY I L GHSERRQI FHESNEQVAEKVQA I D AGLKV I AC I GEL 138
GgTIM..... 61 V . AAQNCYKVPKGAFTGE I SPAMIKD IGAAW I L GHSERRHVFGE SDEL IGQKV AHALAEGLGV I AC I GEL 131
GtTIM..... 62 I . AAQN VYLEGN GAWTGET SVEMLQDMGLKHV I V GHSERRR I MGETDEQS AKKAKRALEKGMTV I FCVGETL 132
GsTIM..... 61 I . GAQTMHFADQGA YTGEVSPVMLKD LGVTYV I L GHSERRQMF AETDET VNKVLA AFRGL I P I I CCGESL 131
HsTIM..... 62 V . AAQNCYKVTNGAFTGE I SPGM I KDCGATWVVL GHSERRHVFGE SDEL IGQKV AHALAEGLGV I AC I GEL 132
LmTIM..... 63 I . SAENA I . AKSGAFTGEV SMP I LKD IGVHW I L GHSERRTYG ETD E I VAQKVSEACKQGF MV I AC I GEL 132
MmTIM..... 63 L . GAQNTDLNNSGAF I GDMSPAM LKEFGATH I I I GHSERREYHAE SDEFVAKKFAFLKENG LTPVLC I G ESD 133
MtTIM..... 64 LTYGAQDLSPHDSGAYTGDVSGAFLAKLGCSYVVV GHSERRTYHNE DDA LVA AKAAT ALKHGLT P I VC I GEHL 136
OcTIM..... 62 V . AAQNCYKVTNGAFTGE I SPGM I KDCGATWVVL GHSERRHVFGE SDEL IGQKV AHALAEGLGV I AC I GEL 132
PftTIM..... 61 T . GIQNVSKFGN GSYTGEVSAE IAKDLN I EYV I I GHSERRYFHETDEDVREK LQAS LKNNLKAVVCFGESL 131
SsTIM..... 62 V . AAQNCYKVANGAFTGE I GPGMIKDLGATWVVL GHSERRHVFGE SDEL IGQKV AHALAEGLGV I AC I GEL 132
TbTIM..... 62 I . AAQNA I . AKSGAFTGEVSLP I LKD IGVNWI V L GHSERRAYYGETNE I VADKVA AAVASGFMV I AC I GEL 131
TcTIM..... 63 I . AAQNA I . TRSGAFTGEVSLQ I LKDYGIKWV L GHSERRLYYGETNE I VAEKVAQACAAGFHV I VCVGETN 132
TtTIM..... 60 I . AAENCYKVGSGAFTGE ISTEMIKDCGCEW I L GHSERRH I FGE SNE L I GEKVHALDSGLNV I PC I GEL 130
TvTIM..... 60 V . SAENVFTKPNGAFTGEVTPMIKSGF I EWT I L GHSERRD I LKEDDEF LAAKAKFALENGMK I I YCCGEHL 130
YtTIM..... 61 V . GAQNAYLKASGAF TGENSV DQ I KDVGAKWV I L GHSERRSYFHEDDKF I ADKTKFALGOGVGV I L C I GEL 131
NpTIM..... 62 L . GAQN IHWEEFGAYTGE I SGPMLTESGVRFV I V GHSERRQYFGETDATVNLRLRTAQRFLG LTP I LCVGETK 132
GoTIM..... 62 L . GAQDVSSEKKGAF TGEVSPAM LLETGCKYAL I GHSERRH I I G ESETF I NHKVHTALEEGLSVVLCMGETL 132
DrTIM..... 59 F . GGQDVS SAHESGAYTGE I SAAMLK D GASCVVV GHSERRREYHDESDATVAAKARQAQANGLLP I VCVGENL 129
ScTIM..... 65 I KYGAQD I SAHDGGAYTGE I GPM LAKLKCTYVAV GHSERRQYHAE TDE I VNAKVA AAYKHGLTP I LCVGEEL 137
CpTIM..... 60 V . GAQNMHFEKGAFTGE I APRMLEAMN IDYV I I GHSERRREYFN ETD ETCNKVKVAAFAHNLTP I LCCGETL 130

Consensus

AQN . GA-TGE S MLKD G -V- GHSERR -E-DE -K- A -GL- C-GE L
 V-YGAQN VYKVGSAFTGE+SPAM LKDLGATWV I L GHSERRHYFGE SDELVA+KVA+ALAEGLTV I +C I GEL



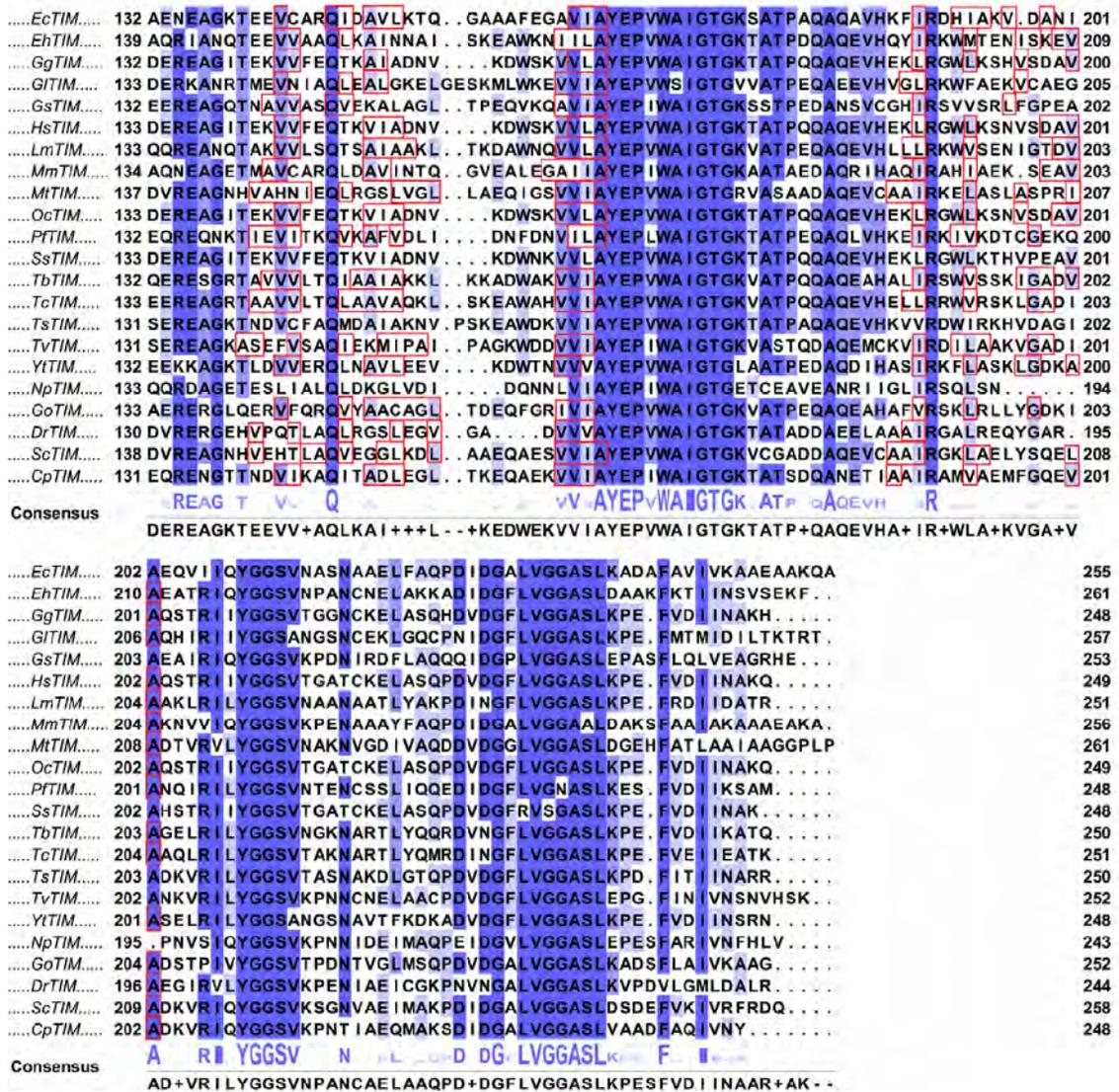


Figura 40. Alineamiento de las secuencias de las TIM usadas para la comparación. Los residuos conservados se muestran en la línea consenso y representados de azul claro a oscuro de acuerdo al incremento de conservación. En cuadros rojos se muestran los residuos (tanto internos como vecinos) involucrados en la formación de la cavidad más grande de las estructuras.



Tabla 21. Parámetros fisicoquímicos de las RevTIMs e IrrevTIMs.

	TIM	pI ^a	Volumen de cavidades(Å ³) ^b	T_m (°C) ^c	Referencia ^d
IrrevTIMs	<i>Eh</i> TIM ^e	6.1	3777	40.5 / 56.3	Téllez <i>et al.</i> , 2008
	<i>Mm</i> TIM	4.8	2927	41.0	Alvarez <i>et al.</i> , 1998
	<i>Ec</i> TIM	5.6	4667	54.0	Alvarez <i>et al.</i> , 1998
	<i>Mt</i> TIM	5.9	3087	55.0	Mathur <i>et al.</i> , 2006
	<i>Np</i> TIM	5.0	ND	55.2	Este trabajo
	<i>Tb</i> TIM	9.1	3780	55.8	Costas <i>et al.</i> , 2009
	<i>Oc</i> TIM	7.1	3485	56.0	Sun <i>et al.</i> , 1995
	<i>G</i> /TIM	7.1	3920	57.5	Enríquez-Flores <i>et al.</i> , 2007
	<i>yt</i> TIM	5.8	3699	59.1	Benitez-Cardoza <i>et al.</i> , 2001
	<i>Tc</i> TIM	8.6	3414	59.2	Costas <i>et al.</i> , 2009
	<i>Ts</i> TIM	6.6	ND	62.3	Jiménez <i>et al.</i> , 2003
	<i>Lm</i> TIM	8.2	3531	62.5	Costas <i>et al.</i> , 2009
	<i>Hs</i> TIM	6.5	3259	63.1	Aguirre <i>et al.</i> , 2014
	<i>Ss</i> TIM	7.1	ND	63.9	Jiménez <i>et al.</i> , 2003
	<i>Pf</i> TIM	6.1	4452	65.0	Ray <i>et al.</i> , 1999
	<i>Go</i> TIM	6.3	4010	65.6	Este trabajo
	<i>Tv</i> TIM ^f	5.6	3383	67.1	Lara-González <i>et al.</i> , 2014
	<i>Gg</i> TIM	6.7	3518	67.2	Shi <i>et al.</i> , 2008
	<i>Gs</i> TIM	5.2	2503	76.0	Alvarez <i>et al.</i> , 1999
	Promedio (μ)	6.5	3651	59.3	
Desviación estándar (σ)	1.2	514	8.5		
$\mu + \sigma$	7.7	4165	67.8		
$\mu - \sigma$	5.3	3137	50.7		
RevTIMs	<i>Dr</i> TIM	4.9	2134	44.6	Este trabajo
	<i>Sr</i> TIM	5.2	2921	45.3	Este trabajo
	<i>Cp</i> TIM	5.0	2241	58.3	Este trabajo

^a Determinado con el software Protparam (Gasteiger *et al.*, 2005).

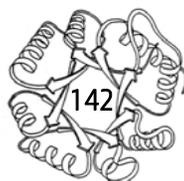
^b El volumen de las cavidades fue calculado sin considerar las moléculas de agua.

^c Los experimentos reportados están en un rango de concentraciones entre 0.2 y 0.5 mg mL⁻¹ y a una velocidad de barrido de 60 K h⁻¹, a excepción de *Tv*TIM donde es 120 K h⁻¹.

^d Los valores de la T_m fueron obtenidos de estas referencias.

^e Los valores corresponden a la disociación del dímero y desplegamiento del monómero.

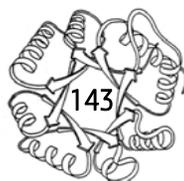
^f Las abreviaturas de las enzimas se encuentran en un apéndice al inicio del texto.



Sobre las relaciones filogenéticas, *S. coelicolor* y *C. perfringens* pertenecen a los fila Actinobacteria y Firmicutes, respectivamente. Ambos fila son del grupo Posibacteria, mientras que *D. radiodurans* pertenece al filo Deinococcus-Thermus, parte del grupo Eobacteria, un grupo distante. Debido a esto, la reversibilidad parece no estar restringida a un grupo filogenético en particular.

Propiedades relacionadas con los estados estables.

Se analizaron las propiedades de los estados estables de ambos grupos de proteínas. Primero, se observó que las IrrevTIMs muestran espectros de DC en el estado desnaturalizado que están asociados a agregación irreversible (figura 41). En contraste, el espectro para las RevTIMs no muestra evidente estructura secundaria residual en el estado desnaturalizado, lo cual contrasta con lo observado para la TIM de levadura, donde sí se observa claramente un espectro con estructura residual y los autores reconocen el hecho como una posible causa de la histéresis que sufre la proteína en su proceso de replegamiento (Benítez-Cardoza *et al.*, 2002). Debido a esto, es posible que la ausencia o la poca estructura residual en el estado desnaturalizado puedan estar relacionado con la reversibilidad.



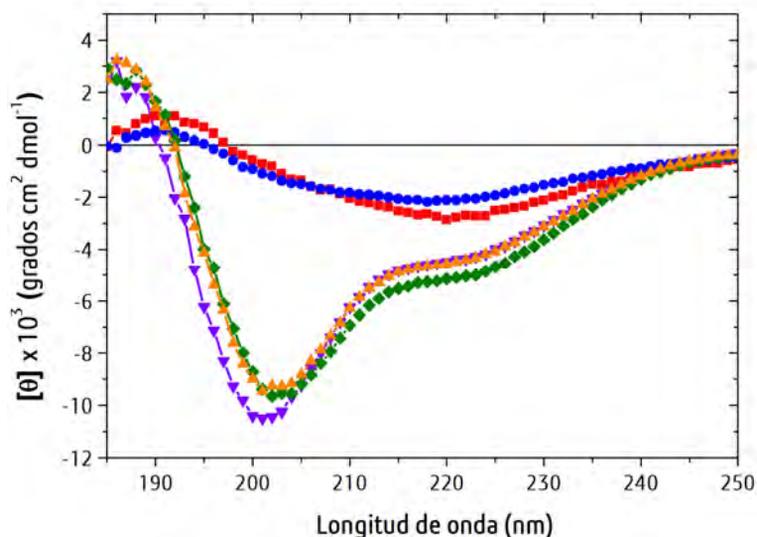


Figura 41. Espectro de Dicroísmo Circular desnaturalizado. En todos los casos la temperatura fue de 90 °C y en amortiguador G. La concentración de proteína fue de 0.4 mg mL⁻¹. *DrTIM* (--- ▽ ---), *NpTIM* (--- ■ ---), *GøTIM* (--- ● ---), *CpTIM* (--- ◆ ---), *ScoTIM* (--- ▲ ---).

Cuando se compararon las propiedades estructurales de los dos grupos, no se encontraron diferencias significativas en la cadena principal del estado nativo de las proteínas comparadas, ya que el RMSD es menor a 1.5 Å (tabla 19). Respecto a las características estructurales de los monómeros, no se encontraron diferencias cuando las interacciones intramonómeros (puentes de hidrógeno y sal) fueron comparadas; tampoco se observa diferencia entre los elementos de estructura secundaria (tabla 22). Lo mismo ocurre cuando se compara la intercara del dímero y también en el área accesible al solvente que se esconde en la dimerización (tabla 23). En resumen, en todas las propiedades estructurales involucradas con interacciones estabilizantes no se encuentra una relación clara con la reversibilidad.

Tabla 22. Elementos de estructura secundaria e interacciones estabilizantes de las RevTIMs e IrrevTIMs.

Propiedad estructural	<i>D</i> rTIM	<i>S</i> cTIM	<i>C</i> pTIM	Promedio RevTIMs	Promedio IrrevTIMs
Elementos de estructura secundaria (%)					
α -hélices	44	48	43	45.0 \pm 2.2	44.9 \pm 2.1
Hebra- β	21	20	21	20.7 \pm 0.5	21.8 \pm 1.0
Estructura no repetitiva	35	32	36	34.3 \pm 1.7	33.3 \pm 1.8
Interacciones estabilizantes					
Total de puentes de hidrógeno en el oligómero	490	503	510	501 \pm 8	521 \pm 24
Puentes de hidrógeno en la intercara	27	19	24	23 \pm 3	28 \pm 5
Total de puentes salinos en el oligómero	37	50	44	44 \pm 5	34 \pm 7
Puentes salinos en la intercara	2	2	12	5 \pm 5	6 \pm 3

Tabla 23. Propiedades estructurales utilizadas para la comparación de las RevTIMs e IrrevTIMs.

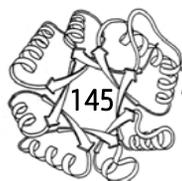
Propiedad estructural	<i>D</i> rTIM	<i>S</i> cTIM	<i>C</i> pTIM	Promedio RevTIMs	Promedio IrrevTIMs
ASA total del oligómero plegado (\AA^2)	18107	19475	19947	19176 \pm 780	19040 \pm 629
ASA hidrofóbica del oligómero plegado (\AA^2)	10389	11367	11481	11079 \pm 490	11498 \pm 600
% Hidrofobicidad ^a	57.4	58.4	57.6	57.8 \pm 0.4	60.4 \pm 1.3
Δ ASA total escondida en el plegamiento del monómero (\AA^2)	25222	26550	26679	26151 \pm 658	27056 \pm 740
Δ ASA hidrofóbica escondida en el plegamiento del monómero (\AA^2)	18367	19465	19269	19034 \pm 478	19552 \pm 375
% Hidrofobicidad ^b	72.8	73.3	72.2	72.8 \pm 0.5	72.3 \pm 1.9
Δ ASA total escondida del monómero plegado en la asociación (\AA^2)	3113	3425	3467	3335 \pm 158	3275 \pm 207
Δ ASA hidrofóbica escondida del monómero plegado en la asociación (\AA^2)	2131	2286	2339	2252 \pm 89	2005 \pm 354
% Hidrofobicidad ^c	68.5	66.7	67.5	67.6 \pm 0.7	61.2 \pm 2.8
Δ ASA total del monómero desplegado en la asociación (\AA^2)	53558	53524	56826	55636 \pm 1474	57388 \pm 1365
Δ ASA hidrofóbica del monómero desplegado en la asociación (\AA^2)	38864	41217	40878	40320 \pm 1038	41110 \pm 835
% Hidrofobicidad ^d	72.6	77.0	71.9	73.8 \pm 2.3	71.6 \pm 2.1
Δ ASA _{dissoc} / Δ ASA _{tot} (%)	5.8	6.4	6.1	6.1 \pm 0.3	5.7 \pm 0.4
Δ ASA _{dissoc} / Δ ASA _{tot} (%)	5.5	5.5	5.7	5.6 \pm 0.1	4.8 \pm 0.7
Volumen molecular del oligómero (\AA^3)	62052	67188	69885	66375 \pm 3249	67339 \pm 2221
Número de cavidades en el oligómero (sin solvente)	10	11	9	10 \pm 1	16 \pm 3
Volumen total de las cavidades en el oligómero (\AA^3) (sin solvente)	2134	2921	2241	2432 \pm 349	3651 \pm 498
Volumen total de la cavidad más grande en el oligómero (\AA^3)	324	683	290	432 \pm 178	1001 \pm 209
Volumen total de las cavidades en el oligómero (\AA^3) (con solvente)	928	1335	1008	1090 \pm 176	1908 \pm 366

^a % hidrofobicidad = (ASA hidrofóbica en el oligómero / ASA total del oligómero) * 100%.

^b % hidrofobicidad = (Δ ASA hidrofóbica oculta en el plegamiento del monómero / Δ ASA total oculta en el plegamiento del monómero) * 100%.

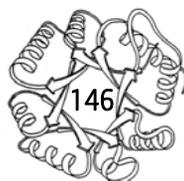
^c % hidrofobicidad = (Δ ASA hidrofóbica en el monómero plegado oculta en la dimerización / Δ ASA total en el monómero plegado oculta en la dimerización) * 100%.

^d % hidrofobicidad = (Δ ASA hidrofóbica en el monómero desplegado oculta en la dimerización / Δ ASA total en el monómero desplegado oculta en la dimerización) * 100%.



La única diferencia detectada hasta el momento entre los estados nativos de las RevTIMs y las IrrevTIMs está relacionada con las cavidades internas. Las RevTIMs muestran un menor número de cavidades, su volumen es aproximadamente 40% menor que el encontrado en las IrrevTIMs (tabla 23 y figura 42). El tamaño de la cavidad mayor (figura 43), localizada en la intercara entre las hebras β 4-6 y las hélices α 4-6 (residuos 90-120, 145-160 y 195-200), es la mayor responsable de la diferencia en el volumen entre los dos grupos de TIMs. Hasta el momento no se ha reportado previamente alguna relación entre el volumen de las cavidades y la reversibilidad en el desplegamiento térmico, por lo que el tema se deberá explorar a mayor profundidad. Además, resulta interesante investigar si el tema de las cavidades también correlaciona con la reversibilidad en otras proteínas con topología distinta a los barriles TIM e, incluso, para familias de proteínas con distinto origen evolutivo, ya que hasta ahora no tenemos conocimiento de que se hayan correlacionado estas dos propiedades propuestas en el presente trabajo.

Asimismo, en relación con el bajo punto isoeléctrico y el bajo volumen de cavidades que presentan las RevTIMs, un análisis estructural de todas las cavidades reveló que no existe una relación específica entre estos dos parámetros; es decir, con respecto a que las cavidades de las RevTIMs contengan residuos más ácidos que las IrrevTIMs, o que la composición aminoacídica de las cavidades entre ambos grupos sea distinta de acuerdo a las propiedades fisicoquímicas de los residuos. Al parecer, la diferencia en los puntos isoeléctricos otorgada por los aminoácidos ácidos se encuentra distribuida en residuos de diferentes partes de la proteína y no específicamente involucrados en la formación cavidades. Se propone hacer un análisis más exhaustivo para ver si existen algunas diferencias claras entre las RevTIMs e IrrevTIMs en relación con las cavidades y el punto isoeléctrico; por ejemplo, estudiar el papel de ambas propiedades mediante mutagénesis sitio-dirigida en cavidades importantes y/o regiones específicas de la proteína y con residuos de distinta naturaleza fisicoquímica (ácidos, básicos, neutros polares, neutros apolares, etcétera).



Por último, se ha sido sugerido que el volumen de las cavidades internas tiene influencia en la estabilidad y la cinética de desplegamiento de las proteínas (Silva y Weber, 1993; Roche *et al.*, 2013); sin embargo, hasta ahora no hay evidente relación reportada entre el volumen de las cavidades y la reversibilidad, por lo que es un campo de estudio para el futuro.

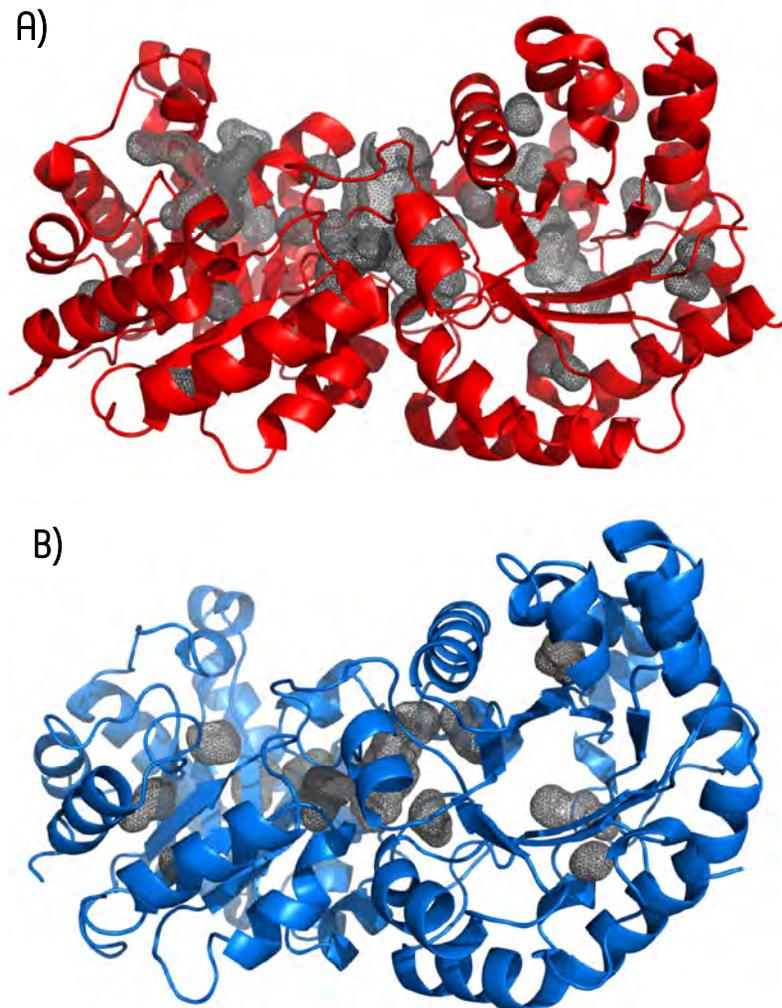


Figura 42. Cavidades internas en la estructura tridimensional de las BacTIMs. La estructura y cavidades mostradas son de: **A)** *G* ϕ TIM, un ejemplo de IrrevTIM y **B)** *C* ρ TIM, un ejemplo de RevTIM. El volumen total de las cavidades en el oligómero es 4010 Å³ y 2241 Å³, respectivamente.

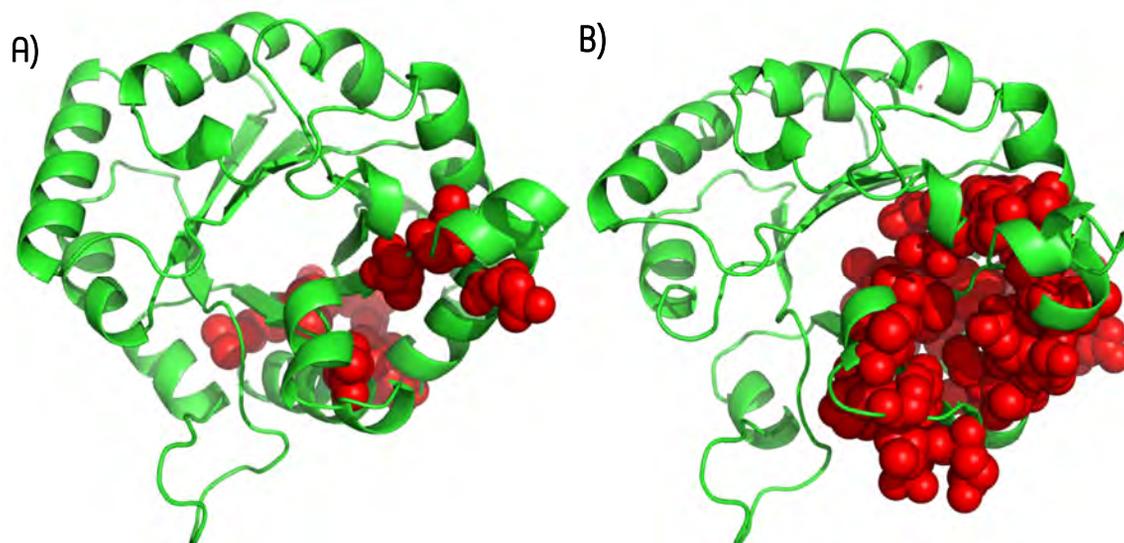


Figura 43. Comparación estructural entre RevTIMs e IrrevTIMs. A) *Go*TIM y **B)** *Cp*TIM para distinguir la presencia de cavidades y el volumen presente en éstas. En rojo se indican los aminoácidos presentes en la cavidad principal de la proteína.

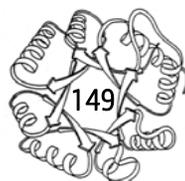
Propiedades fisicoquímicas de la transición desplegamiento/replegamiento.

Al comparar los valores de la T_m para las TIMs estudiadas, se encontró que de forma general las RevTIMs muestran una T_m baja en comparación con las IrrevTIMs. Sin embargo, es de notar que para *Cp*TIM no se muestra así y no aplica esta relación establecida (tabla 21). Las posibles razones de que las RevTIMs muestren una T_m baja es que es más común encontrar transiciones reversibles en proteínas que poseen un valor de T_m bajo ya que a altas temperaturas (>75 °C) y largos tiempos de desnaturalización (de forma general mayor a 10 minutos) se promueven las modificaciones químicas de los aminoácidos (como la desaminación, descarboxilación y descomposición) y esto evita un replegamiento favorable (Creighton, 1996; Fersht, 1999).

Con respecto a la estabilidad termodinámica, los valores de ΔG_{dis} obtenidos de los experimentos de dilución son muy similares en ambos grupos de TIMs y esto sugiere que la estabilidad de la intercara no correlaciona con la reversibilidad (barras blancas, figura 44). La

inconsistencia de los valores de disociación entre experimentos de dilución y los obtenidos mediante desplegamiento químico pudieran deberse a que en los trabajos previamente reportados, en ningún caso se probó la reversibilidad del proceso de dilución, por lo que los valores previamente son aparentes.

Al comparar la estabilidad total del dímero esto es, el valor de ΔG_{tot} a 25 °C entre dichas TIMs, se observa que los valores para las TIMs bacterianas cuyo desplegamiento térmico es reversible son más bajos, casi de la mitad, en comparación con las TIMs que se despliegan irreversiblemente (figura 44). Debido a esto, es razonable considerar que existe una relación entre reversibilidad y baja estabilidad conformacional. En estos momentos desconocemos las causas que subyacen a este comportamiento, pero creemos que esta baja estabilidad pudiera ser una de las razones de la reversibilidad en el proceso de desnaturalización térmica, es decir, proponemos un “*pago fisicoquímico*” que otorgan estas proteínas para tener reversibilidad en su proceso de desplegamiento térmico.



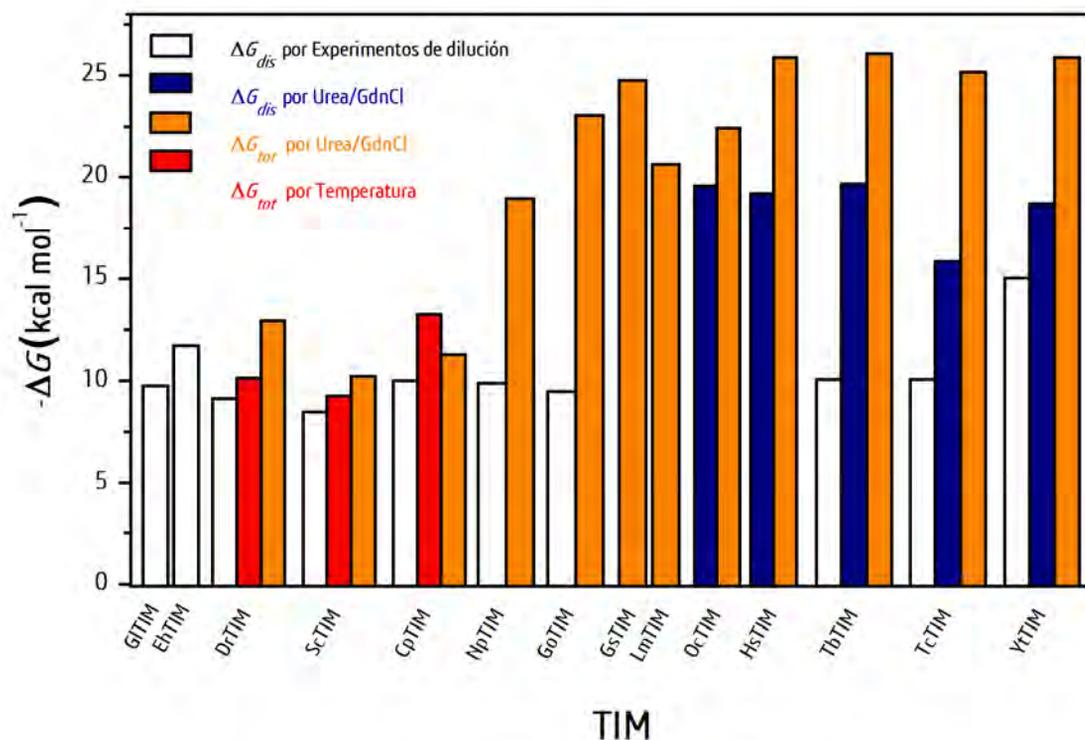
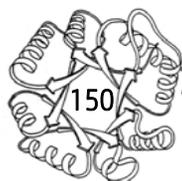


Figura 44. ΔG total y de asociación de las TIMs bacterianas y de otras TIMs eucariontes reportadas hasta el momento. Las barras rojas y naranjas indican el ΔG total a 25 °C, las barras azules son el ΔG a 25 °C y las barras blancas muestran el ΔG derivado de experimentos de dilución.

Se puede observar que la naturaleza del agente desnaturizante y la técnica empleada no cambia la tendencia de ese pago fisicoquímico propuesto. Aun con estas discrepancias en el ΔG máximo de estabilidad, se puede observar que el valor de estabilidad a 25 °C tiene un componente mayoritario que corresponde al ΔG de asociación, como se explicó anteriormente (alrededor del 80-95%), por lo cual el plegamiento de los monómeros contribuye en menor porcentaje en la estabilidad de la proteína (menos del 20% del total), característica que se mantiene en esta familia de proteínas. Es indudable que la estabilidad conformacional se incrementa en gran medida durante el proceso de oligomerización, ya que la estabilidad del monómero es mucho menor cuando se compara con la energía del dímero. A partir de los datos disponibles en la literatura, se estima que el ΔG total promedio del proceso es aproximadamente de 20-25 kcal mol⁻¹, que corresponde al 100% de la



energía libre total del proceso. Por otro lado, la estabilidad del monómero corresponde únicamente al 15% de la energía total. La baja estabilidad del intermediario monomérico es quizá una de las principales razones de la ausencia de catálisis en el monómero aislado (Saab-Rincón *et al.*, 2001).

El desplegamiento térmico al equilibrio para las RevTIMs es bien descrito por un mecanismo de dos estados y la mayoría de las IrrevTIMs estudiadas muestran una transición cooperativa descrita por un único valor de T_m , por lo tanto, parece ser que no hay intermediarios estables o significativamente poblados involucrados en la reversibilidad del plegamiento. Sin embargo, es bien sabido que las propiedades de agregación de los intermediarios transientes juegan un rol importante en determinar la reversibilidad de las transiciones de plegamiento. Al respecto, el intermediario monomérico observado en el desplegamiento térmico de *Eh*TIM se despliega irreversiblemente (Téllez *et al.*, 2008). Del mismo modo, la evidencia de intermediarios cinéticos se encontró en el desplegamiento térmico altamente irreversible de *Tc*TIM (Mixcoha-Hernández *et al.*, 2007); de esta forma, es probable que la presencia de intermediarios en la vía de plegamiento pudiera favorecer la irreversibilidad del proceso y por tanto vale la pena explorar la posibilidad de un desplegamiento térmico de dos estados y la reversibilidad.

Existen dos escenarios extremos para un proceso de disociación; en el caso de cuerpo rígido, la disociación de las subunidades se realiza sin el desplegamiento de los monómeros; en el otro extremo, la disociación está acoplada con el desplegamiento. Sea cual sea el mecanismo que sucede en la TIM, la reversibilidad es comúnmente asociada con la cooperatividad del proceso de disociación/desplegamiento. Por ejemplo, si los monómeros de la TIM fueran aislados de un dímero nativo por un mecanismo de cuerpo rígido, cada uno de ellos debería exponer un área de aproximadamente 1645 \AA^2 , de la cual cerca del 70% de esta área es hidrofóbica, incrementando las posibilidades de agregación; por lo tanto, la presencia de intermediarios monoméricos estables podría reducir la reversibilidad, reafirmando lo que previamente se planteó sobre los intermediarios. Los análisis estructurales para varias TIMs muestran que la asociación de los monómeros para formar

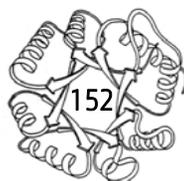


un dímero nativo contempla alrededor de 5-6% del total del ΔASA calculado para el proceso donde los monómeros desplegados forman un dímero plegado. De las parametrizaciones estructurales se estimó que, para un cuerpo rígido, la disociación contribuye entre un 5-6% del total del ΔC_p . En contraste, los valores experimentales muestran que la disociación contribuye alrededor del 75% de la estabilidad del dímero, además, la disociación representa el 55% del total de cambio de entalpía. Todo esto permite concluir que la disociación del dímero no es un proceso de cuerpo rígido (Télez *et al.*, 2008) y por lo tanto existe un acoplamiento de la disociación y el desplegamiento de los monómeros, tal cual se ha visto experimentalmente en algunos casos.

En conformidad con lo anterior, $D\alpha$ TIM, $C\beta$ TIM y $S\alpha$ TIM muestran un comportamiento extremadamente cooperativo en el desplegamiento, esto es, su desplegamiento térmico es un proceso de dos estados donde los monómeros no son estables por sí mismos; esto concuerda con que el ΔG_{dis} y ΔG_{tot} sean similares en estas proteínas. Por todo esto, es probable que la reversibilidad pueda relacionar con un comportamiento de dos estados. No obstante, debe ser notado que el desplegamiento térmico para las IrrevTIMs muestra una transición simple, por lo tanto, una cooperatividad extrema por sí misma no es suficiente para determinar la reversibilidad del plegamiento.

Búsqueda de nuevas secuencias para estudiar

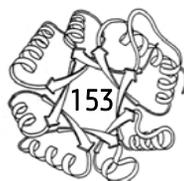
Por último, se llevó a cabo la búsqueda de nuevas secuencias para el análisis de las causas que subyacen a la reversibilidad en el desplegamiento térmico. Sabemos que sobre las relaciones filogenéticas, *S. coelicolor* y *C. perfringens* pertenecen a los fila *Actinobacteria* y *Firmicutes*, respectivamente. Ambos fila son del grupo Posibacteria, mientras que *D. radiodurans* pertenece al filo *Deinococcus-Thermus*, parte del grupo Eobacteria, un grupo distante. Por eso, la reversibilidad parece no estar restringida a un grupo filogenético en particular. Debido a esto, realizamos una



búsqueda para encontrar nuevos candidatos potenciales a estudiar. Uno de los acercamientos iniciales fue ubicar otras proteínas filogenéticamente cercanas a los grupos donde se encontraron TIMs reversibles (en especial al grupo Posibacteria donde se encuentran *S. coelicolor* y *C. perfringens*) tanto parecidas como distintas (alta y baja identidad) para tener una población más grande que nos sirva para analizar las propiedades que hemos propuesto que se podrían relacionar con la reversibilidad. Además, esto estaría complementado con la reconstrucción de proteínas ancestrales en estos fila bacterianos para ver si somos capaces de encontrar una relación evolutiva de la reversibilidad en estos grupos de bacterias. Hasta ahora se tienen recolectadas 1430 secuencias que nos ayudará en el caso de que decidamos continuar con esta segunda opción y realizar el análisis de nuevas proteínas existentes y la caracterización de algunos ancestros de las TIMs bacterianas.

La búsqueda se realizó utilizando el protocolo Rd.HMM publicado previamente (Martínez-Castilla y Rodríguez-Sotres, 2010). Para esto, se pusieron dos anzuelos: *Sc*TIM como ejemplo de RevTIM y *Go*TIM como ejemplo de IrrevTIM. Ambas estructuras se rasuraron y se quitaron iones, moléculas de agua y dobles conformaciones de aminoácidos (si es que tenían). Posteriormente, se utilizó Rosetta para rediseñar las secuencias con el fin de únicamente mantener la cadena principal de $C\alpha$ y cambiar todo el resto. De este rediseño se obtuvieron 150 archivos con coordenadas que cumplieran con esta característica. Para comprobar que el rediseño fuera correcto, se alinearon algunos pdb's y se observó que en efecto, el RMSD entre las estructuras fue de 0 Å, lo que indica que todas mantienen la misma topología. Además, se alinearon algunas secuencias y de la misma forma se observó que la secuencia era completamente variable, sin conservar regiones claras, lo que indicaba que el rediseño sí exploró un espacio conformacional grande (datos no mostrados).

Después, con estos archivos generados se construyó el modelo de Markov para así realizar la búsqueda en la base de datos de secuencias. Se utilizó la base de datos Uniref90 (actualizada a julio de 2016) que contiene secuencias que presentan un porcentaje de identidad menor al 90%. Con el

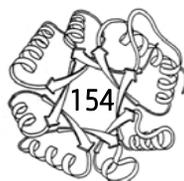


modelo construido se buscaron los resultados para cada anzuelo. Se seleccionaron los primeros 100 resultados para analizar y ubicar las secuencias en un contexto filogenético.

Se observó que en el caso de *G*oTIM, da resultados de especies no estudiadas previamente y presentes en al menos 4 fila diferentes. Además, en toda la lista de resultados no se encuentran las 3 enzimas que son reversibles. Para *Sc*TIM se encontró una cosa interesante: en los primeros resultados se encuentran organismos pertenecientes al mismo género (*Streptomyces*) pero más debajo de la lista aparece *Clostridium perfringens* y una especie de *Deinococcus*, lo cual es relevante porque al parecer se pueden encontrar resultados de TIMs que son reversibles en el desplegamiento térmico. Esto es, cuando se pone de anzuelo una RevTIM, aparecen otros miembros que sabemos que son RevTIMs; mientras cuando se pone un anzuelo de IrrevTIM, no se obtiene ninguna RevTIM, por lo que pareciera que el método es efectivo para encontrar nuevos candidatos a estudiar. Queda analizar a profundidad los resultados y ver las secuencias en un contexto filogenético para ver si es posible observar alguna región en donde aparezcan la mayor cantidad de secuencias. Además, de las secuencias obtenidas, del grupo de RevTIMs todas cumplen con el criterio de tener un punto isoeléctrico bajo ($\mu=5.2 \pm 0.3$) por lo que al parecer, los determinantes propuestos pudieran relacionar efectivamente con la reversibilidad en el desplegamiento térmico.

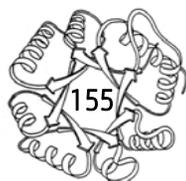
Discusión final sobre el plegamiento de los barriles TIM naturales

De toda la discusión anterior acerca de las posibles causas de la reversibilidad, parece ser que hay un balance delicado entre las diversas contribuciones que permiten la reversibilidad en el desplegamiento térmico de las enzimas oligoméricas y, en específico, de los barriles TIM. Todas estas correlaciones entre las propiedades moleculares y la reversibilidad fueron obtenidas del análisis de la estabilidad y la estructura de un número limitado de proteínas. Claramente, son necesarios más



datos cuantitativos procedentes de una caracterización fisicoquímica de proteínas procedentes de fila no explorados.

Bajo este fin de tener más información del tema y con la perspectiva de comprobar si las 5 propiedades propuestas en realidad se relacionan con la reversibilidad, seleccionamos dos TIMs bacterianas más para analizar, una que correspondiera, bajo nuestros criterios propuestos, a una RevTIM y otra a una IrrevTIM. Los criterios para la selección fue que las propiedades que se pudieran predecir con los datos hasta el momento reportados estuvieran claramente dentro de alguno de los dos grupos, como lo son el punto y el volumen de cavidades; debido a esto se eligieron proteínas con la estructura tridimensional determinada. Además, para poner a prueba también la idea de la temperatura óptima de crecimiento y su relación con la estabilidad, se seleccionaron una proteína termófila y una psicrófila. Para el grupo de las IrrevTIM, se seleccionó la TIM del organismo termófilo *Thermus thermophilus* (*TtTIM*), cuyo punto isoeléctrico es de 5.54 y su volumen de cavidades calculado a partir de la estructura tridimensional es de 3224 \AA^3 . Para el caso de las RevTIMs, se seleccionó la TIM del organismo psicrófilo *Moritella marina* (*MmTIM*), con un punto isoeléctrico de 4.8 y volumen de cavidades estimado de 2927 \AA^3 . Las características fisiológicas y bioquímicas de ambas especies se encuentran en el apéndice B). La razón de la elección de la proteína fue porque cuando se analizaron todas las propiedades asociadas a la reversibilidad, se encontró que una constante fue que *MmTIM* siempre se encontraba en la sección de las propiedades que caracterizaban a las proteínas con desplegamiento térmico reversible aunque su desplegamiento había sido reportado como irreversible (Álvarez, *et al.*, 1998). En otras palabras, en nuestras predicciones, *MmTIM* debería estar dentro del grupo de las RevTIMs (figura 45). Debido a esto, decidimos solicitar la síntesis del gen codificante para esta proteína en la misma compañía donde se pidieron las otras BacTIMs y, con éste, caracterizar el desplegamiento de la enzima con mayor profundidad, para comprobar o descartar nuestra hipótesis de la reversibilidad.



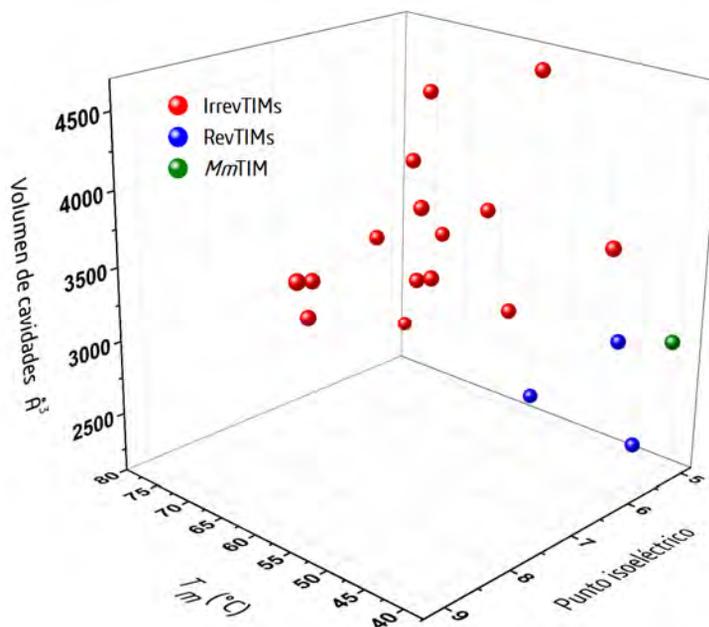
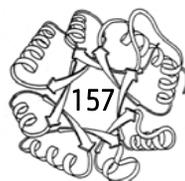


Figura 45. Gráfica 3D que muestra la relación de las 3 propiedades cuantitativas elegidas que se relacionan con la reversibilidad. En rojo se muestra el grupo de las IrrevTIMs, en azul están indicadas las RevTIMs y en verde está indicada *MmTIM*. Nótese que el punto verde se encuentra en el grupo de las RevTIMs.

Las proteínas se subclonaron en un vector pET-28b(+) como todas las BacTIMs y se trabajaron bajo las mismas condiciones. Después de purificarlas, se caracterizó el desplegamiento térmico de las dos enzimas. En el caso de *MmTIM*, contrario a lo reportado previamente, se encontró que el desplegamiento es reversible y que, bajo las condiciones experimentales ocupadas, el sistema está en equilibrio. Lo anterior se puede observar porque se recupera el espectro de DC una vez que se calienta y se enfría la muestra, lo que indica que la enzima se replegó (figura 46a). Además, las rampas de temperatura coinciden la ida con el regreso del experimento y muestran valores de T_m casi idénticos (rampa desplegamiento: 36.9 °C; rampa replegamiento: 36.8 °C) (figura 46b), indicando que el proceso está en equilibrio. También, los experimentos mediante DSC muestran una reversibilidad mayor al 85% (figura 46c), comprobando que el proceso de desplegamiento térmico es reversible; las trazas calorimétricas se ajustan correctamente a un modelo de dos estados con disociación ($N2 \rightleftharpoons 2D$), lo que permitió obtener los parámetros termodinámicos necesarios (ΔC_p , ΔH

y T_m) para construir la curva de estabilidad de la enzima (figura 66) y realizar la caracterización completa. La caracterización fisicoquímica encontró que la T_m de *Mm*TIM es baja (33.8 °C) y su ΔG a 25 °C es de tan sólo 2 kcal mol⁻¹, presentando una curva de estabilidad completamente pequeña, con una estabilidad marginal, la más baja que hasta ahora tenemos conocida para una TIM. Con este resultado, se confirma que *Mm*TIM en realidad sí es parte del grupo de las RevTIMs, a pesar de que estaba reportada como IrrevTIM; tal vez la proteína utilizada no era fresca y por tanto ya estaba dañada, lo que provocaría que la desnaturalización térmica fuera irreversible. Por todo lo anterior, se obtiene una cuarta proteína con desplegamiento térmico reversible y, lo más importante, reafirma las 5 propiedades que encontramos que correlacionan con la reversibilidad: 1) la proteína tiene un bajo punto isoeléctrico (4.8), 2) la proteína tiene baja T_m (\approx 34 °C), 3) el estado desplegado indica poca estructura residual (figura 45a), 4) la proteína presenta pequeño volumen en las cavidades internas (2927 Å³) y 5) la estabilidad de proteína es reducida $\Delta G_{25^\circ C} \approx$ 2 kcal mol⁻¹.



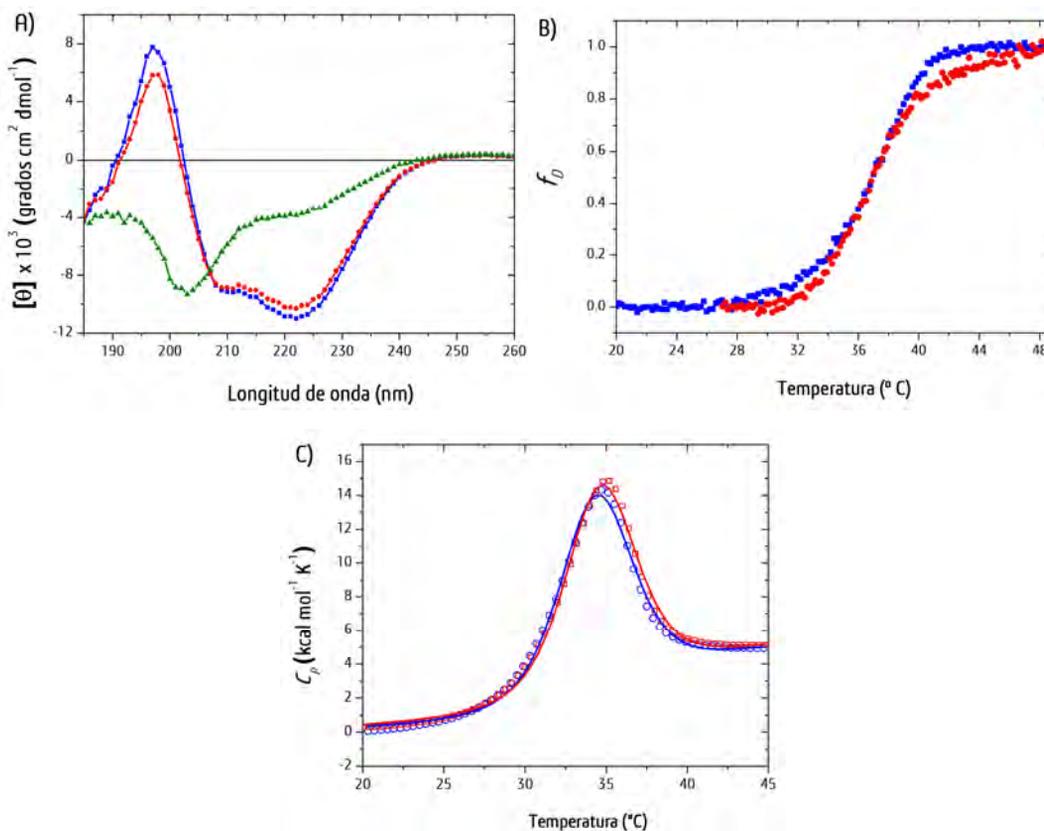
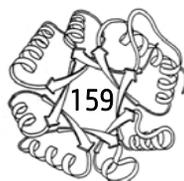


Figura 46. Despliegamiento térmico de *MmTIM*. A) Espectros de DC: nativo (■), desnaturalizado (▲) y nativo después de desnaturalizar (●). B) Despliegamiento inducido por temperatura: nativo a desplegado (●) y desplegado a nativo (■). C) Desnaturalización mediante DSC: primera endoterma (■) y segunda endoterma (●). Las líneas continuas muestran el mejor ajuste a un modelo de disociación de dímeros de dos estados ($N_2 \rightleftharpoons 2D$).

Posteriormente, se analizó el despliegamiento térmico de *TtTIM*. Al inicio, en una rampa normal de temperatura que fue de los 20-97 °C, no se observó transición alguna (figura 47c). Debido a esto, se trató de estudiar el despliegamiento agregándole a la enzima un poco de urea (3 M), sin embargo, tampoco se observó transición en presencia de este agente caotrópico. Debido a esto, se prosiguió a realizar el despliegamiento térmico en presencia de GndHCl. De esta forma, se realizaron desnaturalizaciones térmicas en presencia de 3, 4, 5, 6 y 7 M de GndHCl. En todas ellas se observaron

transiciones simples, altamente cooperativas (figura 47c) e irreversibles (figura 47a). Posteriormente, de cada traza se obtuvo un valor de T_m aparente y se graficaron en dependencia de la concentración de desnaturalizante para así obtener un valor extrapolado a 0 M de concentración de desnaturalizante (figura 47d). El valor obtenido fue de 101.93 °C con una R^2 de la recta de 0.9928. Cabe aclarar que siempre se partió de proteína nativa, como se puede observar al comparar los espectros de DC en cada concentración de GndHCl y que son idénticos a la proteína nativa, al menos en el rango de longitud de onda observado (260-212 nm), ya que a menores longitudes de onda, el buffer con GndHCl ya imprime gran cantidad de ruido en la muestra (figura 47b). El valor de T_m obtenido es el más alto reportado para una TIM, incluso, para la mayoría de las proteínas estudiadas hasta el momento. Con estos datos, se prosiguió a realizar experimentos de desnaturalización térmica mediante DSC. En todos los casos la transición fue irreversible (figura 47e) corroborando lo observado en los experimentos de DC. Además, la T_m calculada fue de 101.62 °C, lo que está en completa coincidencia con los valores extrapolados con los experimentos de DC en presencia de GdnHCl. Las trazas a distintas velocidades de barrido se ajustaron al modelo de dos estados irreversible $N \rightarrow F$ (figura 47f) y se observó que hay un desplazamiento de la T_m al aumentar la velocidad de barrido (≈ 5 °C), lo que comprueba que el desplegamiento térmico de esta enzima está controlado cinéticamente (figura 47f). Además, de estos ajustes se calculó la energía de activación ($132 \text{ kcal mol}^{-1}$), comparándolos con los de las otras TIMs reportadas y se puede observar que el valor está en el rango para las otras TIMs, asemejando energéticamente las barreras cinéticas que están presentes en el proceso. De lo anterior, parece ser que las 5 propiedades propuestas sí relacionan con la reversibilidad, en embargo, habrá que estudiar más a fondo para determinar en qué proporción y de qué forma están correlacionadas entre ellas (si es que lo están).



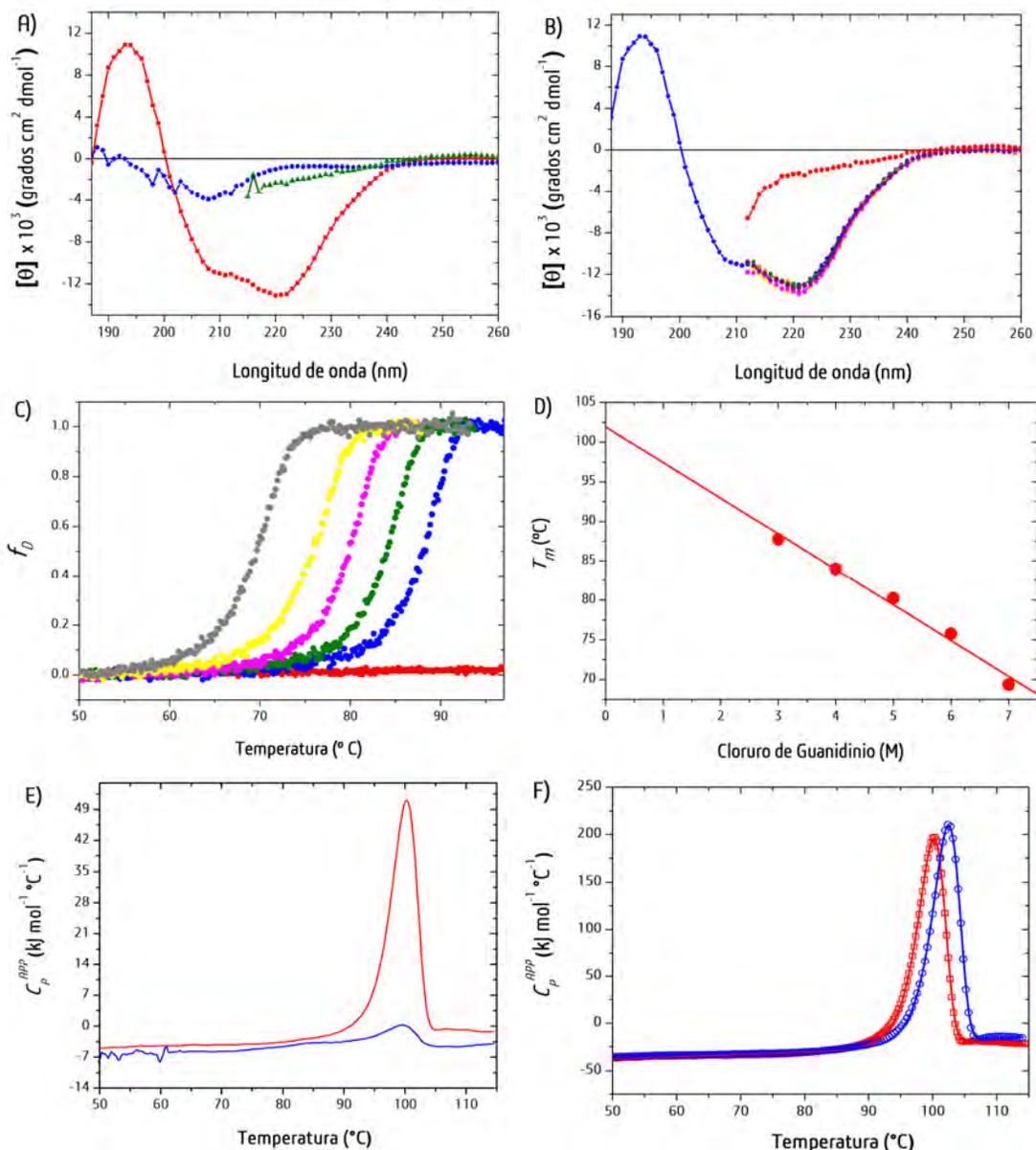
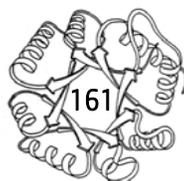


Figura 47. Despliegamiento térmico de 7TIM. **A)** Espectros de DC: nativo (—■—), desnaturalizado (—▲—) y “nativo” después de desnaturalizar (—●—). **B)** Espectros de DC: proteína nativa (0 M de GndHCl) a 25 °C (curva azul), proteína desnaturalizada a 90 °C y 7 M de GndHCl (curva roja), proteína con 3, 4, 5, 6 y 7 M GndHCl (curvas verde, naranja, morada, magenta y amarilla, respectivamente). **C)** Las curvas presentan cantidades de GndHCl incrementadas: 0 M (proteína nativa; curva roja), 3 M (curva gris), 4 M (curva amarilla), 5 M (curva magenta), 6 M (curva verde) y 7 M (curva azul). **D)** Dependencia de la T_m con respecto a la concentración de GndHCl. Se muestra la recta de ajuste a los datos para la extrapolación de una T_m a 0 M de desnaturalizante (R^2 : 0.9928). **E)** Desnaturalización mediante DSC: primera endoterma (curva roja) y segunda endoterma (curva azul). **F)** Desnaturalización mediante DSC: endoterma a 30 K h⁻¹ (—■—) y endoterma a 180 K h⁻¹ (—●—). Las líneas continuas muestran el mejor ajuste a un modelo de dos estados irreversible (N → F).

Una perspectiva clara para obtener una descripción detallada y atomística del proceso de plegamiento/desplegamiento de la TIM es el uso de dinámica molecular, ya que esta metodología es una herramienta invaluable para conocer las interacciones moleculares con un detalle atomístico muy fino. Dentro de esta rama, en específico la dinámica molecular dirigida ha sido exitosamente utilizada para la estimación de los parámetros de proteína-ligando (Li *et al.*, 2011) y la energética de la interacción proteína-proteína (Cheung *et al.*, 2015). El acercamiento de la dinámica molecular aplicada al estudio del desplegamiento *in silico* es un área prometedora (Abriata y Dal Peraro, 2015) y podrá dar información de los cambios conformacionales que toman lugar durante la disociación del dímero y el desplegamiento de los monómeros, así como su rol en la reversibilidad del proceso. Además, una caracterización cinética del desplegamiento térmico de las RevTIMs debería traer más información con respecto a la vía de desplegamiento y sus propiedades fisicoquímicas.

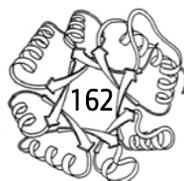
Los resultados aquí presentados muestran que el desplegamiento térmico de una misma topología dada, como lo es el barril TIM, puede exhibir un control cinético o termodinámico en su plegamiento. Ha sido sugerido que la estabilidad cinética es más fuertemente seleccionada por la evolución que la estabilidad termodinámica porque la primera facilita procesos de alteración irreversibles incluso cuando se mantiene una alta estabilidad termodinámica (Sánchez-Ruiz, 2010). La estabilidad y la reversibilidad son biotecnológicamente relevantes ya que el modular la estabilidad de proteínas utilizadas en diversos procesos podría traer grandes beneficios tanto industriales como económicos (como el de la estabilización de enzimas útiles en la industria alimentaria, la producción de compuestos químicos o la estabilización de fármacos proteicos). Debido a esto, el estudiar a profundidad estos procesos nos arrojarán información acerca de cómo podemos modificar racionalmente las propiedades fisicoquímicas de las proteínas y, en específico, de los barriles TIM, quienes presentan una alta relevancia en el mundo proteico.



Para todos los organismos vivos, desde bacterias a mamíferos, la temperatura es probablemente el agente de estrés más común y natural a nivel molecular; de hecho, la vida a "temperatura constante" es un evento evolutivo reciente (Lindquist, 1986; Muga y Moro, 2008). Diversos mecanismos moleculares, como el sistema de chaperoninas, han evolucionado para ayudar a las células a lidiar con el estrés por temperatura (Gething y Sambrook, 1992; Ryabova *et al.*, 2013; Bozaykut *et al.*, 2014). Sin embargo, la TIM no ha sido identificada como un sustrato para el sistema GroEL, aunque los barriles TIM son sustratos comunes del sistema de chaperoninas en *E. coli* (Kerner *et al.*, 2005).

Los barriles TIM oligoméricos son frecuentes en las células y debido a esto, la reversibilidad en su proceso de plegamiento/desplegamiento pudiera ser una ventaja evolutiva, ya que esta propiedad reduce tanto la probabilidad de agregación irreversible no productiva así como la carga de trabajo de la maquinaria celular de degradación y del sistema de chaperoninas. Sin embargo, investigaciones con mayor detalle son necesarias para comprobar esto, ya que aún permanece incierto si la reversibilidad tiene una relevancia biológica específica en ambientes celulares *in vivo*.

Debido a todo lo anteriormente discutido, que una topología tan común, como lo es el barril TIM, pueda desplegarse y replegarse sin la ayuda de la maquinaria celular es una característica biológicamente relevante que requiere de mayor análisis. La conclusión de que las tres proteínas reversibles proceden de organismos de diferentes fila sugiere que la reversibilidad en el desplegamiento térmico puede ser más común de lo que se conoce actualmente y, por tanto, es un campo para seguir estudiándolo en el futuro y del cual puede obtenerse información global sobre la estabilidad y el plegamiento de las proteínas.

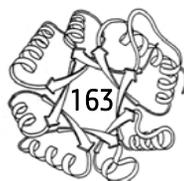


Parte 2: El paisaje conformacional de barriles TIM diseñados *de novo*

En búsqueda de incrementar la estabilidad de sTIM11: diseño de nuevos barriles TIM

Una vez descritas las propiedades de los barriles TIM naturales, esto es, “modelados por la evolución”, nos adentramos al estudio de los barriles TIM “modelados *in silico*.” Desde la perspectiva del diseño, sTIM11 es sin duda una proteína interesante ya que es el primer barril TIM diseñado *de novo*. Esta proteína presenta, sin embargo, algunos inconvenientes desde la perspectiva del plegamiento y la estabilidad. En primer lugar, sTIM11 fue diseñada con dos cisteínas posicionadas para formar un puente disulfuro, aunque en la estructura cristalográfica este puente no se observa. Las cisteínas libres pueden dar lugar a puentes disulfuro intercadena, los cuales complican la interpretación de los estudios de estabilidad y disminuyen la reversibilidad en los procesos de desplegamiento/replegamiento. Por otra parte, la cisteína es un nucleófilo interesante para su introducción en sitios activos. El objetivo central de esta parte del trabajo es modificar la secuencia de sTIM11 para aumentar su estabilidad, para posteriormente utilizar a la proteína como plataforma para el diseño de enzimas; de esta manera, decidimos trabajar tomando como base a sTIM11-SS (que podría llamarse como ReTIM0), una versión de sTIM11 sin cisteínas, en la cual la simetría de orden 4 se conserva (ver figura 13). De manera interesante, sTIM11-SS presenta una T_m mucho menor que sTIM11, lo cual implica que aun sin formar el puente disulfuro, las cisteínas estabilizan al estado nativo de la molécula.

Al observar la estructura de sTIM11 es posible observar algunas cavidades grandes, las cuales se encuentran principalmente en el núcleo interior del barril formado por las hebras β y el núcleo periférico formado por la cara exterior de las hebras β y la cara interna de las hélices α . De esta manera, los diseños se realizaron incrementando el empaquetamiento en distintas regiones del barril (figura 48) e introduciendo de forma racional algunas interacciones estabilizantes como puentes salinos o disulfuro.



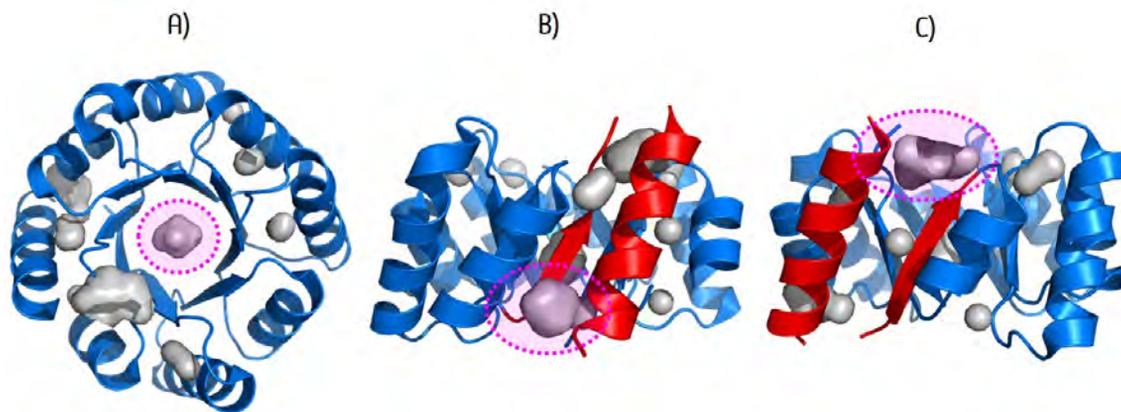
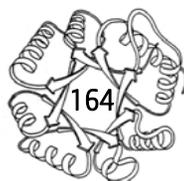


Figura 48. Regiones de sTIM11 que fueron rediseñadas para aumentar la estabilidad. A) Región del núcleo interno, **B)** núcleo periférico A y **C)** núcleo periférico B.

Todas las proteínas diseñadas se nombraron con el prefijo “ReTIM” (de TIM rediseñada) seguido del número consecutivo de diseño experimental analizado (la ReTIM0 fue sTIM11-SS, la proteína madre de todos los diseños). El tamaño de todas las proteínas diseñadas es de 184 residuos (a excepción de ReTIM 15 y 16, las cuales tienen 185 y 188 residuos, respectivamente), al igual que sTIM11 y sTIM11-SS. A continuación se muestran los diseños que se obtuvieron en cada grupo y el listado de las mutaciones para cada ReTIM (tabla 25; para mayor detalle revisar el apéndice C y D):

1. Diseños con mutaciones únicamente en el núcleo interno del barril: de este grupo surgieron las variantes ReTIM1, ReTIM2, ReTIM3 y ReTIM4.
2. Diseños con mutaciones en el núcleo periférico A: este grupo contiene a ReTIM5, ReTIM6 y ReTIM7.
3. Diseños con mutaciones en el núcleo periférico B: este grupo contiene a ReTIM8, ReTIM9 y ReTIM10.
4. Diseños con mutaciones en sitios específicos:



- ReTIM11: esta variante consiste en probar el efecto de una metionina en la posición en la que se encontraba la cisteína original de sTIM11.
 - ReTIM12 y ReTIM13: variantes con dos mutaciones para colocar prolinas: K2P y W27P. Estos diseños fueron realizados para comprobar el efecto de las prolinas sugeridas en varios de los diseños que realizó Rosetta.
 - ReTIM14: variante con una mutación en A45R, con el fin de generar un puente salino entre las hebras β pares y la siguiente hélice α non.
 - ReTIM15 y ReTIM16: variantes para la formación de nuevos puentes disulfuros.
5. Diseños quimeras de los grupos anteriores. Como se podrá ver más adelante, ReTIM1 fue el diseño más exitoso del primer grupo, ReTIM6 del segundo y ReTIM8 del tercero (aunque también se integraron algunas otras ReTIMs como la 7, 12 y 14, las cuales no habían sido posible sobreexpresar). Por esta razón, se combinaron estas proteínas con otras variantes obteniendo las quimeras indicadas en la tabla 24.

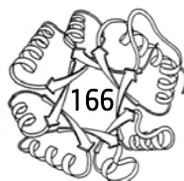
Tabla 24. Quimeras de ReTIMs.

ReTIM	Tipo	Componentes de la quimera
ReTIM17	Doble	ReTIM1 + ReTIM8
ReTIM18	Doble	ReTIM6 + ReTIM8
ReTIM19	Doble	ReTIM7 + ReTIM8
ReTIM20	Doble	ReTIM11 + ReTIM8
ReTIM21	Doble	ReTIM12 + ReTIM8
ReTIM22	Doble	ReTIM14 + ReTIM8
ReTIM23	Doble	ReTIM1 + ReTIM6
ReTIM24	Triple	ReTIM1 + ReTIM6 + ReTIM8
ReTIM25	Cuádruple	ReTIM1 + ReTIM6 + ReTIM8 + ReTIM14
ReTIM26	Triple	ReTIM1 + ReTIM8 + ReTIM14
ReTIM27	Triple	ReTIM6 + ReTIM8 + ReTIM14
ReTIM28	Triple	ReTIM1 + ReTIM6 + ReTIM14



Tabla 25. Mutaciones de las ReTIMs.

Proteína	Número de residuos	Simetría (orden)	Número de mutaciones en unidad repetitiva (total)	Mutaciones	Tipo de mutación*
sTIM11	184	Pseudo 4	2 (2)	Q8C, V181C	Generación de un puente disulfuro
sTIM11-SS (ReTIM0)	184	4	NA	Proteína silvestre	NA
ReTIM1	184	2	9 (18)	A21V, R23I, W34V, A38G, A67W, R69I, R80V, A84G, I86V	SC / CP-NP / SC / CP-NP / A-NP / SC / NP-A / CP-NP / CP-NP / SC / SC
ReTIM2	184	2	11 (22)	A21V, R23Q, W34V, A38G, I42M, A67M, R69I, R80V, A84G, I86F, I88V	SC / CP-NCP / A-NP / SC / SC / CP-NP / CP-NP / SC / NP-A / SC
ReTIM3	184	2	12 (24)	A21M, R23V, W34V, A38G, I40S, I42L, A67L, R69K, R80V, A84G, I86T, I88V	SC / CP-NP / A-NP / SC / SC / SC / SC / CP-NP / SC / CP-NP / SC
ReTIM4	184	2	11 (22)	A21M, R23F, W34V, A38G, I40L, I42V, A67L, R69I, R80V, A84G, I88L	SC / CP-A / A-NP / SC / SC / SC / SC / CP-NP / CP-NP / SC / SC
ReTIM5	184	4	4 (16)	E15L, K31R, W34V, A38G	CN-NP / SC / A-NP / SC
ReTIM6	184	4	6 (24)	Q11I, E15L, T18K, K31Q, W34L, A38G	NCP-NP / CN-NP / NCP-CP / CP-NCP / A-NP / SC
ReTIM7	184	4	6 (24)	Q11I, E15L, T18K, Q31D, W34R, A38G	NCP-NP / CN-NP / NCP-CP / NCP-CN / A-CP / SC
ReTIM8	184	4	7 (28)	K2V, A5M, W6L, Y22V, D29I, W34V, A38G	CP-NP / SC / A-NP / A-NP / CN-NP / A-NP / SC
ReTIM9	184	4	6 (24)	K2P, W6Q, Y22I, D29Y, W34V, A38G	CP-NCP / A-NCP / A-NP / NCP-NP / CN-CP / A-NP / SC
ReTIM10	184	4	7 (28)	K2S, W6I, Y22V, S24M, D29K, W34V, A38G	CP-NCP / A-NP / A-NP / CN-CP / A-NP / SC
ReTIM11	184	4	3 (12)	Q8M, W34V, A38G	NCP-NP / A-NP / SC
ReTIM12	184	4	3 (12)	K2P, W34V, A38G	CP-NCP / A-NP / SC
ReTIM13	184	4	3 (12)	W27P, W34V, A38G	A-NCP / A-NP / SC
ReTIM14	184	4	3 (12)	W34V, A38G, A45R	A-NP / SC / NP-CP
ReTIM15	185	4	3 (12+1)	D1C, W34V, A38G, inserción N-terminal: D	CN-NCP / A-NP / SC / Generación de un puente disulfuro
ReTIM16	188	4	3 (12+4)	D1A, W34V, A38G, inserción N-terminal: DCAA	CN-NP / A-NP / SC / Generación de un puente disulfuro
ReTIM17 (ReTIM1 + ReTIM8)	184	2	19 (38)	K2V, A5M, W6L, A21V, Y22V, R23I, D29I, W34V, A38G, K48V, A51M, W52L, A67W, Y68V, R69I, D75I, R80V, A84G, I86V	CP-NP / SC / A-NP / SC / A-NP / CP-NP / CN-NP / A-NP / SC / CP-NP / SC / A-NP / NP-A / A-NP / CP-NP / CN-NP / CP-NP / SC / SC
ReTIM18 (ReTIM6 + ReTIM8)	184	4	11 (44)	K2V, A5M, W6L, Q11I, T18K, Y22V, D29I, K31Q, W34L, A38G, K48V	CP-NP / SC / A-NP / NCP-NP / NP-CP / A-NP / CN-NP / CP-NCP / A-NP / SC / CP-NP
ReTIM19 (ReTIM7 + ReTIM8)	184	4	11 (44)	K2V, A5M, W6L, Q11I, E15L, T18K, Y22V, D29I, K31D, W34R, A38G	CP-NP / SC / A-NP / NCP-NP / CN-NP / NP-CP / A-NP / CN-NP / CP-CN / A-CP / SC
ReTIM20 (ReTIM11 + ReTIM8)	184	4	9 (36)	K2V, A5M, W6L, Q8M, Y22V, D29I, W34V, A38G, A45R	CP-NP / SC / A-NP / NCP-NP / A-NP / CN-NP / A-NP / SC / NP-CP
ReTIM21 (ReTIM12 + ReTIM8)	184	4	7 (28)	K2P, A5M, W6L, Y22V, D29I, W34V, A38G	CP-NCP / SC / A-NP / A-NP / CN-NP / A-NP / SC
ReTIM22 (ReTIM14 + ReTIM8)	184	4	8 (32)	K2V, A5M, W6L, Y22V, D29I, W34V, A38G, A45R	CP-NP / SC / A-NP9:F25 / A-NP / CN-NP / A-NP / SC / NP-CP



Proteína	Número de residuos	Simetría (orden)	Número de mutaciones en unidad repetitiva (total)	Mutaciones	Tipo de mutación*
ReTIM23 (ReTIM1 + ReTIM6)	184	2	17 (34)	Q11I, E15L, T18K, A21V, R23I, K31Q, W34L, A38G, Q57I, E61L, A67W, T64K, R69I, K77Q, W80L, A84G, I86V	NCP-NP / CN-NP / NP-CP / SC / CP-NP / CP-NCP / A-NP / SC / NCP-NP / CN-NP / NP-A / NP-CP / CP-NP / CP-NCP / A-NP / SC / SC
ReTIM24 (ReTIM1 + ReTIM6 + ReTIM8)	184	2	27 (54)	K2V, A5M, W6L, Q11I, E15L, T18K, A21V, Y22V, R23I, D29I, K31Q, W34L, A38G, K48V, A51M, W52L, Q57I, E61L, T64K, A67W, Y68V, R69I, D75I, K77Q, W80L, A84G, I86V	CP-NP / SC / A-NP / NCP-NP / CN-NP / NP-CP / SC / A-NP / CP-NP / CN-NP / CP-NCP / A-NP / SC / CP-NP / SC / A-NP / NCP-NP / CN-NP / NP-CP / NP-A / A-NP / CP-NP / CN-NP / CP-NCP / A-NP / SC / SC
ReTIM25 (ReTIM1 + ReTIM6 + ReTIM8 + ReTIM14)	184	2	29 (58)	K2V, A5M, W6L, Q11I, E15L, T18K, A21V, Y22V, R23I, D29I, K31Q, W34L, A38G, A45R, K48V, A51M, W52L, Q57I, E61L, T64K, A67W, Y68V, R69I, D75I, K77Q, W80L, A84G, I86V, A90R	CP-NP / SC / A-NP / NCP-NP / CN-NP / NP-CP / SC / A-NP / CP-NP / CN-NP / CP-NCP / A-NP / SC / NP-CP / CP-NP / SC / A-NP / NCP-NP / CN-NP / NP-CP / NP-A / A-NP / CP-NP / CN-NP / CP-NCP / A-NP / SC / SC / NP-CP
ReTIM26 (ReTIM1 + ReTIM8 + ReTIM14)	184	2	21 (42)	K2V, A5M, W6L, A21V, Y22V, R23I, D29I, W34V, A38G, A45R, K48V, A51M, W52L, A67W, Y68V, R69I, D75I, W80V, A84G, I86V, A90R	CP-NP / SC / A-NP / SC / A-NP / CP-NP / CN-NP / A-NP / SC / NP-CP / CP-NP / SC / A-NP / NP-A / A-NP / CP-NP / CN-NP / A-NP / SC / SC / NP-CP
ReTIM27 (ReTIM6 + ReTIM8 + ReTIM14)	184	4	12 (48)	K2V, A5M, W6L, Q11I, T18K, Y22V, D29I, K31Q, W34L, A38G, A45R, K48V	CP-NP / SC / A-NP / NCP-NP / NP-CP / A-NP / CN-NP / CP-NCP / A-NP / SC / NP-CP / CP-NP
ReTIM28 (ReTIM1 + ReTIM6 + ReTIM14)	184	2	19 (38)	Q11I, E15L, T18K, A21V, R23I, K31Q, W34L, A38G, A45R, Q57I, E61L, A67W, T64K, R69I, K77Q, W80L, A84G, I86V, A90R	NCP-NP / CN-NP / NP-CP / SC / CP-NP / CP-NCP / A-NP / SC / NP-CP / NCP-NP / CN-NP / NP-A / NP-CP / CP-NP / CP-NCP / A-NP / SC / SC / NP-CP

Para mayor detalle revisar el apéndice C y D.

*: El tipo de mutación se indica con respecto a la proteína base. Por ejemplo: "CP-NP" es una mutación que cambió un aminoácido cargado positivamente a un aminoácido no polar.

NA: No aplica.

SC: Sin cambio en polaridad.

NCP: Polar no cargado.

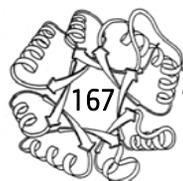
NP: No polar.

CP: Cargado positivamente.

CN: Cargado negativamente.

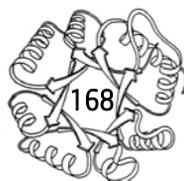
A: Aromático.

NP: No polar.



Sobreexpresión de las ReTIMs

De los 28 diseños seleccionados, no fue posible subclonar la ReTIM7, ReTIM11, ReTIM12, ReTIM14 y ReTIM16 por cuestiones metodológicas que no se lograron resolver. Sin embargo, como se puede ver en los diseños, estas mutaciones se incluyeron en las ReTIMs que presentan mezclas de diseños. Para observar la sobreexpresión de las proteínas, inicialmente se realizaron pruebas en volumen pequeño. Tras la realización de éstas, se encontró que 22 de las 23 proteínas se sobreexpresaron de manera correcta y su peso corresponde al esperado de acuerdo a la secuencia aminoacídica (figura 49). Para ReTIM19 no se logró sobreexpresar la proteína con las condiciones establecidas ni modificando los protocolos y variando temperatura de inducción y 5 cepas de *E. coli* distintas. Debido a esto, todos los resultados siguientes serán únicamente abarcando las 22 proteínas que se lograron sobreexpresar.



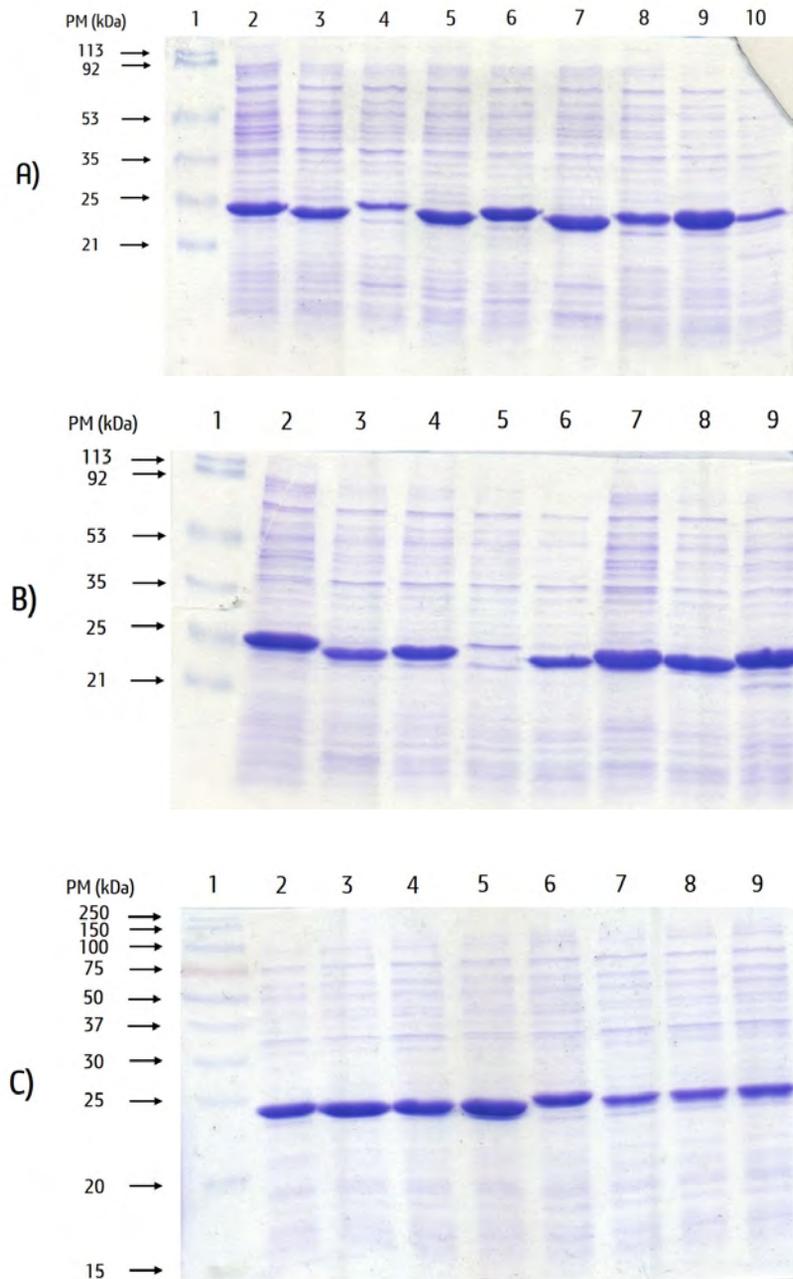
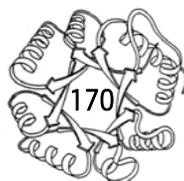


Figura 49. Gel de poliacrilamida SDS-PAGE al 12 % con las pruebas de sobreexpresión de las ReTIMs. **Gel A:** 1) marcador de peso molecular, 2) ReTIM1, 3) ReTIM2, 4) ReTIM3, 5) ReTIM4, 6) ReTIM5, 7) ReTIM8, 8) ReTIM9, 9) ReTIM10, 10) ReTIM13. **Gel B:** 1) marcador de peso molecular, 2) ReTIM15, 3) ReTIM17, 4) ReTIM18, 5) ReTIM19, 6) ReTIM20, 7) ReTIM21, 8) ReTIM22, 9) sTIM11. **Gel C:** 1) marcador de peso molecular, 2) ReTIM6, 3) ReTIM13, 4) ReTIM23, 5) ReTIM24, 6) ReTIM25, 7) ReTIM26, 8) ReTIM27, 9) ReTIM28.

Se sabe que la estabilidad de la proteína es uno de los aspectos importantes en la producción de éstas de forma recombinante. Se ha observado que las proteínas plegadas adecuadamente son estables durante la expresión y la purificación; caso contrario sucede con algunas de las proteínas que son inestables y se producen en una cantidad insuficiente (Bilgimol *et al.*, 2015). Diversos aspectos que afectan a la estabilidad proteica son la secuencia de aminoácidos de la proteína, la construcción génica en la que se inserta el gen, el estrés químico de la célula huésped, las condiciones de expresión, entre otras muchas. De esta forma, la falta de sobreexpresión puede tener su causa en términos de estabilidad, ya que se ha sugerido que proteínas poco estables tienen a degradarse rápidamente para evitar la formación de agregación no-productiva en la célula. Hay casos en que la secuencia de aminoácidos de una proteína en sí misma es propensa a la degradación. Ciertos aminoácidos como Arg, Lys, Leu, Phe, Tyr y Trp en la región N-terminal pueden conducir a la degradación de la proteína (Hosur *et al.*, 2014). Reemplazar estos aminoácidos con aminoácidos compatibles puede mejorar en gran medida la estabilidad de la macromolécula. De igual forma, se sabe que expresar proteínas con etiquetas o proteínas de fusión en muchos casos puede evitar la degradación proteolítica y aumentar la estabilidad (Bilgimol *et al.*, 2015). A futuro, podrían aplicarse algunas de estas estrategias cuando se tengan problemas de sobreexpresión, ya que es claro que la estabilidad de la proteína influye directamente en la sobreexpresión de la misma.

Purificación de las ReTIMs

Las 22 proteínas se purificaron mediante cromatografía de afinidad y exclusión molecular en columna acopladas a FPLC. Como ejemplo se muestra la purificación de ReTIM1 en la figura 50. Los resultados de las 22 proteínas purificadas analizadas en geles de SDS-PAGE se señalan en la figura 51. En todas las proteínas se observó el mismo patrón cromatográfico y de pureza al final de la purificación, lo único que varió ampliamente fue el rendimiento de la purificación, el cual estuvo



entre 6-55 miligramos de proteína purificada por litro de cultivo bacteriano, dependiendo de la ReTIM de la que se trate. Las proteínas purificadas se almacenaron para realizar los experimentos que se describirán a continuación. En todos los casos, el estado estructural cuaternario fue monomérico (con base en experimentos de cromatografía de exclusión molecular).

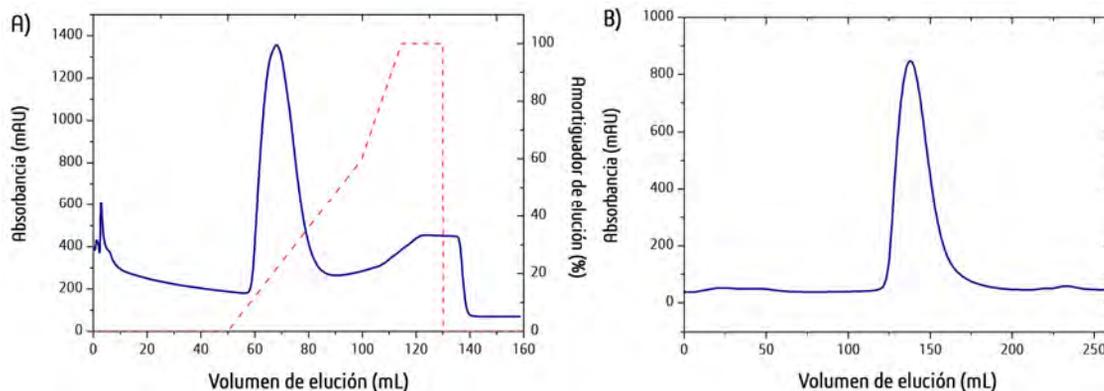


Figura 50. Cromatogramas de la purificación de ReTIM1. A) Cromatograma del primer paso de purificación utilizando una columna de afinidad. El pico a los 60-80 mL muestra la elución de la proteína deseada con etiqueta de histidinas. Amortiguador de elución: amortiguador B. B) Cromatograma del segundo paso de purificación para *D*TIM utilizando una columna de exclusión molecular. El pico a los 125-160 mL muestra la elución de la proteína purificada. Amortiguador de elución: amortiguador G más NaCl 150 mL.

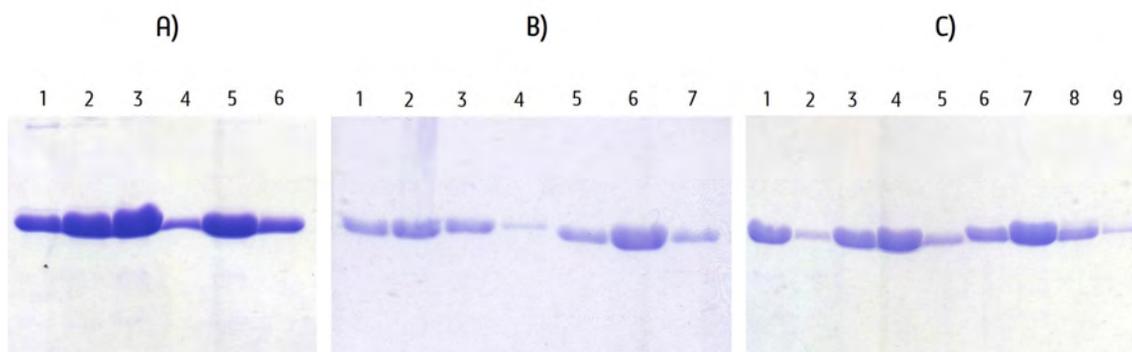


Figura 51. Gel de poliacrilamida SDS-PAGE al 12 % con las ReTIMs purificadas. En todos los casos se muestra la proteína pura tras el paso final de purificación. **Gel A:** 1) ReTIM1, 2) ReTIM2, 3) ReTIM3, 4) ReTIM4, 5) ReTIM5, 6) ReTIM6. **Gel B:** 1) ReTIM8, 2) ReTIM9, 3) ReTIM10, 4) ReTIM13, 5) ReTIM15, 6) ReTIM17, 7) ReTIM18. **Gel C:** 1) ReTIM20, 2) ReTIM21, 3) ReTIM22, 4) ReTIM23, 5) ReTIM24, 6) ReTIM25, 7) ReTIM26, 8) ReTIM27, 9) ReTIM28.

Propiedades espectroscópicas de las ReTIMs: espectro de dicroísmo circular

Para comprobar que las proteínas estaban plegadas adecuadamente, se obtuvo el espectro de DC para todas las ReTIMs, encontrando que las proteínas presentan dos mínimos característicos a 208 y 222 nm, propios de estructuras tridimensionales que contienen α -hélices y hebras- β . En la figura 52 se muestran los espectros de DC para algunas ReTIMs (los diseños más exitosos, como se verá más adelante). Para la mayoría de las ReTIMs se observó un espectro típico para proteínas plegadas, pero para 4 de ellas (ReTIM2, ReTIM5, ReTIM13 y ReTIM15; datos no mostrados) se observó un espectro con los mismos mínimos pero con una señal disminuida de dicroísmo y con un cambio ligero en la forma, lo cual nos habla que tal vez la proteína está plegada pero no del todo compacta (esto se comprueba con los datos de desnaturalización que se presentarán en las secciones siguientes).

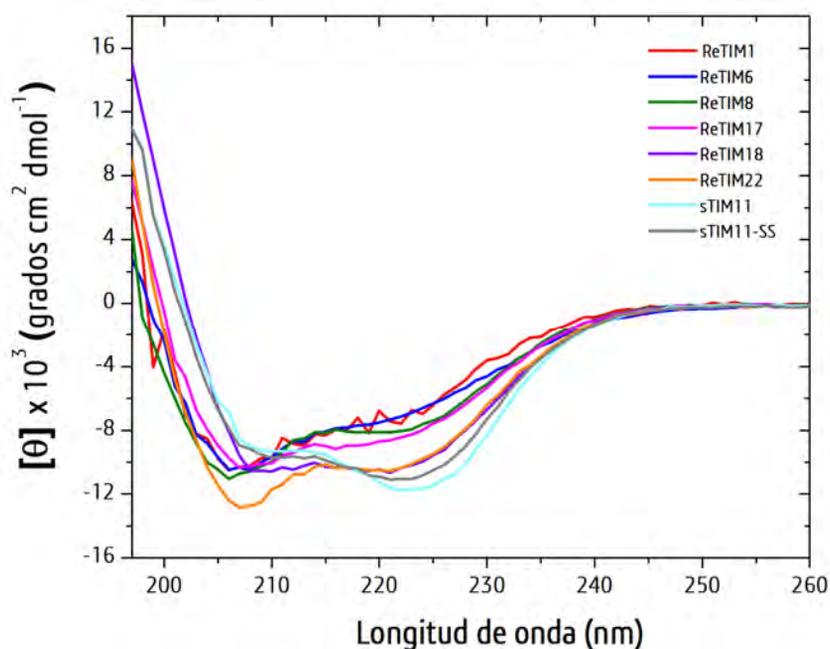
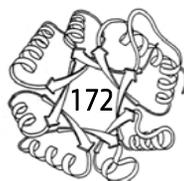


Figura 52. Espectro nativo de dicroísmo circular para algunas ReTIMs.



Propiedades fisicoquímicas de los diseños del núcleo interno: ReTIM1-4

Estas proteínas son variantes del núcleo interno. A diferencia de su diseño original, estas macromoléculas tienen simetría de orden 2, ya que no fue posible generar diseños con orden 4 que satisficieran las restricciones geométricas en las proteínas (tabla 25). Una visión estructural del modelo y de la ubicación de las mutaciones se presenta en la figura 53.

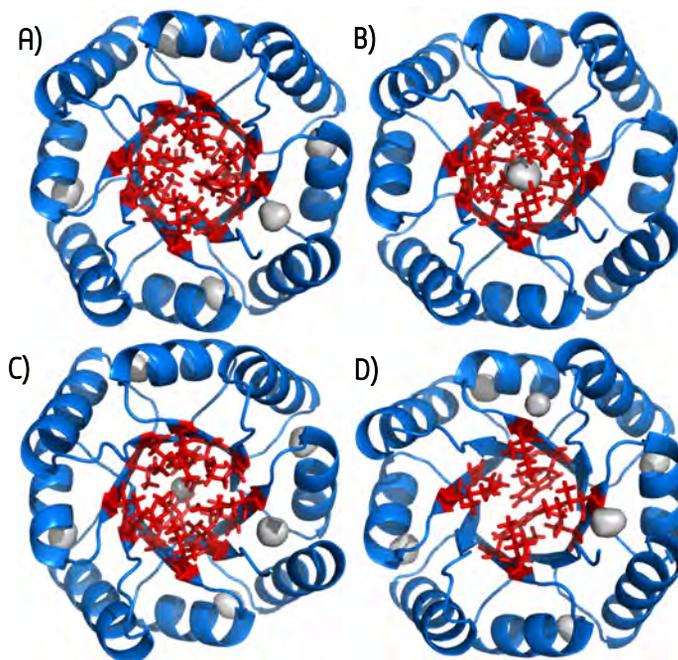


Figura 53. Modelos estructurales de las ReTIM1-4. A) ReTIM1, B) ReTIM2, C) ReTIM3, D) ReTIM4. Las mutaciones se encuentran marcadas en rojo y se indican en la tabla 25.

En esta sección de los barriles TIM diseñados *de novo*, definimos a un diseño como exitoso si mejora una o varias de las propiedades fisicoquímicas que posee la proteína madre: cambio en la entalpía calorimétrica (ΔH), termoestabilidad (T_m), estabilidad conformacional a 25 °C (ΔG), o cooperatividad. Los cambios en estos parámetros serán los que nos guíe para seleccionar a los mejores diseños de los que no lo son.

Con base en lo anterior, el desplegamiento térmico para estas proteínas fue seguido por DC y DSC. Como se puede ver en las figura 54a-b, aún se observa el mismo comportamiento en la pretransición en sTIM11; además, la cooperatividad de las proteínas es baja (observado en ambas técnicas). Si se comparan los valores con la proteína base (tabla 26), se puede concluir que estos diseños no fueron totalmente exitosos, sólo ReTIM1 parece ser el mejor candidato para seguir estudiando, ya que al comparar con sTIM11-SS, las demás proteínas no tienen una mayor T_m y el área de la endoterma es muy pequeña (es decir, un ΔH pequeño) y una baja cooperatividad. Es interesante que en este grupo de diseños se encuentran proteínas con desplegamiento térmico reversible e irreversible (tabla 26), a pesar de compartir una misma topología y todas partir de una proteína madre cuyo desplegamiento térmico es reversible. Adelante se discutirá en conjunto las implicaciones de la reversibilidad en el desplegamiento térmico de los barriles TIM diseñados *de novo*. Cuando se analiza el desplegamiento químico de las proteínas y los datos se ajustan a un modelo de dos estados (figura 54c), se puede ver que la estabilidad a temperatura ambiente no se ve incrementada significativamente en ninguna de ellas, siendo de nuevo ReTIM1 el mejor representante de este grupo, al poseer un ligero incremento del ΔG en comparación con sTIM11, pasando de 4.2 a 4.7 kcal mol⁻¹ (tabla 26). Por todas estas razones, se eligió a ReTIM1 como el diseño exitoso de este grupo de proteínas y fue una de las que se usaron para realizar las quimeras de diseños.



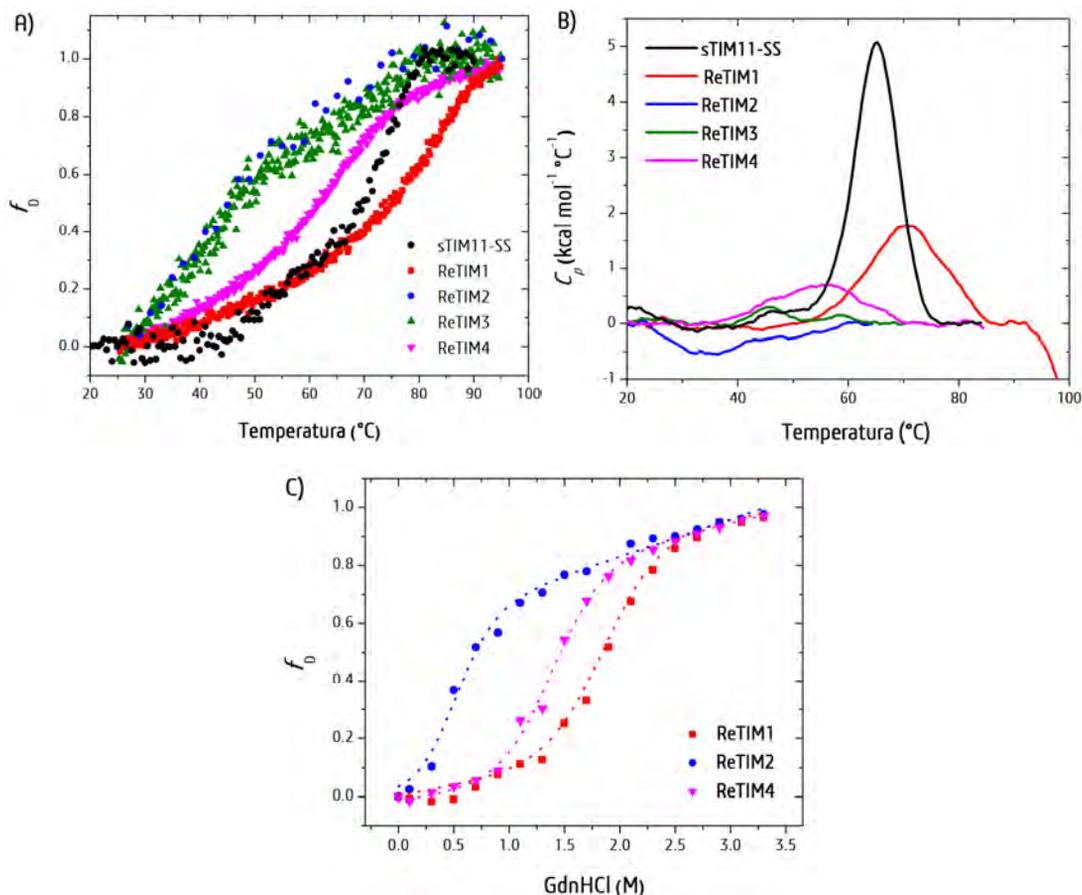


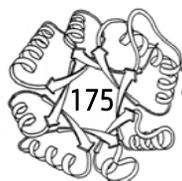
Figura 54. Caracterización fisicoquímica de las ReTIM1-4. **A)** Despliegamiento térmico seguido por DC. **B)** Despliegamiento térmico seguido por DSC. **C)** Despliegamiento químico usando GdnHCl y seguido por CD; la línea punteada representa el ajuste de los datos a un modelo de dos estados reversible ($N \rightleftharpoons D$).

Tabla 26. Parámetros termodinámicos de las ReTIM1-4.

Proteína	Despliegamiento térmico				Despliegamiento químico	
	T_m (°C)	ΔH (kcal mol ⁻¹)	ΔC_p (kcal mol ⁻¹ K ⁻¹)	Reversibilidad	ΔG 25°C (kcal mol ⁻¹)	C_m (M)
ReTIM1	70.6	22.9	NA	No	4.7	2.1
ReTIM2	41.2	ND	NA	No	1.6	0.7
ReTIM3	45.7	3.8	ND	Sí	ND	ND
ReTIM4	55.2	13.2	ND	Sí	3.6	1.6
sTIM11	79.8	93.7	2.13	Sí	4.2	2.1
sTIM11-SS	64.5	71.2	1.66	Sí	ND	ND

ND: parámetro no determinado hasta el momento.

NA: este parámetro no se puede determinar debido a que el despliegamiento térmico es irreversible.



Propiedades fisicoquímicas de los diseños del núcleo periférico A: ReTIM5-7

Estas proteínas están enfocadas en variantes para llenar una cavidad inferior que se encuentra en el barril. Al igual que su diseño original, estas macromoléculas tienen simetría de orden 4. Los modelos estructurales y la ubicación de las mutaciones se presentan en la figura 55.

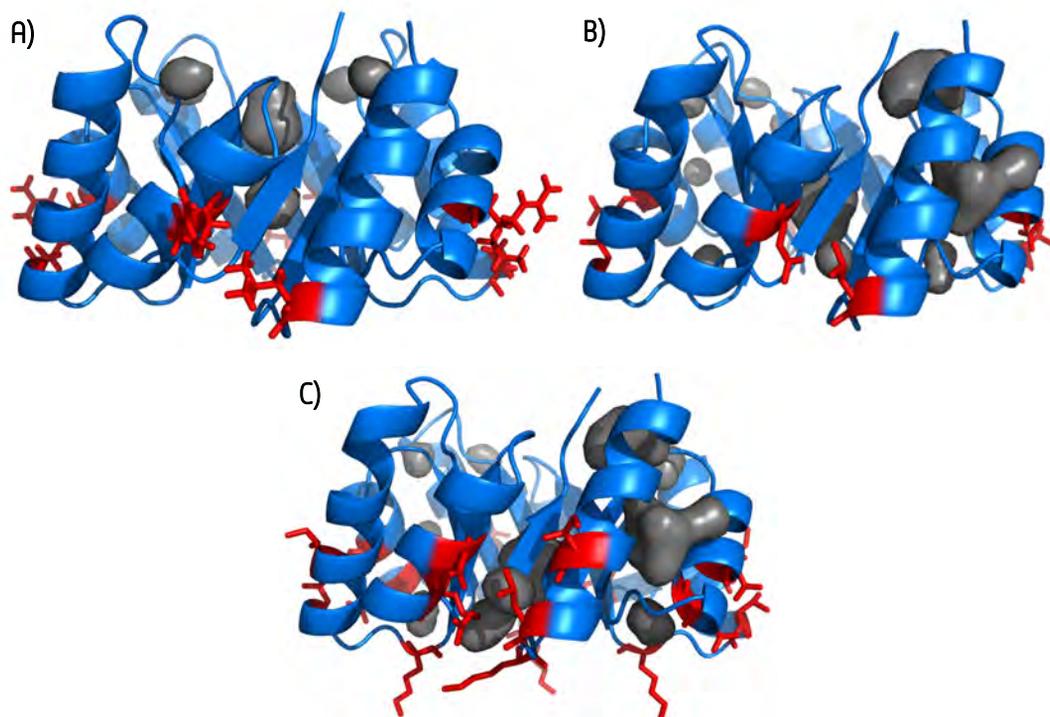


Figura 55. Modelos estructurales de las ReTIM5-7. A) ReTIM5, B) ReTIM6, C) ReTIM7. Las mutaciones se encuentran marcadas en rojo y se indican en la tabla 25.

Para ReTIM7 no fue posible subclonar el gen con el kit de Gibson Assembly, por lo que únicamente se pudo trabajar con ReTIM5 y ReTIM6. Para ReTIM5, los resultados de desnaturalización térmica no son alentadores ya que la cooperatividad es baja (figura 53a), el ΔH de la endoterma es pequeño (figura 56b y tabla 27) y, aunque la estabilidad ante un agente químico no pudo ser determinada (debido a la poca proteína sobreexpresada), no parece ser un buen candidato para seguir siendo estudiada por su baja termoestabilidad. Para el caso de ReTIM6 se obtuvieron

resultados completamente diferentes. Para esta proteína la termoestabilidad (T_m) se incrementó hasta 93.6 °C y, aunque el ΔH es menor al de sTIM11 y no existe reversibilidad en el desplegamiento térmico, es una buena candidata para seguir estudiándola, ya que es el diseño más exitoso de este grupo de proteínas. Por esta razón, esta proteína fue elegida como el diseño exitoso de este grupo y fue considerada para las mezclas de diseños. Debido a problemas de rendimiento en la purificación, no se pudo realizar la caracterización detallada del desplegamiento de estas proteínas, sin embargo, se espera hacer en un futuro próximo.

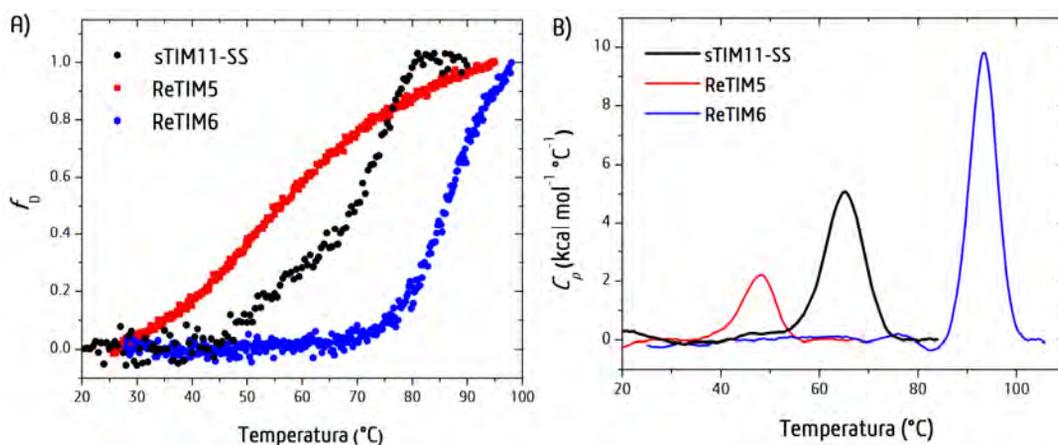


Figura 56. Caracterización fisicoquímica de la ReTIM5 y ReTIM6. A) Desplegamiento térmico seguido por DC. B) Desplegamiento térmico seguido por DSC.

Tabla 27. Parámetros termodinámicos de la ReTIM5 y ReTIM6.

Proteína	Desplegamiento térmico				Desplegamiento químico	
	T_m (°C)	ΔH (kcal mol ⁻¹)	ΔC_p (kcal mol ⁻¹ K ⁻¹)	Reversibilidad	ΔG 25°C (kcal mol ⁻¹)	C_m (M)
ReTIM5	48.3	16.2	NA	No	ND	ND
ReTIM6	93.2	64.7	NA	No	ND	ND
sTIM11	79.8	93.7	2.13	Sí	4.2	2.1
sTIM11-SS	64.5	71.2	1.66	Sí	ND	ND

ND: parámetro no determinado hasta el momento.

NA: este parámetro no se puede determinar debido a que el desplegamiento térmico es irreversible.



Propiedades fisicoquímicas de los diseños del núcleo periférico B: ReTIM8-10

Estas proteínas están enfocadas en variantes para rellenar una cavidad superior que se encuentra en el barril. Al igual que su diseño original, estas macromoléculas tienen simetría de orden 4. Los modelos estructurales y la ubicación de las mutaciones se presentan en la figura 57.

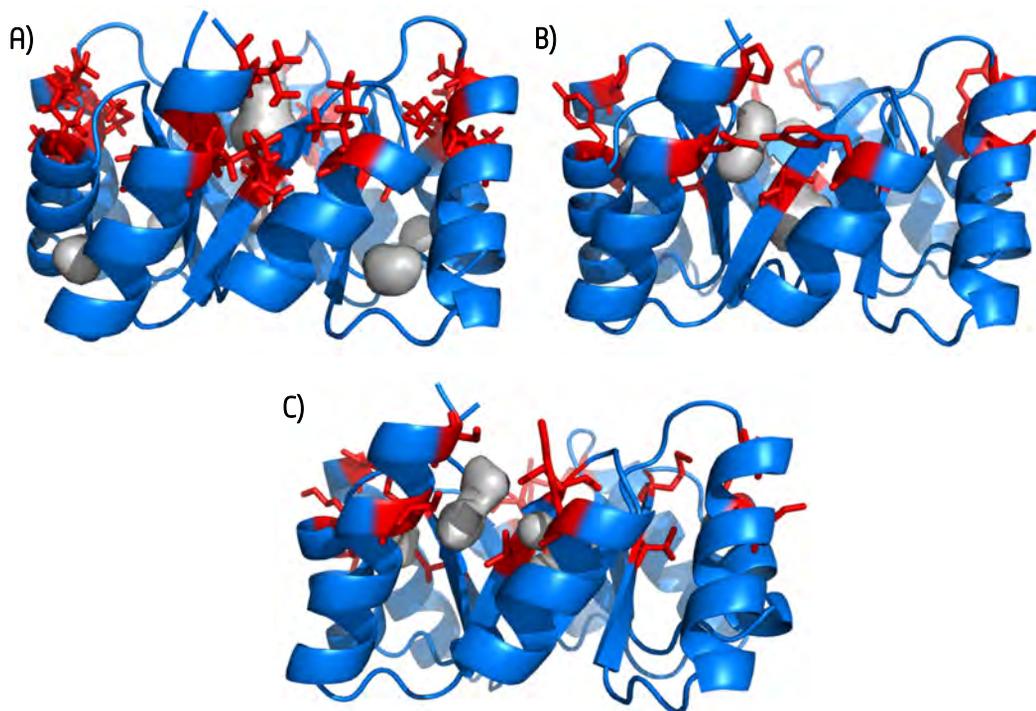


Figura 57. Modelos estructurales de las ReTIM8-10. A) ReTIM8, B) ReTIM9, C) ReTIM10. Las mutaciones se encuentran marcadas en rojo y se indican en la tabla 25.

Sobre la caracterización de las proteínas, cuando se observan sus curvas de desnaturalización seguidas por DC, se puede notar que se ven mucho mejor que las proteínas anteriores en términos de forma de la rampa (ya no se observa la pendiente en la pretransición) y de termoestabilidad (figura 58a); aunque si se comparan sus endotermas por calorimetría se encuentra que aún siguen teniendo una entalpía pequeña y el ancho de las mismas indica que la cooperatividad de la proteína está disminuida (figura 58b). De todas estas, ReTIM8 es el mejor candidato para seguir siendo caracterizado tanto por valor de T_m (10 °C más que sTIM11-SS), entalpía y reversibilidad parcial en

el desplegamiento térmico (tabla 28). Además, cuando se comparan los resultados de desnaturalización química (figura 58c), se puede observar que la estabilidad de ReTIM18 a 25 °C se incrementó ligeramente con respecto a sTIM11.

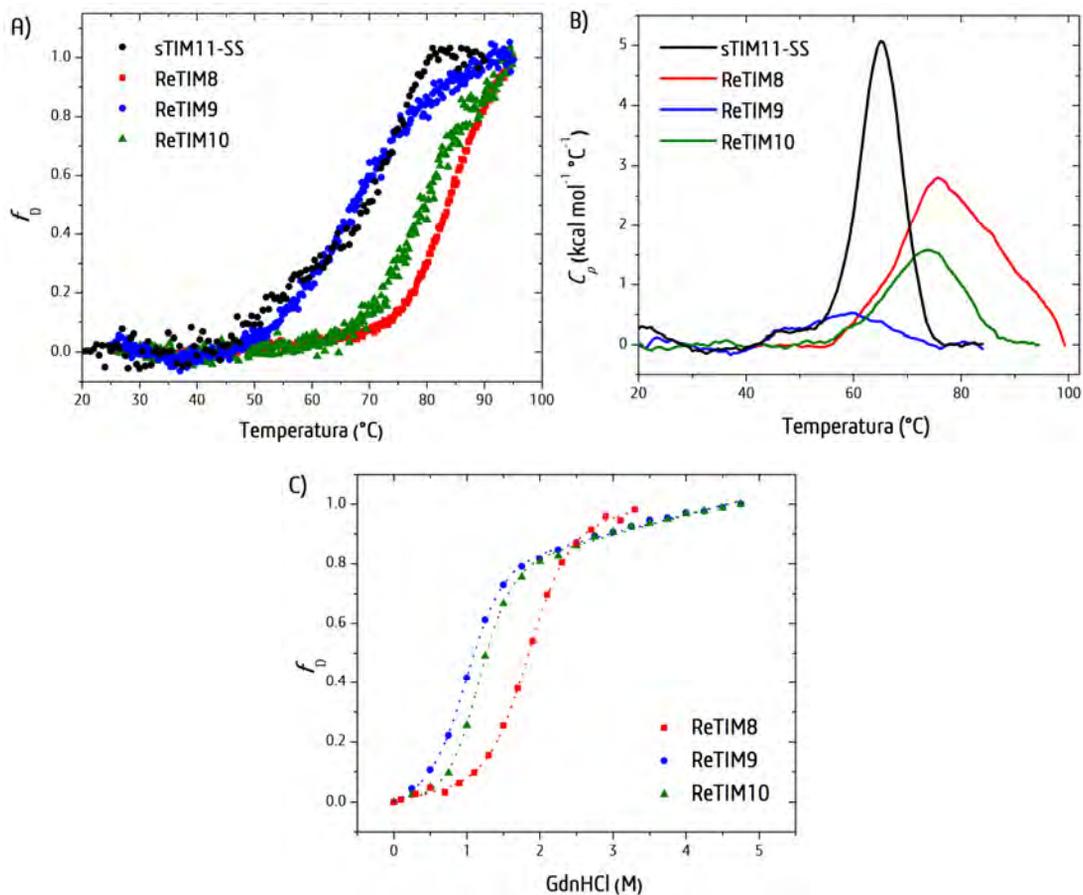


Figura 58. Caracterización fisicoquímica de las ReTIM8-10. A) Desplegamiento térmico seguido por DC. B) Desplegamiento térmico seguido por DSC. C) Desplegamiento químico usando GdnHCl y seguido por CD; la línea punteada representa el ajuste de los datos a un modelo de dos estados reversible ($N \rightleftharpoons D$).

Tabla 28. Parámetros termodinámicos de las ReTIM8-10.

Proteína	Desplegamiento térmico				Desplegamiento químico	
	T_m (°C)	ΔH (kcal mol ⁻¹)	ΔC_p (kcal mol ⁻¹ K ⁻¹)	Reversibilidad	ΔG 25°C (kcal mol ⁻¹)	C_m (M)
ReTIM8	75.7	65.2	ND	Parcialmente	4.3	2.2
ReTIM9	59.9	8.7	NA	No	1.6	1.2
ReTIM10	74.1	26.4	ND	Parcialmente	2.1	1.4
sTIM11	79.8	93.7	2.13	Sí	4.2	2.1
sTIM11-SS	64.5	71.2	1.66	Sí	ND	ND

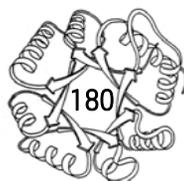
ND: parámetro no determinado hasta el momento.

NA: este parámetro no se puede determinar debido a que el despliegamiento térmico es irreversible.

Reversibilidad parcial se refiere a una recuperación de la endoterma entre 50-75%.

Propiedades fisicoquímicas de los diseños con mutaciones puntuales: ReTIM11-16

Este grupo de proteínas incluye diseños con mutaciones en posiciones específicas para probar distintas contribuciones de estos residuos en la estabilidad de la proteína. Los modelos estructurales con las mutaciones realizadas son los siguientes (figura 59).



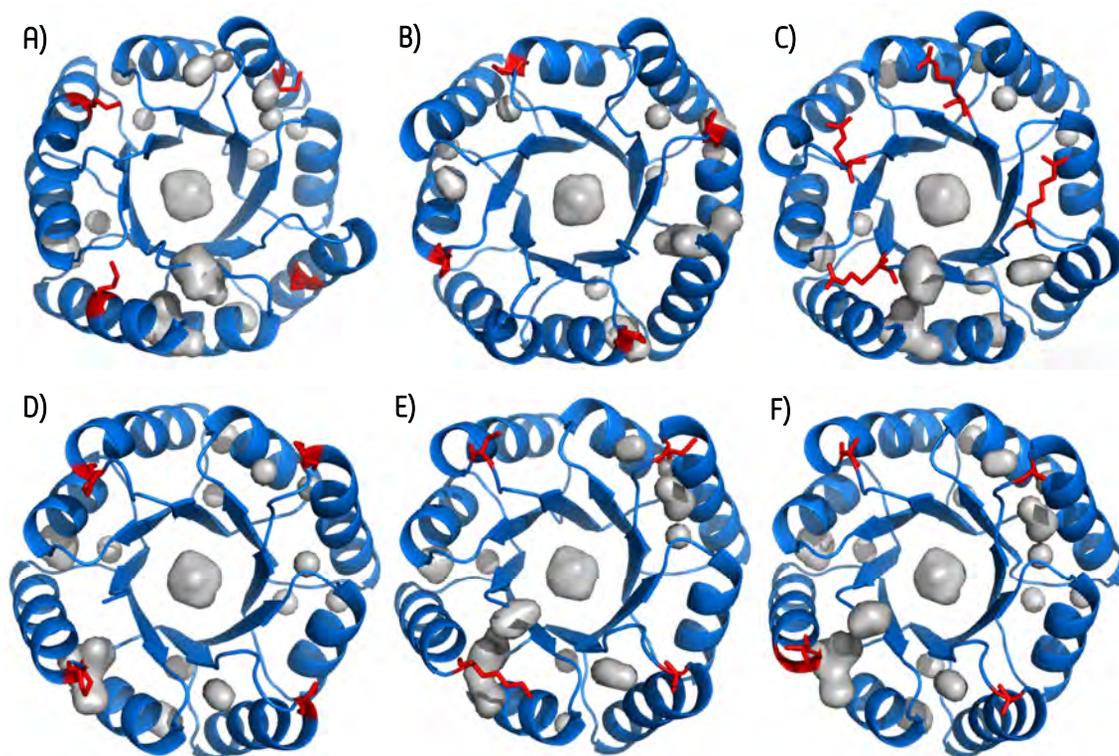


Figura 59. Modelos estructurales de las ReTIM11-16. A) ReTIM11, B) ReTIM12, C) ReTIM13, D) ReTIM14, E) ReTIM15, F) ReTIM16. Las mutaciones se encuentran marcadas en rojo y se indican en la tabla 25.

La caracterización de estas proteínas se limita a ReTIM13 y ReTIM15, ya que para el resto (ReTIM11-14) se tuvo problemas con el subclonado de los genes y no fue posible sobreexpresarlas. Para estas dos proteínas, no se observa un desplegamiento cooperativo y con entalpía comparable con alguno de los diseños previos (figura 60 y tabla 29); además, la proteína es tan poco estable que en poco tiempo (un par de días) se agrega irreversiblemente, por lo que no fue posible caracterizar el desplegamiento químico de ambas proteínas. Estas variantes tienen pocas mutaciones y cuando se analizan los cambios que se hicieron, podemos darnos cuenta de que mutar los triptófanos 27 y 34 en el barril tiene una consecuencia muy grande en la estabilidad de la proteína, por lo que estos residuos no deben ser tocados en el diseño o al menos no deben ser mutados a aminoácidos no polares como se hizo en estas proteínas. Debido a todo esto, estos dos candidatos fueron descartados para un análisis fisicoquímico posterior. Para los que hubo problemas de subclonado de genes, se

consideraron estos diseños para las mezclas que se incluyen en el último grupo de proteínas diseñadas.

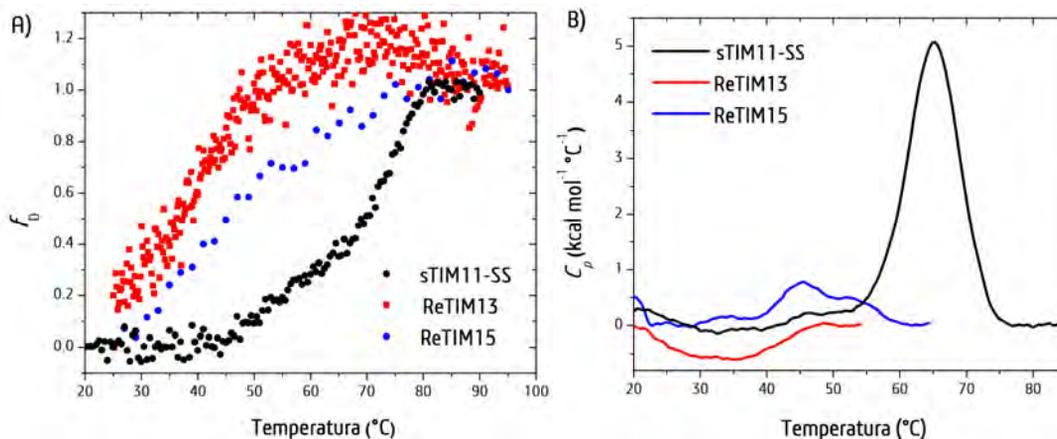


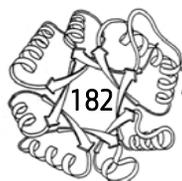
Figura 60. Caracterización fisicoquímica de la ReTIM13 y ReTIM15. A) Despliegamiento térmico seguido por DC. B) Despliegamiento térmico seguido por DSC.

Tabla 29. Parámetros termodinámicos de la ReTIM3 y ReTIM15.

Proteína	Despliegamiento térmico				Despliegamiento químico	
	T_m (°C)	ΔH (kcal mol ⁻¹)	ΔC_p (kcal mol ⁻¹ K ⁻¹)	Reversibilidad	ΔG 25°C (kcal mol ⁻¹)	C_m (M)
ReTIM13	40.1	ND	ND	No	ND	ND
ReTIM15	45.2	ND	ND	No	ND	ND
sTIM11	79.8	93.7	2.13	Sí	4.2	2.1
sTIM11-SS	64.5	71.2	1.66	Sí	ND	ND

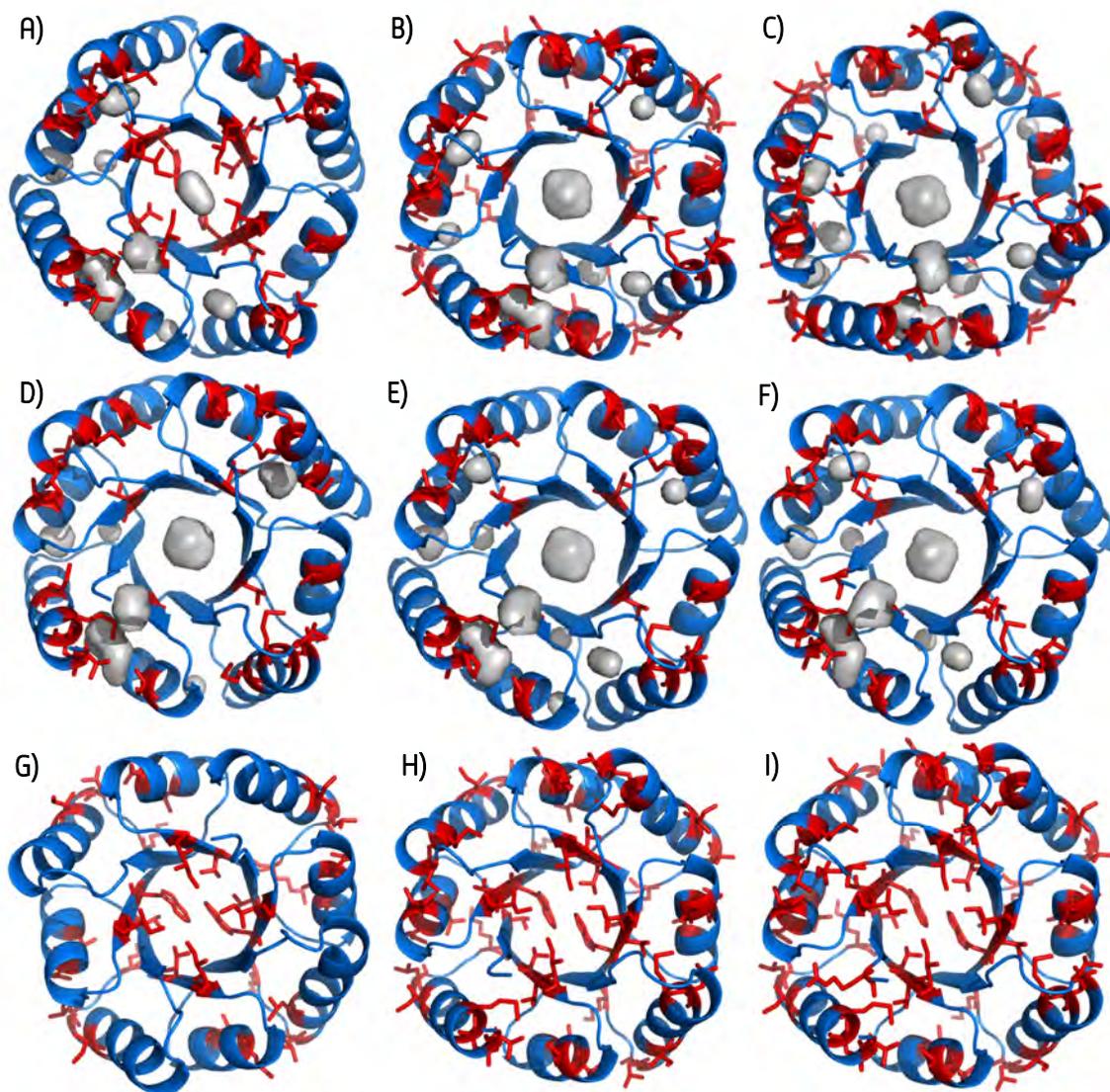
ND: parámetro no determinado hasta el momento.

NA: este parámetro no se puede determinar debido a que el despliegamiento térmico es irreversible.



Propiedades fisicoquímicas de los diseños quimeras: ReTIM17-28

En este último grupo de proteínas se incluyen quimeras de los diseños anteriores, todas teniendo como base la proteína ReTIM8, el mejor candidato de la primera ronda de diseños. La localización de las mutaciones en cada una de las quimeras se muestra en la figura 61.



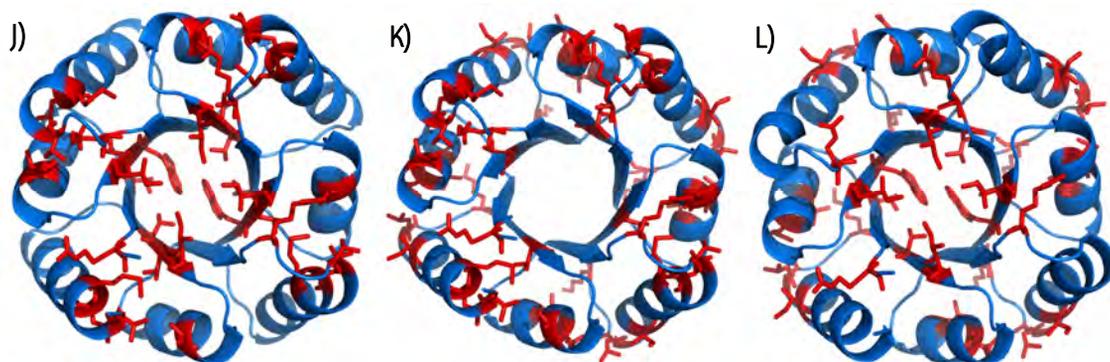


Figura 61. Modelos estructurales de las ReTIM17-28. A) ReTIM17, B) ReTIM18, C) ReTIM19, D) ReTIM20, E) ReTIM21, F) ReTIM22, G) ReTIM23, H) ReTIM24, I) ReTIM25, J) ReTIM26, K) ReTIM27, L) ReTIM28. Las mutaciones se encuentran marcadas en rojo y se indican en la tabla 25.

La caracterización fisicoquímica de este grupo de proteínas resultó interesante. Por un lado, en las rampas de temperatura seguidas por DC (figura 62a) se observa un comportamiento completamente distinto que la proteína “madre” (ReTIM0: sTIM11-SS). En varios de los diseños no se observa la pre-transición con una pendiente grande indicativa de que el estado nativo se desplegaba antes de la transición principal y/o tenía una cooperatividad baja. Además, en las proteínas ReTIM17, ReTIM18 y ReTIM24, se observa una transición altamente cooperativa y con valores de una T_m mayores a 90 °C (figura 62b y tabla 30), 40 grados mayor a la de la proteína madre sTIM11-SS (ReTIM0). Aunado a esto, el ΔH para su desplegamiento es de los más altos encontrados en todas las rondas de los diseños (tabla 30). Sobre este parámetro, el diseño de ReTIM24 (quimera de ReTIM17 y ReTIM18) resulta prometedor, ya que presenta la mayor T_m y, en nuestro conocimiento, de las más altas reportadas para una proteína; además, el valor de ΔH (114 kcal mol⁻¹) es casi exacto al valor esperado (128 kcal mol⁻¹) para una proteína de ese tamaño de acuerdo a las parametrizaciones de Robertson y Murphy, 1997. Por todas estas razones, los 3 mejores diseños de todo este proceso fueron ReTIM17, ReTIM18 y ReTIM24. Cuando se compara la estabilidad utilizando un agente químico desnaturizante (figura 62c), se puede observar que, aunque para varios de los diseños la estabilidad se mantiene igual o se ve ligeramente reducida, para ReTIM17 se incrementa la estabilidad en un 50% pasando de 4.2 kcal/mol a 6.5 kcal/mol con respecto a sTIM11.

Para los diseños ReTIM23-28 no se han obtenido los valores de estabilidad utilizando un agente químico, pero estos experimentos ya se encuentran en proceso; es interesante obtenerlos ya que los valores nos comprobarán si los 3 diseños mencionados fueron exitosos, logrando el objetivo de incrementar la estabilidad del barril TIM para utilizarlo en futuras aplicaciones de funcionalización molecular.

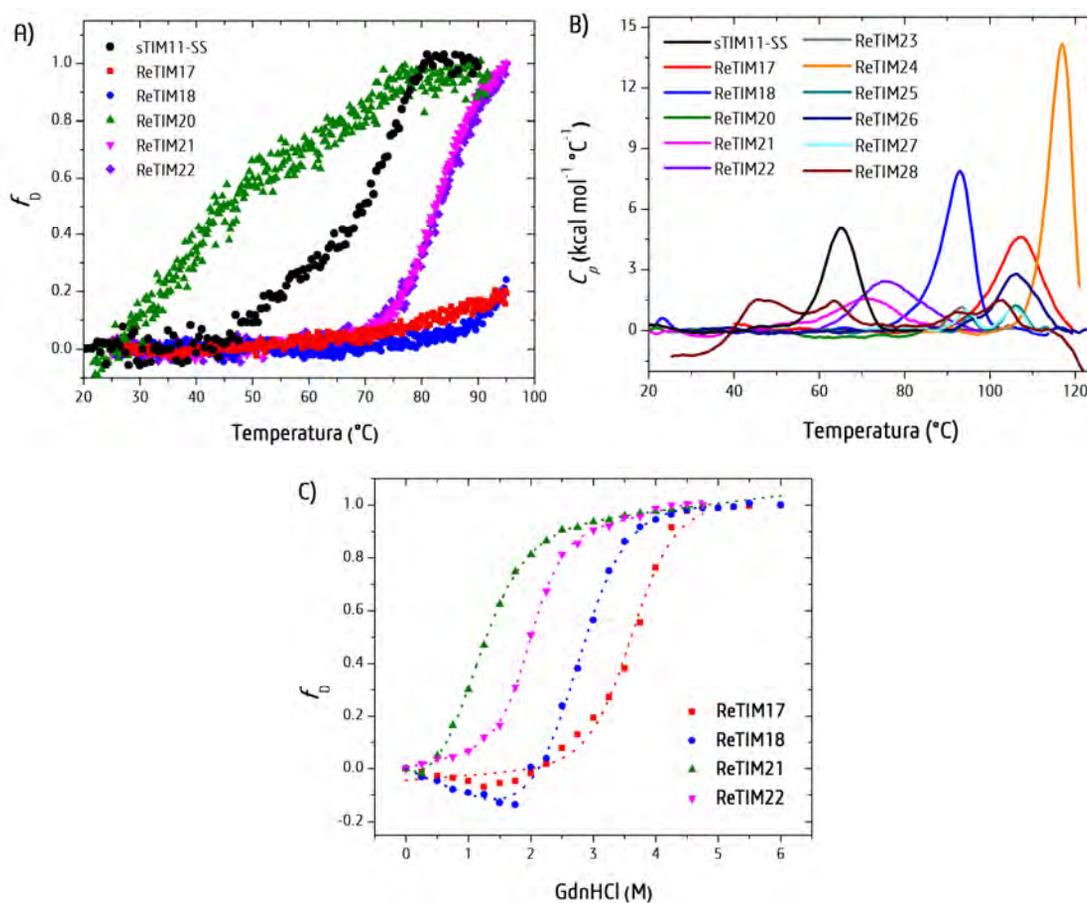


Figura 62. Caracterización fisicoquímica de las ReTIM17-28. A) Despliegamiento térmico seguido por DC. B) Despliegamiento térmico seguido por DSC. C) Despliegamiento químico usando GdnHCl y seguido por CD; la línea punteada representa el ajuste de los datos a un modelo de dos estados reversible ($N \rightleftharpoons D$).

Tabla 30. Parámetros termodinámicos de las ReTIM17-28.

Proteína	Desplegamiento térmico				Desplegamiento químico	
	T_m (°C)	ΔH (kcal mol ⁻¹)	ΔC_p (kcal mol ⁻¹ K ⁻¹)	Reversibilidad	ΔG 25°C (kcal mol ⁻¹)	C_m (M)
ReTIM17	107.2	80.4	0.95	Sí	6.5	3.8
ReTIM18	92.9	76.6	ND	Parcialmente	4.1	3.0
ReTIM20	49.5	ND	NA	No	ND	ND
ReTIM21	69.1	50.8	ND	Sí	2.1	1.3
ReTIM22	75.1	42.9	1.67	Sí	4.7	2.1
ReTIM23	93.9	63.1	NA	No	ND	ND
ReTIM24	117.3	114.3	ND	Sí	ND	ND
ReTIM25	106.1	30.93	NA	No	ND	ND
ReTIM26	106.2	32.7	ND	Sí	ND	ND
ReTIM27	105.8	30.9	NA	No	ND	ND
ReTIM28	50.2	ND	NA	No	ND	ND
sTIM11	79.8	93.7	2.13	Sí	4.2	2.1
sTIM11-SS	64.5	71.2	1.66	Sí	ND	ND

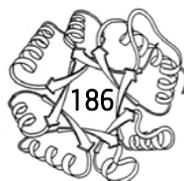
ND: parámetro no determinado hasta el momento.

NA: este parámetro no se puede determinar debido a que el despliegamiento térmico es irreversible.

Reversibilidad parcial se refiere a una recuperación de la endoterma entre 50-75%.

Descripción fisicoquímica de la estabilidad de las ReTIMs: la curva de estabilidad

Se determinó la curva de estabilidad para los diseños que presentaron reversibilidad en el despliegamiento térmico y cuya estabilidad (ΔG a 25 °C) fuera mayor o igual a la de la proteína sTIM11 (no se tiene comparación de los valores de ΔG a 25 °C para sTIM11-SS porque no se han realizado estos experimentos). Bajo estos criterios, las proteínas seleccionadas fueron: ReTIM17, ReTIM22, ReTIM24, sTIM11 y sTIM11-SS (las dos últimas con fines de comparación). Para todas ellas se realizaron experimentos de DSC a varias concentraciones de proteína y los datos experimentales se ajustaron a un modelo de dos estados simple ($N \rightleftharpoons D$) (los experimentos no se realizaron para ReTIM24 debido a problemas metodológicos con el calorímetro, pero se realizarán próximamente para completar el trabajo). Del ajuste de los datos se obtuvieron los parámetros



fisicoquímicos T_m , ΔC_p y ΔH (tabla 31). Vale resaltar que para todos los diseños, tanto el ΔC_p como el ΔH son más bajos que los estimados para una proteína de ese tamaño (184 residuos: $\Delta C_p = 2.6$ kcal mol⁻¹ y $\Delta H = 128.4$ kcal mol⁻¹); sin embargo, para ReTIM24 se tiene calculada una entalpía mayor, por lo que resulta interesante saber si este diseño cumple ya con todos los parámetros “ideales” para una proteína de esa cantidad de residuos; si es así, habremos avanzado en el objetivo y podríamos aventurarnos a funcionalizar la proteína diseñada.

Usando la ecuación correspondiente y los parámetros termodinámicos determinados (tabla 31), las curvas de estabilidad, es decir, la dependencia del ΔG con la temperatura para las ReTIMs y sTIMs fueron calculadas (figura 63). A partir de estas curvas, se obtuvieron los valores de la temperatura de estabilidad máxima, el valor de ΔG a esta temperatura y el valor de la T_m en frío, valores que se muestran en la tabla 31.

Como se puede observar en la figura 63, las curvas de estabilidad de estas proteínas son distintas entre sí, a pesar de que todos ellos son barriles TIM diseñados con el esqueleto fijo (el de sTIM11-SS). Se puede ver claramente que las proteínas presentan diferente dependencia del ΔG con respecto a la temperatura. La curva de estabilidad puede ser modulada mediante cambios en la T_s (temperatura donde $\Delta S = 0$), el ΔH_{T_s} (valor de entalpía en la T_s), y/o el ΔC_p ; en el caso de los barriles TIM diseñados, la forma de la curva de estabilidad es tan distinta entre ellas porque los tres parámetros cambian importantemente en las distintas ReTIMs; caso contrario ocurre con los barriles TIM naturales, donde sólo los dos primeros parámetros termodinámicos cambian y el ΔC_p permanece casi constante.

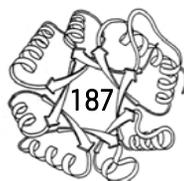
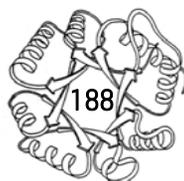


Tabla 31. Parámetros termodinámicos para el desplegamiento de ReTIM17, ReTIM22, sTIM11 y sTIM11-SS.

Parámetro fisicoquímico	ReTIM17	ReTIM22	sTIM11	sTIM11-SS
T_m en calor (°C)	107.3 ± 0.2	74.2 ± 0.2	79.8 ± 0.3	64.5 ± 0.3
T_m en frío (°C)	-32.4	-11.2	-1.2	-14.3
ΔH (kcal mol ⁻¹)	80.4 ± 0.7	55.9 ± 1.5	93.7 ± 0.9	71.2 ± 4.5
ΔC_p (kcal mol ⁻¹ K ⁻¹)	1.1 ± 0.03	1.2 ± 0.11	2.2 ± 0.03	1.5 ± 0.09
$\Delta G_{25^\circ C}$ (kcal mol ⁻¹) (químico)	6.5 ± 0.6	4.7 ± 0.2	4.2	ND
$\Delta G_{25^\circ C}$ (kcal mol ⁻¹) (temperatura)	6.9	3.5	4.6	4.7
ΔG de máxima estabilidad (kcal mol ⁻¹)	7.9	3.6	5.6	4.4
Temperatura de máxima estabilidad (°C)	39.3	29.3	37.6	23.3



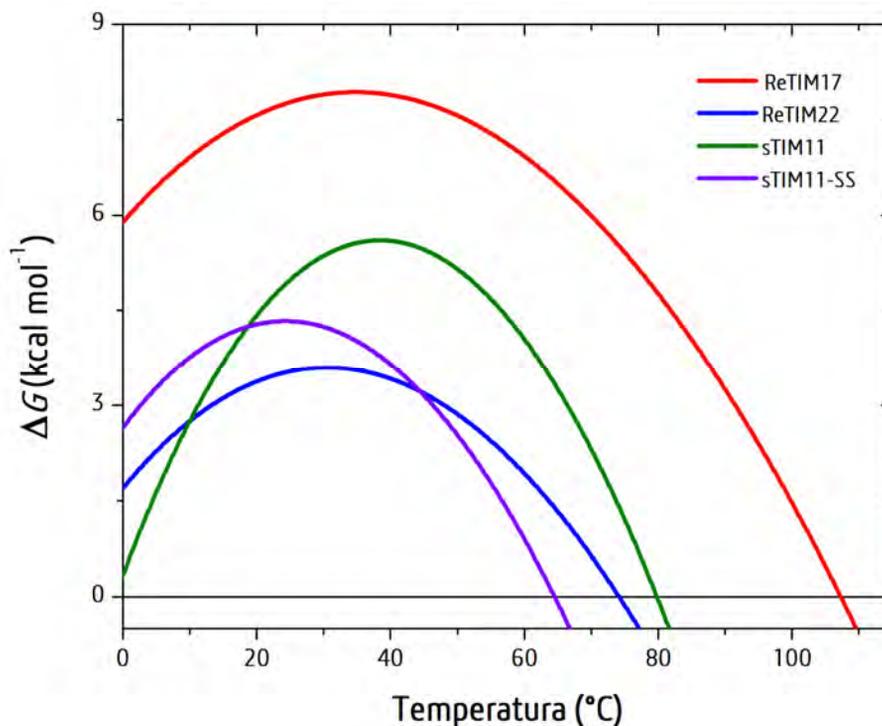
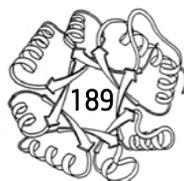


Figura 63. Curvas de estabilidad de barriles TIM diseñados *de novo*: ReTIM17, ReTIM22, sTIM11 y sTIM11-SS. Las líneas sólidas son las curvas de estabilidad construidas con los datos obtenidos de los experimentos de DSC y la ecuación de Gibbs-Helmholtz (ecuación 19). Las curvas de estabilidad de los barriles TIM diseñados *de novo* se comparan con la curva de estabilidad de un barril TIM natural ($C\beta$ TIM).

Estructura tridimensional de las ReTIMs

Con la finalidad de establecer una relación energético-estructural entre los parámetros termodinámicos observados en las ReTIMs y los componentes estructurales diseñados en los barriles TIM, se intentó determinar la estructura tridimensional de los diseños más exitosos de cada grupo: ReTIM1, ReTIM6, ReTIM8, ReTIM17, ReTIM18 y ReTIM24. De todas las condiciones exploradas, hasta el momento hemos obtenido una condición exitosa de cristalización que permitió obtener datos de difracción de alta calidad para la ReTIM18. Las condiciones de cristalización para esta proteína fueron tartrato de Sodio-Potasio 0.2 M, Polietilenglicol 3350 20% w/v y la proteína cristalizó a una concentración de 10 mg mL⁻¹.



El cristal se difractó en el sincrotrón de Suiza (Swiss Lightsource SLS) y se resolvió la estructura como se explica en la estrategia metodológica. En la tabla 32 se muestran las estadísticas cristalográficas. Como se puede observar, la estructura determinada alta calidad y las estadísticas son confiables para ésta, por lo que está lista para depositarse en el PDB cuando sea necesario.

Los datos cristalográficos muestran un estado de asociación monomérico, en concordancia con lo esperado para toda esta familia de barriles TIM diseñados. La topología canónica de un barril $(\beta/\alpha)_8$ es idéntica en el barril diseñado (figura 64a) y, si se compara con la proteína sTIM11-SS, su RMSD es de 1.4 Å, lo cual nos indica de que sí hubo ciertos cambios locales entre ambas proteínas (figura 64b), tal como se esperaba ya que sus propiedades fisicoquímicas son diferentes. En cambio, cuando la proteína se compara con el modelo que arroja Rosetta tras el diseño, podemos ver que el RMSD es más bajo (0.87 Å; figura 64c), lo que confirma que el programa para algunos sistemas es capaz de diseñar con una precisión casi a nivel atómico, tal cual se había visto para otras proteínas previamente diseñadas con esta herramienta (Kuhlman *et al.*, 2003; Huang *et al.*, 2014; Joh *et al.*, 2014; Koga *et al.*, 2012; Smadbeck *et al.*, 2014; Bellows-Paterson *et al.*, 2012; Khoury *et al.*, 2014; Correia *et al.*, 2014; Huang *et al.*, 2016). En los casos en que esto no sucede, es decir, cuando el diseño no se parece a la estructura real que adquiere la proteína, consideramos que la razón recae en que básicamente aún no comprendemos a profundidad las reglas o principios que fundamentan la estructura-función-estabilidad de las proteínas. Sin embargo, trabajos como estos aportan conocimiento al campo para mejorar la tasa de éxito de las proteínas diseñadas.

Buscamos determinar la estructura tridimensional de las otras ReTIMs (1, 6, 8, 17 y 24), por lo que actualmente ya se tienen pruebas de cristalización preliminares para ellas como se puede ver en la figura 65. Estamos en la etapa de optimizar tales condiciones para obtener cristales con buena calidad de difracción. Se espera que pronto se puedan obtener las otras estructuras para así realizar una completa comparación energético-estructural entre los barriles TIM diseñados *de novo*.

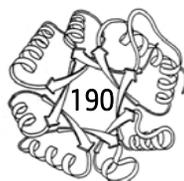


Tabla 32. Estadísticas tras la colecta, integración, reducción, escalamiento y afinamiento de los datos para ReTIM18.

Colecta de datos ^a	
PDB ID	Por asignar
Límites de resolución (Å)	44.79 – 1.64 (1.74-1.64)
Grupo espacial	P3(2) (145)
Dimensiones de la celda unitaria	
a, b, c, (Å)	51.72, 51.72, 63.97
α , β , γ , (°)	90.0, 90.0, 120.0
Contenido en la unidad asimétrica	Monómero
Reflexiones totales	79664
Reflexiones únicas	23407 (2279)
Multiplidad	3.4
Integridad (%)	99.5 (98.5)
I / sigma (mean)	14.27 (1.74)
R-merge (%)	5.1 (58.2)
CC1/2 (%)	98.3 (74.7)
Estadísticas del afinamiento	
R _{work} / R _{free} (%)	18.5 / 21.0
Factor B promedio (Å ²)	33.8
Proteína	32.5
Ligando (calcio y crioprotector: glicerol)	45.6
Solvente	46.0
Número de átomos	3242
Proteína	1464
Ligando	7
Solvente	150
Residuos de proteína	181
RMS (enlaces) (Å)	0.006
RMS (ángulos) (°)	0.8
Ramachandran favorecidos (%)	99.44
Ramachandran permitidos (%)	0.56
Ramachandran outliers (%)	0.00
Clashscore	3.95

^a Entre paréntesis se muestran las estadísticas para la capa de más alta resolución.



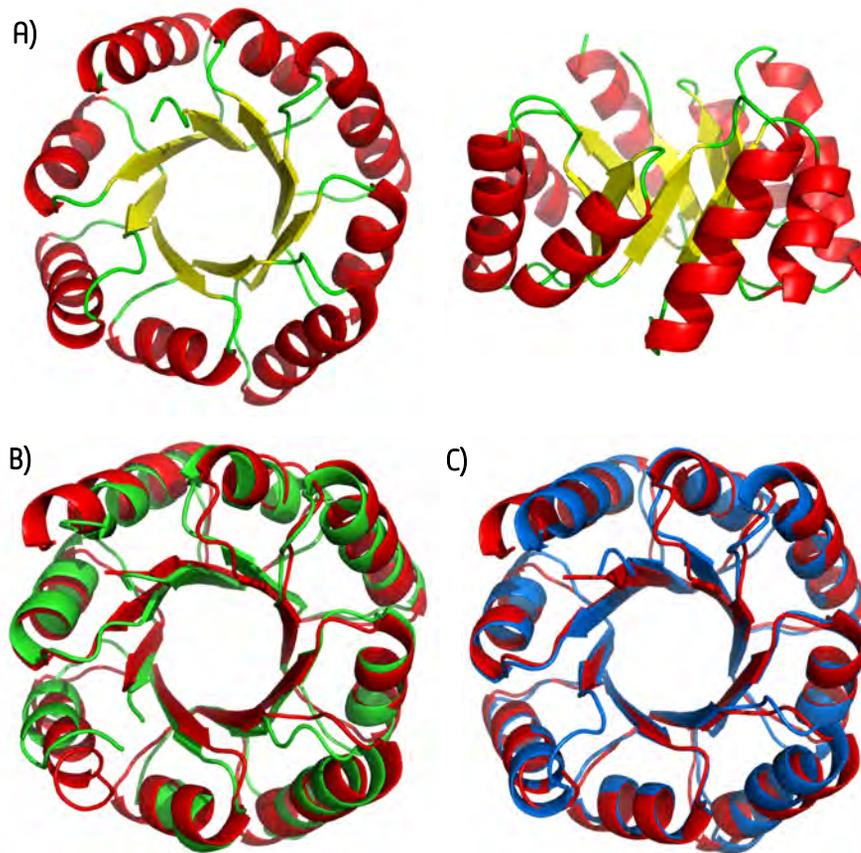


Figura 64. Estructura tridimensional de ReTIM18. A) Vista superior (izquierda) y lateral (derecha) del barril TIM diseñado; los elementos de estructura secundaria están presentados con colores diferentes. B) Alineamiento estructural de ReTIM18 con sTIM11-SS (RMSD: 1.4 Å). C) Alineamiento estructural de ReTIM18 con el modelo arrojado por Rosetta tras el diseño inicial (RMSD: 0.87 Å).

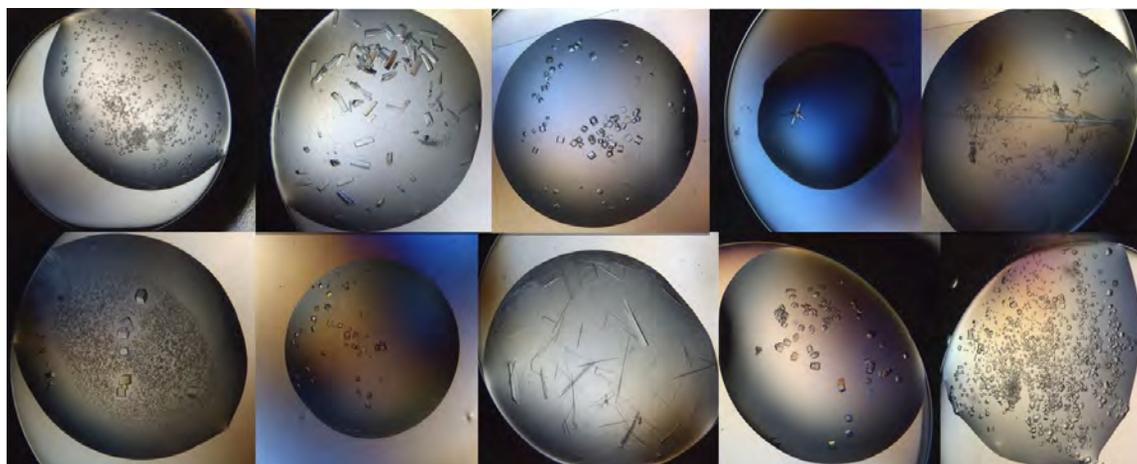
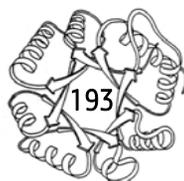


Figura 65. Pruebas de cristalización de las ReTIMs. Se muestran algunas gotas donde se han obtenido estructuras cristalinas para las proteínas ensayadas (ReTIMs 1, 6, 8, 17 y 24).

Discusión final sobre el plegamiento de los barriles TIM diseñados de novo

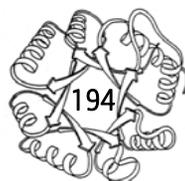
Los resultados presentados muestran claramente varios aspectos que es necesario resaltar. Por una parte, al igual que en los barriles TIM naturales se observan dos familias de proteínas: las que se despliegan térmicamente de forma irreversible y las que se pliegan reversiblemente. Es interesante señalar que para algunos casos existen diseños que se repliegan parcialmente, como ReTIM8, y que al mezclarse con otro conjunto de mutaciones provoca que la proteína resultante se despliegue/repliegue reversiblemente. Si bien desconocemos la causa de este comportamiento, resulta interesante analizarlo a profundidad en un futuro. Una gran ventaja del diseño por regiones, es que permite explorar la posibilidad de estabilizar la proteína de manera sistemática, esto es, creando diseños que incorporen en una misma secuencia mutaciones que probaron ser estabilizantes por separado.

Cabe resaltar que un diseño racional de proteínas, como el que se realizó, permite saber específicamente qué regiones promueven una mayor estabilidad en las proteínas diseñadas y cuáles



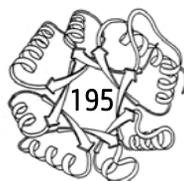
tienen efectos negativos. Igualmente, la propiedad de los diseños de ser simétricos permite que el efecto de las mutaciones también sea aditivo, porque están presentes de forma repetitiva. Al respecto, los diseños que presentan mutaciones en la cavidad superior del barril (ReTIM8-10 y sus quimeras) al parecer son los más efectivos para estabilizar a sTIM11-SS, debido a que creemos que las mutaciones tienen un efecto considerable en cómo se conforman estructuralmente las asas superiores del barril, ya que justamente esas regiones son las más desordenadas y móviles en sTIM11-SS, cuestión que ya no se observa en la estructura de ReTIM18. Debido a esto, posiblemente las mutaciones en la parte superior del barril son las que más contribuyen a estabilizar la proteína por su efecto en cómo se estructuran las asas que conectan las distintas regiones del barril.

Conjuntamente, algo importante que se debe analizar en un futuro es si los diseños “no exitosos” nos pueden dar información de qué regiones no deben modificarse en la proteína o qué mutaciones no deben realizarse. Hasta el momento, aunque no es clara una tendencia que divida a los diseños exitosos de los que no lo fueron, sí se puede observar que los diseños exitosos de la primera ronda realizada, es decir, ReTIM1, ReTIM6 y ReTIM8, presentan mayor número de mutaciones a aminoácidos no polares, a diferencia de los diseños más inestables o no cooperativos en los cuales no se cambió la naturaleza fisicoquímica del aminoácido. Al parecer, el cambiar la polaridad del residuo a uno no polar da mayores beneficios que mutarlo por otro de la misma naturaleza fisicoquímica. La razón de esto es porque dichas mutaciones pueden estar incrementando el empaquetamiento de la proteína y esto tiene un efecto directo en los parámetros fisicoquímicos asociados como la entalpía o el cambio en la capacidad calorífica. De igual forma, con los diseños fallidos como ReTIM13, 15 y 16 se puede observar que mutar los triptófanos 27 y 34 tiene un efecto negativo en la estabilidad de la proteína, por lo que esos residuos no deben ser tocados durante las rondas de diseños posteriores. Vale hacer notar que esas dos mutaciones están en algunos otros diseños, pero ahí no se ve un efecto deletéreo considerable tal vez porque la mezcla con otras mutaciones ayuda a compensar las consecuencias de esa mutación.

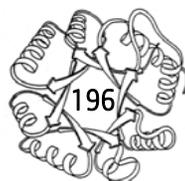


En la siguiente sección se analizará lo referente a las curvas de estabilidad, ya que estas son relevantes cuando se comparan con barriles TIM naturales. Con el conocimiento que hemos obtenido fuimos capaces de modificar la curva de estabilidad de estos diseños, obteniendo 3 diseños prometedores para utilizarse como andamiaje para los estudios de funcionalización de los barriles TIM diseñados. Como plan para funcionalizar los barriles, se planean tres vertientes principales. Primero, se desea generar un sitio de unión a fosfato ya que casi todos los barriles TIM naturales tienen uno, por lo que sería interesante imitar esos sitios de unión pero con proteínas nuevas. Segundo, también se quiere diseñar un sitio de unión a ATP; esto tiene dos razones de fondo. Por un lado, de las 6 clases de actividades enzimáticas que están descritas, los barriles TIM catalizan 5 de ellas, la única que no se ha descrito en un barril TIM es la actividad de ligasa. Además, se sabe que todas las ligasas utilizan ATP, por lo que un primer paso para que un barril diseñado tenga actividad de ligasa es que éste una ATP. Asimismo, no existe un barril TIM natural que una a esta molécula, por lo cual es un reto novedoso en el mundo del diseño de proteínas. Finalmente, también se planea a futuro diseñar una enzima con alguno de los barriles, una posibilidad es colocarle actividad de Triosafosfato Isomerasa, ya que conocemos bien el sistema y como todos los barriles diseñados son monoméricos, sería novedoso tener un barril con actividad TIM pero sin tener que ser un oligómero obligado como lo son todas las Triosafostato Isomerasas naturales.

También es importante señalar que otra posibilidad en este ámbito del diseño de proteínas es haber primero diseñado la función deseada y, posteriormente, estabilizar la proteína. Sin embargo, consideramos que esta estrategia presenta dos problemas: primero, la estabilidad del barril madre (sTIM11-SS) es muy baja, lo cual reduciría considerablemente la tasa de éxito en el diseño; segundo, todo el proceso únicamente serviría para el diseño de esa actividad en específico, mientras que en la estrategia que seguimos, el barril estabilizado puede ser utilizado para diseñar diversas cosas, como sitios de unión a un ligando, actividad enzimática, etc.



A continuación, se dará una discusión general sobre las propiedades fisicoquímicas de los paisajes conformacionales de los barriles TIM tanto naturales como diseñados y sobre cómo estas propiedades que hemos determinado nos pueden ayudar a modular la estabilidad y el plegamiento de las proteínas, ya sea en mecanismos de diseño racional , en el mejoramiento de las características de las proteínas que existen actualmente en la naturaleza (como enzimas con propiedades biotecnológicas interesantes o fármacos proteicos que se distribuyen comercialmente) o en el entendimiento de cómo estas propiedades han evolucionado hasta lo que son hoy en día.

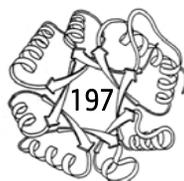


Parte 3: Lo que hemos aprendido sobre el paisaje conformacional de los barriles TIM. Una comparación entre barriles TIM naturales y diseñados *de novo*

El universo proteico es de gran complejidad. Muchas proteínas poseen una elevada masa molecular, forman complejos con otras biomoléculas, sufren cambios al llevar a cabo su actividad, poseen estabilidades distintas, etc. En casi todas de ellas, la adquisición de su estructura nativa es de vital importancia. Uno de los retos más importantes a los que se enfrenta la bioquímica actual consiste en discernir de qué modo una cadena polipeptídica recién sintetizada adquiere su estructura tridimensional y funcional y cómo la estabilidad de esta estructura nativa ha sido modulada por la evolución.

El estudio del plegamiento de proteínas ha evolucionado durante los últimos 50 años para convertirse en un campo de investigación maduro (Dill y MacCallum, 2012). Aunque el plegamiento y la estabilidad de diversas proteínas se ha estudiado a resolución atómica y se han diseñado diversas proteínas (Feng *et al.*, 2005; Koga *et al.*, 2012), nuestra comprensión actual de la relación secuencia \rightleftharpoons estructura \rightleftharpoons estabilidad de las proteínas está limitada por la información disponible en estos diferentes niveles. El número actual de secuencias de proteínas depositadas en el GenBank ($\approx 10^8$) es varios órdenes de magnitud mayor que el número de estructuras depositadas en el PDB ($\approx 10^5$). Del mismo modo, el número de proteínas de estructuras tridimensionales en el PDB es mucho mayor que su número en la Base de Datos de Estabilidad de Proteínas ($\approx 10^3$).

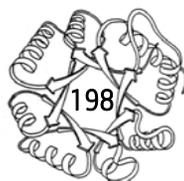
La comprensión de los fundamentos energéticos que subyacen la formación de ciertas topologías, como lo es el barril TIM, no sólo ayuda a la construcción de nuevas proteínas sino que también al mejoramiento de la predicción de la estructura terciaria a partir de la secuencia aminoacídica. Actualmente los métodos informáticos no son precisos y adecuados para el cálculo de la energía asociada a los procesos de plegamiento y estabilidad de las proteínas; sin embargo, poco a poco el conocimiento acumulado ha posibilitado la realización de múltiples trabajos de diseño e



ingeniería de proteínas. Es justamente el conocimiento acumulado y la realización de estudios comparativos permite mejorar los cálculos termodinámicos que se hacen al respecto, ya que el diseño racional de proteínas requiere un conocimiento cuantitativo de las contribuciones e interacciones que estabilizan a estas macromoléculas.

Para entender los cambios en la estabilidad de una familia de proteínas, básicamente se siguen tres aproximaciones: 1) comparar proteínas con distintas topologías para encontrar propiedades comunes o diferentes entre éstas que aumentan la estabilidad, 2) comparar una misma proteína pero proveniente de organismos diferentes, los cuales se enfocan en comparar los cambios en las proteína hechas a lo largo de la evolución y bajo la presión de selección distinta que sufre cada organismo y 3) utilizar mutagénesis dirigida en una proteína particular para ver los efectos de los cambios de residuos en posiciones específicas sobre la estabilidad global de la proteína. Sea cual sea el camino elegido, el procedimiento exige analizar variaciones en la proteína y determinar los cambios que suceden en su estabilidad, esto mediante experimentos de desplegamiento con temperatura, agentes químicos, presión o pH. El cruce de estas aproximaciones entre proteínas naturales y proteínas diseñadas es que justamente los cálculos hechos en proteínas naturales nos sirven para predecir las contribuciones de interacciones específicas que podrían suceder en los diseños realizados *in silico*, posteriormente se comprueban estas predicciones al trabajar las proteínas experimentalmente.

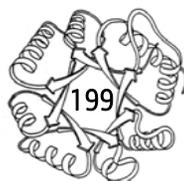
Uno de los primeros trabajos donde se comparan las curvas de estabilidad de distintas proteínas (en este caso, con diferente topología y tamaño, como lo son la lisozima, RNAsa, mioglobina, quimiotripsina y citocromo C) fue el realizado por Privalov y colaboradores en 1974. En el estudio se pueden notar varias propiedades de las curvas: 1) la diferencia de las energías libres molares no difiere mucho aunque son proteínas muy diferentes, 2) la diferencia en estabilidad es de sólo unas 5 kcal mol⁻¹, 3) la diferencia de ΔG alcanza valores máximos a temperaturas donde las diferencias de entalpía y entropía son muy pequeñas (Privalov y Khechinashvili, 1974). Con todo lo



anterior se puede concluir que a pesar de presentar topologías diferentes, en algunas proteínas los parámetros termodinámicos se han mantenido casi idénticos a pesar de presiones selectivas diferentes en el ambiente de los organismos y/o funciones moleculares de las proteínas.

Años más tarde, Nojima y colaboradores en 1977 realizaron el primer estudio comparativo entre una misma proteína pero procedente de organismos mesófilos y termófilos, la fosfoglicerato cinasa. Comparando distintas fosfoglicerato cinasas, ellos proponen tres mecanismos de termoestabilidad, los cuales siguen vigentes. El primero involucra únicamente un aumento en el cambio de entalpía, lo que hace una curva más alta y, por tanto, con una estabilidad máxima y T_m mayores. El segundo mecanismo implica un aumento en el núcleo hidrofóbico de la proteína, lo que tiene efecto directo en el ΔC_p y por tanto un aumento en la T_m , aunque no en la estabilidad máxima de la proteína. Por último, el tercer mecanismo propuesto es únicamente un incremento en la T_m , lo que hace a una proteína más termoestable pero con la misma estabilidad máxima, aunque desplazada a temperaturas mayores (Nojima *et al.*, 1977). Estos mecanismos se han visto involucrados en numerosas proteínas estudiadas a lo largo de varios años.

El efecto de mutaciones puntuales en la estabilidad de una proteína específica ha sido estudiado en algunos trabajos. Un ejemplo representativo es el estudio a profundidad que se ha hecho de la nucleasa de *Staphylococcus aureus* por el grupo de Stites desde 1985. Ellos han mutado aproximadamente 58 residuos de la proteína en diversas publicaciones y han encontrado que el efecto principal de todas mutaciones radica en los cambios de ΔC_p , los cuales tienen efectos completamente distintos en el máximo de estabilidad y la T_m de las proteínas mutantes. Ellos han logrado modular la estabilidad de la proteína únicamente cambiando el núcleo hidrofóbico de ésta y, por lo tanto, cambiando las curvas de estabilidad de acuerdo a cada mutación que se realiza (Shortle *et al.*, 1990; Green *et al.*, 1992; Byrne *et al.*, 1995; Stites *et al.*, 1995; Schwehm y Stites, 1998; Chen y Stites, 2001a-c; Holder *et al.*, 2001; Byrne y Stites, 2007; Talla-Singh y Stites, 2008).

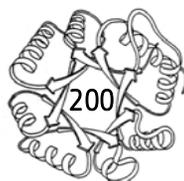


Trabajos como éste demuestran que cambios en un parámetro pueden ser suficientes para tener un efecto grande en la estabilidad global de las proteínas.

Pasando a trabajos más recientes, un estudio comparativo para entender la evolución de la estabilidad y los parámetros termodinámicos de las proteínas es el trabajo de Hart en 2014. En este trabajo se utiliza un método comparativo filogenético en el estudio de la termoestabilidad de la ribonucleasa H1. Los autores señalan que si bien existe una correlación estadística entre el incremento en la T_m y los ambientes de alta temperatura, los parámetros fisicoquímicos que determinan la T_m (como lo son el ΔC_p , el ΔH y la T_s) no correlacionan estadísticamente con el ambiente de vida para organismos mesófilos o termófilos, por lo que se puede concluir que los mecanismos de estabilización no fueron moldeados por la selección natural sino, probablemente, por cambios neutrales, es decir, por cambios evolutivos a nivel molecular causados por la deriva génica (Hart *et al.*, 2014).

Hasta antes de este estudio, las curvas de estabilidad determinadas para barriles TIM abarcaban sólo un estudio realizado por Cipolla y colaboradores en 2012, quienes analizan la correlación entre la estabilidad térmica y el ambiente y temperatura fisiológica de α -amilasas derivadas de organismo mesófilos y termófilos. Los autores reportan que existe un compromiso entre los requerimientos para tener un estado nativo estable y una dinámica estructural apropiada que permita la función a temperaturas ambientales y fisiológicas (Cipolla *et al.*, 2012). Desde una perspectiva evolutiva, las variaciones adaptativas de la estabilidad (reflejadas en la curva de estabilidad) parecen estar restringidas a cambios en el ΔH y la T_m , de la misma forma que se observa en las curvas de estabilidad de las BacTIMs.

En el presente trabajo donde se comparan las curvas de estabilidad de barriles TIM naturales y diseñados *de novo* se puede notar que la forma como se modulan las curvas de estabilidad de estas proteínas es distinta. El cambio en la curva de estabilidad para las BacTIMs es principalmente debida a los cambios en la T_s y el ΔH_{TS} , ya que el ΔC_p es casi idéntico entre las diferentes TIMs bacterianas;



lo mismo sucede en la familia de las α -amilasas previamente reportada. En cambio, para los barriles TIM diseñados, sus curvas de estabilidad están moduladas por T_s , ΔH_{TS} y ΔC_p , observando que existen grandes cambios entre ellos y que, por lo tanto, no hay un parámetro que permanezca constante entre todos los diseños (figura 66).

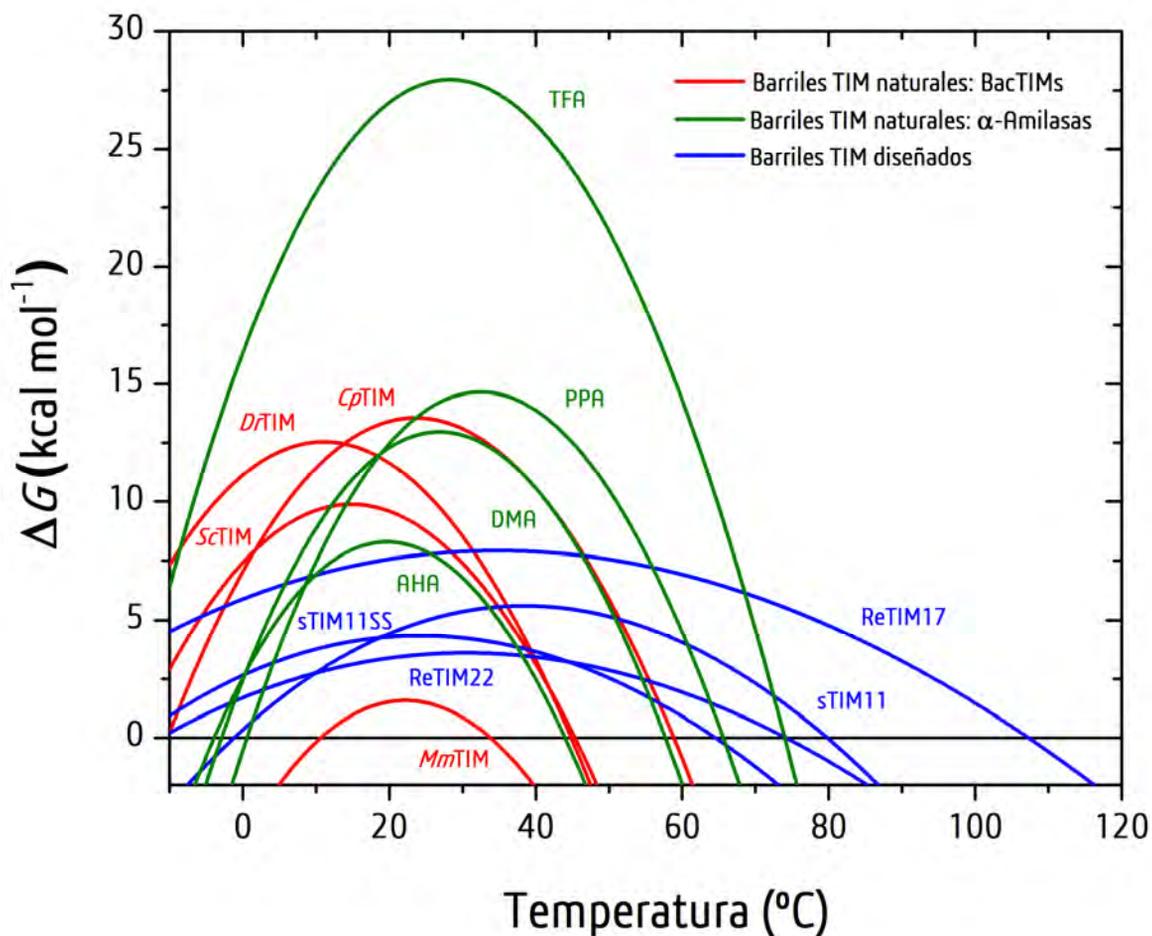
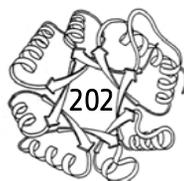


Figura 66. Curvas de estabilidad comparadas entre barriles TIM naturales y diseñados *de novo*. Se muestran las curvas agrupadas en 3 bloques: barriles TIM naturales de las BacTIMs, barriles TIM naturales de las α -amilasas y barriles TIM diseñados. AHA: α -amilasa de *Pseudoalteromonas haloplanktis* –psicrófilo-, DMA: α -amilasa de *Drosophila melanogaster* –mesófilo ectotérmico-, PPA: α -amilasa de *Sus scrofa* –mesófilo endotérmico-, TFA: α -amilasa de *Thermobifida fusca* –termófilo-.



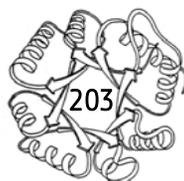
Los resultados presentados en este trabajo confirman la idea de que a pesar de haber estudiado el plegamiento de las proteínas por más de 50 años, aún existen muchas preguntas por contestar. Las pautas que determinan el tipo de control en el desplegamiento (termodinámico o cinético) y la reversibilidad del mismo son prácticamente desconocidas por lo que estudios como el reportado en esta tesis pueden acercarnos a tener una visión más amplia sobre el tema. Lo que hasta ahora conocemos es la punta de un iceberg cuya base día tras día vamos descubriendo. Es probable que las investigaciones futuras permitirán comprender cómo modular de forma racional la estabilidad de las proteínas y, con esto, diseñar proteínas con funciones moleculares novedosas y con alguna aplicación biomédica y/o biotecnológica.

Por último, quisiera terminar el escrito dejando una pregunta abierta: ¿para qué nos sirve aprender a modular la curva de estabilidad? La respuesta a esta pregunta no es única ni definitiva. Además de las posibles aplicaciones que esto tendría en la modificación de proteínas naturales o diseñadas para trabajar bajo condiciones ambientales específicas, el poder modular racionalmente una curva de estabilidad de una proteína comprobaría que somos capaces de lo que hasta ahora no podemos: comprender las bases moleculares de la estabilidad proteica; es decir, el modular una curva de estabilidad demostraría que en verdad entendemos los principios termodinámicos que subyacen a la estructura, función y estabilidad de estas maravillosas macromoléculas llamadas proteínas.



CONCLUSIONES

- i.* Los barriles TIM naturales bacterianos caracterizados en este estudio poseen propiedades estructurales, catalíticas, espectroscópicas e hidrodinámicas compatibles con barriles TIM plegados nativamente y similares a otras Triosafosfato Isomerasas silvestres previamente estudiadas.
- ii.* El desplegamiento térmico de las TIMs bacterianas se agrupan en dos bloques: para *NpTIM* y *GoTIM* el proceso es irreversible, controlado cinéticamente, y para *D τ TIM*, *ScTIM* y *CpTIM* es un proceso reversible bajo control termodinámico, un fenómeno nunca antes observado para una TIM.
- iii.* Los valores de energía de activación para *NpTIM* y *GoTIM* se encuentran entre los reportados previamente para otras TIMs eucariontes y es diferente entre ambas proteínas, lo que se refleja en una estabilidad cinética mayor para *GoTIM* con respecto a *NpTIM*.
- iv.* La estabilidad conformacional a 25 °C (ΔG) para *D τ TIM*, *ScTIM* y *CpTIM* es de la mitad en comparación con *NpTIM*, *GoTIM* y todas las otras TIMs estudiadas anteriormente. Esta característica pudiera ser un “*pago fisicoquímico*” que otorgan las proteínas para tener reversibilidad en su proceso de desplegamiento térmico.



- v. La forma de la curva de estabilidad para $DfTIM$, $ScTIM$ y $CpTIM$ es principalmente debida a los cambios en la T_s y el ΔH_{TS} , ya que el ΔC_p es casi idéntico entre las diferentes RevTIMs.
- vi. Cinco propiedades correlacionan con la reversibilidad en el desplegamiento térmico de los barriles TIM naturales estudiados: 1) bajo punto isoeléctrico, 2) ausencia o baja cantidad de estructura secundaria residual en el estado desplegado, 3) pequeño volumen en las cavidades internas del estado nativo, 4) baja T_m y 5) baja estabilidad conformacional.
- vii. Dado que los tres barriles TIM naturales con desplegamiento térmico reversible provienen de organismos de diferentes fila, esto sugiere que la reversibilidad en el proceso puede ser más común de lo que se conoce actualmente.
- viii. Debido a que los barriles TIM oligoméricos son frecuentes en las células, la reversibilidad en su desplegamiento pudiera ser una ventaja evolutiva al reducir la probabilidad de agregación irreversible no productiva y la carga de la maquinaria celular del sistema de chaperonas.
- ix. Se diseñaron un total de 28 barriles TIM *de novo* con el fin de incrementar la estabilidad de sTIM11. Los diseños más exitosos de cada grupo fueron: ReTIM1, ReTIM6, ReTIM8, ReTIM17, ReTIM18 y ReTIM24.



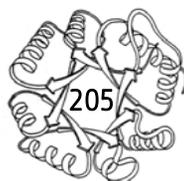
- x. Para la familia de barriles TIM diseñados también existen proteínas con desplegamiento térmico reversible e irreversible.

- xi. De los diseños exitosos, tres de ellos cumplieron el objetivo de incrementar la estabilidad de sTIM11: ReTIM17, ReTIM18 y ReTIM24, quienes poseen los parámetros fisicoquímicos (T_m , ΔH , ΔC_p y ΔG) más altos obtenidos y cercanos a los estimados paramétricamente para una proteína de 184 residuos. De ellos, ReTIM17 posee una estabilidad conformacional mayor en un 50% con respecto de sTIM11.

- xii. Para el caso de los barriles diseñados, T_S , ΔH_{75} y ΔC_p cambian considerablemente entre las distintas ReTIMs, a diferencia de los barriles TIM naturales donde sólo los primeros dos parámetros son los que cambian entre las BacTIMs. El cambio en estos parámetros tiene un efecto específico en la forma de las curvas de estabilidad de las proteínas estudiadas.

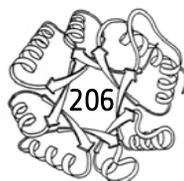
- xiii. Los resultados presentados en este trabajo indican que para distintos miembros, ya sea naturales o diseñados, de una familia que poseen una misma topología, en este caso el barril $(\beta/\alpha)_8$ o barril TIM, el desplegamiento puede estar bajo control cinético o termodinámico.

- xiv. Debido a todo esto, es interesante encontrar que una topología tan común, como lo es el barril TIM, pueda desplegarse/replegarse como propuso Anfinsen en 1973, es decir, sin la ayuda de la maquinaria celular.



PERSPECTIVAS

- i. Resolver las estructuras tridimensionales de las TIMs bacterianas en presencia de PGH.
- ii. Determinar el ΔC_p de unión a PGH para las TIMs bacterianas y proponer una relación energético-estructural que expliquen las diferencias en las firmas termodinámicas del proceso.
- iii. Estudiar el desplegamiento químico de las BacTIMs mediante otras técnicas espectroscópicas, hidrodinámicas y catalíticas.
- iv. Explorar a mayor profundidad las cinco propiedades propuestas que correlacionan con la reversibilidad en el desplegamiento, por ejemplo, estudiar el papel de las cavidades en el proceso mediante mutagénesis dirigida.
- v. Caracterizar cinéticamente el desplegamiento térmico de las BacTIMs.
- vi. Estudiar la evolución de la estabilidad de los barriles TIM naturales mediante métodos de reconstrucción ancestral.
- vii. Completar la caracterización fisicoquímica de los barriles TIM diseñados *de novo*.
- viii. Funcionalizar molecularmente alguno de los barriles TIM diseñados más exitosos: por ejemplo, diseñar un sitio de unión a fosfato o ATP en alguna de las ReTIMs diseñadas.
- ix. Caracterizar otras proteínas, tanto barriles TIM naturales como diseñados, con el objetivo de aumentar el número de miembros en cada grupo con un comportamiento diferencial en su desplegamiento térmico.
- x. Realizar estudios de dinámica molecular para estudiar el desplegamiento de los barriles TIM.

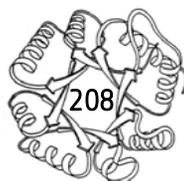


BIBLIOGRAFÍA

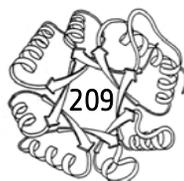
1. **Abendroth, J., T. E. Edwards, y B. Staker.** 2009. **To be published.**
2. **Abbott, I. A. C. A. Lembi, y J. Waaland.** 1989. **1st edition.** New York, USA: Cambridge University Press-Phycological Society of America.
3. **Abriata, L. A., y M. Dal Peraro.** 2015. *Sci Rep.* **5(1):** 1-12.
4. **Adams, P. D., P. V. Afonine, G. Bunkóczy, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, y P. H. Zwart.** 2010. *Acta Cryst Sect D.* **66(2):** 213-221.
5. **Aher, T. J., J. I. Casal, G. A. Petsko, y A. M. Klibanov.** 1987. *Proc Natl Acad Sci USA.* **84(3):** 675-679.
6. **Alahuhta, M., M. Salin, M. G. Casteleijn, C. Kemmer, I. El-Sayed, K. Augustyns, P. Neubauer, y R. K. Wierenga.** 2008. *Prot Eng Design and Select.* **21(4):** 257-266.
7. **Alberg, W. J., y J. R. Knowles.** 1976. *Biochemistry.* **15(25):** 5627-5631.
8. **Alvarez, M., J. P. Zeelen, V. Mainfroid, F. Rentier-Delrue, J. A. Martial, L. Wyns, R. K. Wierenga, y D. Maes.** 1998. *J Biol Chem.* **273(4):** 2199-2206.
9. **Alvarez, M., J. Wouters, D. Maesi, V. Mainfroid, F. Rentier-Delrue, L. Wynsi, E. Depiereux, y J. A. Martial.** 1999. *J Biol Chem.* **274(27):** 19181-19187.
10. **Anantharaman, V., L. Aravind, y E. V. Koonin.** 2003. *Curr Opin Chem Biol.* **7(1):** 12-20.
11. **Andreeva, A., D. Howorth, J. M. Chandonia, S. E. Brenner, T. J. Hubbard, C. Chotia, y A. G. Murzin.** 2008. *Nucleic Acids Res.* **36(D1):** D419-D425.
12. **Anfisen, C. B.** 1973. *Science.* **181(96):** 223-230.
13. **Anfinsen, C. B., E. Haber, M. Sela, y F. H. White.** 1961. *Proc Nat Acad Sci USA.* **47(2):** 1309-1314.
14. **Aparicio, R., S. T. Ferreira, y I. Polikarpov.** 2003. *J Mol Biol.* **334(5):** 1023-1041.
15. **Arolas, J. L., F. X. Aviles, J-Y. Chang, y S. Ventura.** 2006. *Trends Biochem Sci.* **31(5):** 292-301.
16. **Arrhenius, S. A.** 1889. *Z Phys Chem.* **4(1):** 96-116.



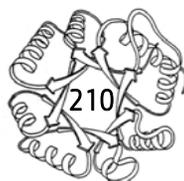
17. Banaszak, L. N. Winter, Z. Xu, D. A. Bernlohr, S. Cowan, y T. A. Jones. 1994. *Adv Prot Chem.* **45(4)**: 89-151.
18. Banner, D. W., A. C. Bloomer, G. A. Petsko, D. C. Phillips, C. I. Pogson, I. A. Wilson, P. H. Corran, A. J. Furth, J. D. Milman, R. E. Offord, J. D. Priddle, y S. G. Waley. 1976. *Nature.* **255(5510)**: 609-614.
19. Battista, J.R. 1997. *Ann Rev Microbiol.* **51(3)**: 203-224.
20. Baugh, L., L. A. Gallagher, R. Patrapuvich, M. C. Clifton, A. S. Gardberg, T. E. Edwards, B. Amour, D. W. Begley, S. H. Dieterich, D. M. Dranow, J. Abendroth, J. W. Fairman, D. Fox, B. L. Staker, I. Phan, A. Gillespie, R. Choi, S. Nakazawa-Hewitt, M. T. Nguyen, A. Napuli, L. Barrett, G. W. Buchko, R. Stacy, P. J. Myler, L. J. Stewart, C. Manoil, y C. Van Voorthis. 2013. *PLoS One.* **8(1)**: e53851.
21. Beaucamp, N., A. Hofmann, B. Kellerer, y R. Jeanicke. 1997. *Prot Sci.* **6(10)**: 2159-2165.
22. Becktel, W. J., y J. A. Schellman. 1987. Protein stability curves. *Biopolymers.* **26(5)**: 1859-1877.
23. Beffa, T., M. Blanc, P. F. Lyon, G. Vogt, M. Marchiani, J. L. Fischer, y M. Aragno. 1996. *Appl Environ Microbiol.* **62(5)**: 1723-1727.
24. Bellows-Peterson, M. L., H. K. Fung, C. A. Floudas, C. A. Kieslich, L. Zhang, D. Morikis, K. J. Wareham, P. N. Monk, O. A. Hawksworth, y T. M. Woodruff. 2012. *J Med Chem.* **55(9)**: 4159-4168.
25. Benítez-Cardoza, C.G., A. Rojo-Domínguez, y A. Hernández-Arana. 2001. *Biochemistry.* **40(30)**: 9049-9058.
26. Bentley, S. D., K. F. Chater, A. M. Cerdeño-Tárraga, G. L. Challis, N. R. Thomson, K. D. James, D. E. Harris, M. A. Quail, H. Kieser, D. Harper, A. Bateman, S. Brown, G. Chandra, C. W. Chen, M. Collins, A. Cronin, A. Fraser, A. Goble, J. Hidalgo, T. Hornsby, S. Howarth, C. H. Huang, T. Kieser, L. Larke, L. Murphy, K. Oliver, S. O'Neil, E. Rabinowitsch, M.-A. Rajandream, K. Rutherford, S. Rutter, K. Seeger, D. Saunders, S. Sharp, R. Squares, S. Squares, K. Taylor, T. Warren, A. Wietzorrek, J. Woodward, B. G. Barrell, J. Parkhill, y D. A. Hopwood. 2002. *Nature.* **417(2)**: 141-147.
27. Berg, J. M., J. L. Tymoczko, y L. Stryer. 2002. 5th edition. New York, USA: WH Freeman.
28. Berka, K., O. Hanák, D. Sehnal, P. Banás, V. Navrátilová, D. Jaiwal, C-M, Ionescu, R. S. Vareková, J. Koca, y M. Otyepka. 2012. *Nucleic Acids Res.* **40(1)**: W222-W227.
29. Bernstein, F. C., T. F. Koetzle, G. J. Williams, E. E. Meyer Jr., M. D. Brice, J. R. Rodgers, O. Kennard, T. Shimanouchi, y M. Tasumi. 1977. *J Mol Biol.* **112(3)**: 535-542.
30. Bilgimol, C. J., P. Suthakaran, S. Sankaranarayanan, M. Musti, S. Kalimuthu, M. Genesan, y M. Sadananda. 2015. *J Cell Sci Ther.* **6(1)**: 5-12.



31. Böhm, G., R. Muhr, y R. Jaenicke. 1992. *Protein Eng.* **5(2)**: 191-195.
32. Borchert, T. V., K. Prat, J. Zeelen, M. Callens, M. E. Noble, F. R. Opperdoes, P. A. M. Michels, y R. K. Wierenga. 1993. *Eur J Bioche.* **211(3)**: 703-710.
33. Borchert, T.V., K. V. Kishan, J. P. Zeelen, W. Schliebs, N. Thanki, R. Abagyan, R. Jaenicke, y R. K. Wierenga. 1995. *Structure.* **3(7)**: 669-679.
34. Borchert, T. V., R. Abagyan, R. Jaenicke, y R. K. Wierenga. 1994. *Proc Natl Acad Sci USA.* **91(2)**: 1515-1518.
35. Bozaykut, P., N. K. Ozer, y B. Karademir. 2014. *Free Radic Biol Med.* **77(2)**: 195-209.
36. Brandts, J. F. 1964. *J Amer Chem Soc.* **86(20)**: 4291-4301.
37. Brenner, D. J., N. R. Krieg, y J. T. Staley. 2005. 2nd edition. New York, USA: Springer.
38. Brim, H., S. C. McFarlan, J. K. Fredrickson, K. W. Minton, M. Zhai, L. P. Wackett, y M. J. Daly. 2000. *Nat Biotech.* **18(1)**: 85-90.
39. Bromley, E. H. C., K. Channon, E. Moutevelis, y D. N. Woolfson. 2008. *ACS Chem Biol.* **3(1)**: 38-50.
40. Brook, I. 2007. *Anaerobe.* **13(5-6)**: 171-177.
41. Bryngelson, J. D. 1995. *Proteins.* **21(4)**: 164-195.
42. Bursy, J., A. U. Kuhlmann, M. Pittelkow, H. Hartmann, M. Jebbar, A. J. Pierik, y E. Bremer. 2008. *Appl Environ Microbiol.* **72(3)**: 4561-4568.
43. Cabrera, N., G. Hernández-Alcántara, G. Mendoza-Hernández, A. Gómez-Puyou, y R. Pérez-Monfort. 2008. *Biochemistry.* **47(11)**: 3499-3506.
44. Caetano-Anollés, G., H. S. Lim, y J. E. Mittenthal. 2007. *Proc Natl Acad Sci USA.* **104(22)**: 9358-9363.
45. Casal, J. I., T. J. Ahern, R. C. Davenport, G. A. Petsko, y A. M. Klibanov. 1987. *Biochemistry.* **100(3)**: 702-710.
46. Cavalier-Smith, T. 2002. *International J SystEvol Microbiol.* **52(2)**: 297-354.
47. Cavalier-Smith, T. 2006a. *Philos Trans R Soc Lond B Biol Sci.* **361(1470)**: 969-1006.
48. Cavalier-Smith, T. 2006b. *Biol Direct.* **1(1)**: 19.
49. Cháñez-Cárdenas, M. E., D. A. Fernández-Velasco, E. Vázquez-Contreras, R. Coria-Ortega, G. Saab-Rincón, y R. Pérez-Montfort. 2002. *Arch Biochem Biophys.* **399(2)**: 117-129.



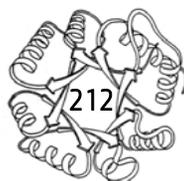
50. Cháñez-Cárdenas, M. E., G. Pérez-Hernández, B. G. Sánchez-Rebollar, M. Costas, y E. Vázquez-Contreras. 2005. *Biochemistry*. **44(32)**: 10883-10892.
51. Chang, C., N. Maltseva, Y. Kim, S. Shatsman, A. Joachimiak, W. F. Anderson, y Center for Structural Genomics of Infectious Diseases. 2017. To be published.
52. Cheung, L. S., D. J. Shea, N. Nicholes, A. Date, M. Ostermeier, y K. Konstantopoulos. 2015. *Sci Rep*. **5(1)**: 1-9.
53. Chiarabelli, C. D. De Lucrezia, P. Stano, y P. L. Luisi. 2009. *Orig Life Evol Biosph*. **39(3)**: 308-309.
54. Chiti, F., N. Taddei, P. M. White, M. Bucciantini, F. Magherini, M. Stefani, y C. M. Dobson. 1999. *Nat Struct Biol*. **6(2)**: 1005-1009.
55. Chu, C-H., Y-J. Lai, H. Huang, y Y. J. Sun. 2008. *Proteins*. **71(2)**: 396-406.
56. Ciccarelli, F.D., T. Doerks, C. von Mering, C. J. Creevey, B. Snel, y P. Bork. 2006. *Science*. **311(5765)**: 1283-1287.
57. Cipolla, A., F. Delbrassine, J. L. Da Lage, y G. Feller. 2012. *Biochimie*. **94(9)**: 1943-1950.
58. Cisneros-Armas, D. A. 2003. Tesis de grado. Licenciatura en Investigación Biomédica Básica. Universidad Nacional Autónoma de México. México, D.F.
59. Claren, J., C. Malisi, B. Höcker, y R. Sterner. 2009. *Proc Natl Acad Sci USA*. **106(10)**: 3704-3709.
60. Clarke, J., E. Cota, S. B. Fowler, y S. J. Hamill. 1999. *Structure*. **7(9)**: 1145-1153.
61. Cohen, N., S. Abramov, Y. Dror, y A. Freeman. 2001. *Trends Biotechnol*. **19(12)**: 507-510.
62. Coleman, R.G., y K. A. Sharp. 2006. *J Mol Biol*. **362(3)**: 441-458.
63. Connor, S. E., G. C. Capodagli, M. K. Deaton, y S. D. Pegan. 2011. *Acta Crystallogr D Biol Crystallogr*. **67(Pt12)**: 1017-1022.
64. Copley, R. R., y P. Bork. 2000. *J Mol Biol*. **303(4)**: 627-641.
65. Correia, B. E., J. T. Bates, R. J. Loomis, G. Baneyx, C. Carrico, J. G. Jardine, P. Rupert, C. Correnti, O. Kalyuzhniy, V. Vittal, M. J. Connell, E. Stevens, A. Schroeter, M. Chen, S. Macpherson, A. M. Serra, Y. Adachi, M. A. Holmes, Y. Li, R. E. Klevit, B. S. Graham, R. T. Wyatt, D. Baker, R. K. Strong, J. E. Crowe, P. R. Johnson, y W. R. Schief. 2014. *Nature*. **507(7491)**: 201-206.
66. Costas, M., D. Rodríguez-Larrea, L. de María, T. V. Borchert, A. Gómez-Puyou, y J. M. Sánchez-Ruiz. 2009. *J Mol Biol*. **385(2)**: 924-937.



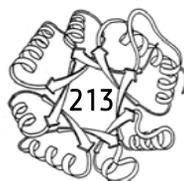
67. Creighton, T. E. 1990. *Biochemistry*. **270(1)**: 1-16.
68. Creighton, T. E. 1996. **2nd edition**. New York, USA: W. H. Freeman and Company.
69. Currin, A., N. Swainston, P. J. Day, D. B. Kell. 2015. *Chem Soc Rev*. **44(3)**: 1172-1239.
70. Dalby, P. A. 2007. *Rec Pat on Biotech*. **1(1)**: 1-9.
71. D'Alessio, G. 1999. *Biochemistry*. **72(1)**: 271-298.
72. Dalessio, P. M., y I. J. Ropson. 2000. *Biochemistry*. **39(2)**: 860-871.
73. Daly, M. J. 2009. *Nat Rev Microbiol*. **7(1)**: 237-245.
74. Daly, M. J., E. K. Gaidamakova, V. Y. Matrosova, A. Vasilenko, M. Zhai, A. Venkateswaran, M. Hess, M. V. Omerchenko, H. M. Kostandarithes, K. S. Makarova, L. P. Wackett, J. K. Fredickson, y D. Ghosal. 2004. *Science*. **306(5698)**: 1025-1028.
75. Dagan, T., M. Roettger, D. Bryant, y W. Martin. 2010. *Gen Biol Evol*. **2(1)**: 379-392.
76. Dahiyat, B. I., y S. L. Mayo. 1996. *Protein Sci*. **5(5)**: 895-903.
77. Dahiyat, B. I., y S. L. Mayo. 1997. *Science*. **278(5335)**: 82-87.
78. Davis, I. W., A. Leaver-Fay, V. B. Chen, J. N. Block, G. J. Kapral, X. Wang, L. W. Murray, W. B. Arendall III, J. Snoeyink, J. S. Richardson, y D. C. Richardson. 2007. *Nucl Acids Res*. **35(W)**: W375-W383.
79. Delboni, L. F., S. C. Mande, F. Reintier-Delrue, V. Mainfroid, S. Turtley, F. M.D. Vellieux, J. A. Martial, y W. G. J. Hol. 1995. *Prot Sci*. **4(12)**: 2594-604.
80. Dill, K. A. 1990. *Biochemistry*. **29(31)**: 7133-7155.
81. Dill, K. A., y D. Shortle. 1991. *Annu Rev Biochem*. **60(5)**: 795-825.
82. Dill, K. A., y H. S. Chan. 1997. *Nat Struct Biol*. **4(3)**: 10-18.
83. Dill, K. A., y J. L. MacCallum. 2012. *Science*. **338(6110)**: 1042-1046.
84. Dobson, C. M. 2003. *Nature*. **18526(6968)**: 884-890.
85. Doyle, C. N., J. A. Rumfeldt, H. R. Broom, A. Broom, P. B. Stathopoulos, K. A. Vassall, J. J. Almey, y E. M. Meiering. 2013. *Arch. Biochem. Biophys*. **531(1-2)**: 44-64.



86. Enríquez-Flores, S., A. Rodríguez-Romero, G. Hernández-Alcántara, J. Oria-Hernández, P. Gutiérrez-Castrellon, G. Pérez-Hernández, I. de la Mora-de la Mora, A. Castillo-Villanueva, I. Garcia-Torres, S. T. Mendez, S. Gómez-Manzo, A. Torres-Arroyo, G. López-Velázquez, y H. Reyes-Vivas. 2011. *Proteins*. **79(9)**: 2711-2724.
87. Emsley, P., y K. Cowtan. 2004. *Acta Crystallogr Sect D*. **60(12)**: 2126-2132.
88. Felter, R. A., R. Colwell, y G. Chapman. 1969. *J Bact*. **99(1)**: 326-335.
89. Feng, H., Z. Zhou, y Y. Bai. 2005. *Proc Natl Acad Sci USA*. **102(14)**: 5026-5031.
90. Ferguson, N., A. P. Capaldi, R. James, C. Kleanthous, y S. E. Radford. 1999. *J Mol Biol*. **286(6)**: 1597-1608.
91. Ferrer, M., F. Martínez-Abarcam y P. N. Golyshin. 2005. *Curr Opin Drug Discov Devel*. **9(2)**: 269-275.
92. Fersht, A. 1999. 1st edition. New York, USA: Freeman.
93. Figueroa-Angulo, E. E., P. Estrella-Hernández, H. Salgado-Lugo, A. Ochoa-Leyva, A. Gómez-Puyou, S. S. Campos, G. Montero-Moran, J. Ortega-López, G. Saab-Rincón, R. Arroyo, C. Benítez-Cardoza, y L. G. Briebe. 2012. *Parasitology*. **139(2)**: 1729-1738.
94. Figueroa, M., N. Oliveira, A. Lejeune, K. W. Kaufmann, B. M. Dorr, A. Matagne, J. A. Martial, J. Meiler, y C. Van de Weerd. 2013. *PLoS One*. **8(8)**: e71858.
95. Filippova, E. V., Z. Wawrzak, G. Minasov, A. Cardona-Correa, B. Bishop, W. F. Anderson, H. Ngo, y Center for Structural Genomics of Infectious Diseases. 2017. To be published.
96. Fleming, P. J., y G. D. Rose. 2005. *Protein Sci*. **14(7)**: 1911-1917.
97. Frank, H. S., y M. W. Evans. 1945. *J Chem Phys*. **13(11)**: 507-532.
98. Freire, E. 1995., en: B. A. Shirley. 1st edition. Berlin, Germany: Springer. 191-218.
99. Fuerst, J. A., y R. I. Webb. 1991. *Proc Natl Acad Sci USA*. **88(1)**: 8184-8188.
100. Garel, J. R. 1992, en: T. E. Creighton. 2nd edition. New York, USA: W. H. Freeman and Company. 405-454.
101. Garza-Ramos, G., M. T. de Gómez-Puyou, y R. W. Gracy. 1992. *Eur J Biochem*. **208(2)**: 389-395.
102. Gasteiger, E., C. Hoogland, A. Gattiker, S. Duvaud, M. R. Wilkins, R. D. Appel, y A. Bairoch. 2005, en: Walker, J. M. 1st edition. New York, USA: Humana Press. 571-607.
103. Gayathri, P., M. Banerjee, A. Vijayalakshmi, S. Azeez, H. Balaram, P. Balaram, y M. R. N. Murthy. 2007. *Acta Crystallogr Sect D*. **63(2)**: 206-220.



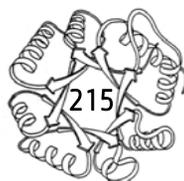
104. Gething, M. J., y J. Sambrook. 1992. *Nature*. **355(6355)**: 33-45.
105. Gokhale, R.S., S. S. Ray, H. Balaram, y P. Balaram. 1999. *Biochemistry*. **38(1)**: 423-431.
106. Gómez-Puyou, A., E. Saavedra-Lira, I. Becker, R. A. Zubillaga, A. Rojo-Domínguez, y R. Pérez-Montfort. 1995. *J Chem Biol*. **2(12)**: 847-55.
107. González-Mondragón, E., R. A. Zubillaga, E. Saavedra, M. E. Cháñez-Cárdenas, R. Pérez-Monfort, y A. Hernández-Arana. 2004. *Biochemistry*. **43(11)**: 3255-3263.
108. Goodsell, D. S., y A. J. Olson. 2000. *Annu Rev Biophys Biomol Struct*. **20(1)**: 105-153.
109. Gopal, B., S. S. Ray, R. S. Gokhale, H. Balaram, M. R. N. Murthy, y P. Balaram. 1999. *Biochemistry*. **38(1)**: 478-486.
110. Goraj, K., A. Renard, y J. A. Martial. 1990. *Protein Eng*. **3(4)**: 259-266.
111. Gruebele, M. 2005. *C R Biol*. **238(8)**: 701-712.
112. Gunasekaran, K., S. J. Eyles, A. T. Hagler, y L. M. Gierasch. 2001. *Curr Opin in Struct Biol*. **11(1)**: 83-93.
113. Guruprasad, K., B. V. Reddy, y M. W. Pandit. 1990. *Prot Eng*. **4(2)**: 155-161.
114. Guzmán-Luna, V., y G. Garza-Ramos. 2012. *Proteins*. **80(2)**: 1669-1682.
115. Hadley, C., y D. T. Jones. 1999. *Structure*. **7(9)**: 1099-1112.
116. Hart, K. M., M. J. Harms, B. H. Schmidt, C. Elya, J. W. Thornton, y S. Marqusee. 2014. *PLoS Biol*. **12(3)**: 1-12.
117. Heinig, M., y D. Frishman. 2004. *Nucl Acids Res*. **32(2)**: W500-W502.
118. Hernández-Alcántara, G., G. Garza-Ramos, G. M. Hernández, A. Gómez-Puyou, y R. Pérez-Montfort. 2002. *Biochemistry*. **41(13)**: 4230-4238.
119. Hernández-Alcántara, G., A. Rodríguez-Romero, H. Reyes-Vivas, J. Peon, N. Cabrera, C. Ortíz, S. Enríquez-Flores, I. de la Mora-de la Mora, y G. López-Velázquez. 2008. *Biochim Biophys Acta*. **1784(11)**: 1493-1500.
120. Höcker, B. 2005. *Biomol Eng*. **22(1-3)**: 31-38.
121. Höcker, B., J. Claren, y R. Sterner. 2004. *Proc Natl Acad Sci USA*. **101(47)**: 16448-16453.
122. Höcker, B. A. Lochner, T. Seitz, J. Claren, y R. Sterner. 2009. *Biochemistry*. **48(6)**: 1145-1157.
123. Höcker, B., S. Schmidt, y R. Sterner. 2002. *FEBS Lett*. **510(3)**: 133-135.



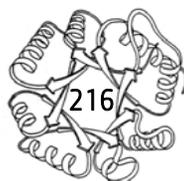
124. Hosur, V., K. R. Johnson, L. M. Burzenski, T. M. Stearns, R. S. Maser, y L. D. Shultz. 2014. *Proc Natl Acad Sci USA*. **111(21)**: E2200-2209.
125. Houbrechts, A., B. Moreau, R. Abagyam, V. Mainfroid, G. Préaux, A. Lamproye, A. Poncin, E. Goormaghtigh, J. M. Ruyschaert, y A. Martial. 1995. *Protein Eng*. **8(3)**: 249-259.
126. Huang, P. S., G. Oberdorfer, C. Xu, X. Y. Pei, B. L. Nannenga, J. M. Rogers, F. DiMaio, T. Gonen, B. Luisi, y D. Baker. 2014. **346(6208)**: 481-485.
127. Huang, P. S., K. Feldmeier, F. Parmeggiani, D. A. Fernández-Velasco, B. Höcker, y D. Baker. 2016a. *Nat. Chem. Biol.* **12(4)**: 29-36.
128. Huang, P. S., S. E. Boyksen, y D. Baker. 2006b. *Nature*. **537(7620)**: 320-327.
129. Hubbard, S. J., y J. M. Thornton. 1993. NACCESS, University College London, London, U. K.
130. Ibarra-Molero, B., y J. M. Sánchez-Ruíz. 2006. en: J. L. R. Arrondo, y A. Alonso. **1st edition**. Berlin, Germany: Springer. 27-48.
131. Jaenicke, R. 1991. *Adv Protein Chem*. **30(13)**: 3147-3461.
132. Jiménez, L., D. A. Fernández-Velasco, K. Willms, y A. Landa. 2003. *J Parasitol*. **89(2)**: 209-214.
133. Jiménez, L., N. Vibanco-Pérez, L. Navarro, y A. Landa. 2000. *Int J Parasitol*. **30(9)**: 1007-1012.
134. Joh, N. H., T. Wang, M. P. Bhate, R. Acharya, Y. Wu, M. Grabe, M. Hong, G. Grigoryan, y W. F. DeGrado. 2014. *Science*. **346(6216)**: 1520-1524.
135. Jogl, G., S. Rozovsky, A. E. McDermott, y L. Tong. 2003. *Proc Natl Acad Sci USA*. **100(1)**: 50-55.
136. Joseph-McCarthy, D., E. Lolis, E. A. Komives, y G. A. Petsko. 1994. *Biochemistry*. **33(10)**: 2815-2823.
137. Joubert, F., A. W. H. Neitz, y A. I. Louw. 2001. *Prot Struct Funct Genet*. **45(2)**: 136-143.
138. Kabsch, W. 2010. *Acta Crystallogr Sect D*. **66(2)**: 125-132.
139. Katoh, K., y D. M. Standley. 2013. *Mol Biol Evol*. **30(4)**: 772-780.
140. Kauzmann, W. 1959. *Adv Prot Chem*. **14(2)**: 1-63.
141. Kerner, M. J., D. J. Naylor, Y. Ishihama, T. Maier, H. C. Chang, A. P. Stines, C. Georgopoulos, D. Frishman, M. Hayer-Hartl, M. Mann, y F. Urlich. 2005. *Cell*. **122(2)**: 209-220.
142. Khoury, G. A., J. Smadbeck, C. A. Kielsich, y C. A. Floudas. 2014. *Trends Biotechnol*. **32(2)**: 99-109.



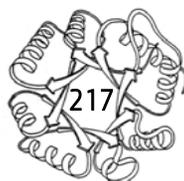
143. Kim, D. E., C. Fisher, y D. Baker. 2000. *J Mol Biol.* **298(5)**: 971–984.
144. Kiss, G., N. Çelebi-Ölçüm, R. Moretti, D. Baker, y K. N. Houk. 2013. *Angew Chem Int Ed Engl.* **52(22)**: 5700-5725.
145. Knobloch, D., A. Schmidt, P. Scheerer, N. Krauss, H. Wessner, Ch. Scholz, G. Küettner, T. Von Rintelen, A. Wessel, y W. Höehne. 2010. *Ins Mol Biol.* **19(1)**: 35-48.
146. Knowles, J. R. 1991. *Nature.* **350(2)**: 121-124.
147. Knowles, J. R., y W. J. Albery. 1977. *Acc Chem Res.* **10(4)**: 105-111.
148. Koga, N., R. Tatsumi-Koga, G. Liu, R. Xiao, T. B. Acton, G. T. Montelione, y D. Baker. 2012. *Nature.* **491(7423)**: 222-227.
149. Kohl, L., M. Callens, R. K. Wierenga, F. R. Opperdoes, y P. A. Michels. 1994. *Eur J Biochem.* **220(2)**: 331-338.
150. Komives, E. A., L. C. Chang, E. Lolis, R. F. Tilton, G. A. Petsko, y J. R. Knowles. 1991. *Biochemistry.* **30(12)**: 3011-3019.
151. Kragelund, B. B., P. Højrup, M. S. Jensen, C. K. Schjerling, E. Juul, J. Knudsen, y F. M. Poulsen. 1996. *J Mol Biol.* **256(1)**: 187–200.
152. Krebs, H., F. X. Schmid, y R. Jaenicke. 1983. *J Mol Biol.* **169(2)**: 619-635.
153. Krietsch, W. K. G., P. G. Pentchev, H. Klingenburg, T. Hofstätter, y T. Bücher. 1970. *Eur J Biochem.* **14(2)**: 289-300.
154. Krissinel, E., y K. Henrick. 2007. *J Mol Biol.* **372(3)**: 774-797.
155. Kuaprasert, B., P. Riengrunroj, W. Pornthanakasem, W. Suginta, M. Mungthin, S. Leelayoova, y U. Leartsakulpanich. 2013. To be published.
156. Kuhlman, B., y D. Baker. 2000. *Proc Natl Acad Sci USA.* **97(19)**: 10383-10388.
157. Kuhlman, B., G. Dantas, G. C. Ireton, G. Varani, B. L. Stoddard, y D. Baker. *Science.* **302(5649)**: 1364-1368.
158. Kumar, K., P. Bhargava, y U. Roy. 2012. *Exp Parasitol.* **130(4)**: 430-436.
159. Lake, J. A., R. G. Skophammer, C. W. Herbold, y J. A. Servin. 2009. *Phil Trans R Soc B.* **364(1)**: 2177-2185.
160. Lambeir, A. M., F. R. Opperdoes, y R. K. Wierenga. 1987. *Eur J Biochem.* **168(1)**: 69-74.
161. Lambeir, A. M., J. Backmann, J. Ruiz-Sanz, V. Filimonov, J. E. Nielsen, I. Kursula, B. V. Norledge, y R. K. Wierenga. 2000. *Eur J Biochem.* **267(9)**: 2516-2524.



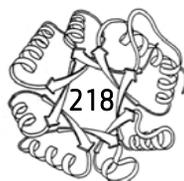
162. Landa, A., A. Rojo-Domínguez, L. Jiménez, y D. A. Fernández-Velasco. 1997. *Eur J Biochem.* **247(2)**: 348-355.
163. Lara-González, S., P. Estrella-Hernández, A. Ochoa-Leyva, M. Del Carmen Portillo-Téllez, L. A. Caro-Gómez, E. E. Figueroa-Angulo, H. Salgado-Lugo, J. F. T. Miranda-Ozuna, J. Ortega-López, R. Arroyo, L. G. Briebe y C. B. Benítez-Cardoza. 2014. *Proteins.* **82(9)**: 22-33.
164. Lecomte, T. J. J., D. A. Vuletich, y A. M. Lesk. 2005. *Curr Opin in Struct Biol.* **15(2)**: 290-301.
165. Letunic, I., y P. Bork. 2011. *Nucleic Acids Res.* **39(3)**: W475-W478.
166. Levitt, M., y C. Chothia. 1976. *Nature.* **261(5561)**: 552-558.
167. Li, D. B. Ji, K. Hwang, y Y. Huang. 2011. *PLoS One.* **6(2)**: e19268.
168. Li, J., y B. A. McClane. 2006. *Appl Environ Microbiol.* **72(2)**: 4561-4568.
169. Lindquist, S. 1986. *Ann Rev Biochem.* **55(3)**: 1151-1191.
170. Lodi, P. J., L. C. Chang, J. R. Knowles, y E. A. Komives. 1994. *Biochemistry.* **33(10)**: 2809-2814.
171. Lolis, E., y G. A. Petsko. 1990. *Biochemistry.* **29(28)**: 6619-6625.
172. López-Castillo, L. M., P. Jiménez-Sandoval, N. Baruch-Torres, C. H. Trasviña-Arenas, C. Díaz-Quezada, S. Lara-González, R. Winkler, y L. G. Briebe. 2016. *Front Plan Sci.* **7(1)**: 1817.
173. López-Velázquez, G., D. Molina-Ortíz, N. Cabrera, G. Hernández-Alcántara, J. Peón-Peralta, L. Yopez-Mulia, R. Pérez-Montfort, y H. Reyes-Vivas. 2004. *Proteins.* **55(4)**: 824-834.
174. López-Zavala, A. A., J. S. Carrasco-Miranda, C. D. Ramírez-Aguirre, M. López-Hidalgom C. G. Benitez-Cardoza, A. Ochoa-Leyva, C. S. Cardona-Felix, C. Díaz-Quezada, E. Rudiño-Piñera, R. R. Sotelo-Mundo, L. G. Briebe. 2016. *Biochim Biophys Acta.* **1864(12)**: 1696-1706.
175. Lorenz, P., y J. Eck. 2005. *Nat Rev Microbiol.* **3(6)**: 510-516.
176. Love, A. 2015. 1st edition. New York, USA: Springer.
177. Lukacs, C. M., J. W. Fairman, T. E. Edwards y D. Lorimer. 2013. To be published.
178. Lyubarev, A. E., B. I. Kurganov, A. A. Burlakova, y V. N. Orlov. 1998. *Biophys Chem.* **70(3)**: 247-257.
179. Madan, B., y K. A. Sharp. 2001. *Biophys J.* **81(4)**: 1881-1887.
180. Madden, T. L., R. L. Tatusov, y J. Zhang. 1996. *Methods Enzymol.* **266(1)**: 131-141.



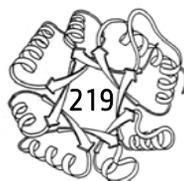
181. Maes, D., J. P. Zeelen, N. Thanki, N. Beaucamp, M. Alvarez, M. H. D. Thi, J. Backmann, J. A. Martial, L. Wyns, R. Jaenicke, y R. K. Wierenga. 1999. *37*(3): 441-453.
182. Mainfroid, V., S. C. Mande, W. G. J. Hol, J. Martial, y K. Goraj. 1996a. *Biochemistry*. **35**(13): 4110-4117.
183. Mainfroid, V., P. Terpstra, M. Beauregard, J-M. Frère, S. C. Mye, W. G. J. Hol, J. A. Martial, y K. Goraj. 1996b. *J Mol Biol*. **257**(2): 441-456.
184. Maithal, K., G. Ravindra, H. Balaram, y P. Balaram. 2002. *J Biol Chem*. **277**(28): 25106-25114.
185. Makhatadze, G. I., V. N. Medvedkin, y P. L. Privalov. 1990. *Biopolymers*. **30**(11-12): 1001-1010.
186. Maldonado, E., M. Soriano-Garcia, A. Moreno, N. Cabrera, G. Garza-Ramos, M. T. de Gómez-Puyou, A. Gómez-Puyou, y R. Pérez-Montfort. 1998. *J Mol Biol*. **283**(1): 193-203.
187. Mande, S. C., V. Mainfroid, K. H. Kalk, K. Goraj, J. A. Martial, y W. G. Hol. 1994. *Prot Sci*. **3**(5): 810-821.
188. Martínez-Castilla, L., y R. Rodríguez-Sotres. 2010. *PLoS One*. **5**(9): e12483.
189. Mathur, D., G. Malik, y L. C. Garg. 2006. *FEMS Microbiol Lett*. **263**(2): 229-235.
190. Mathews, C. K., K. E. Van Holde, D. R. Appling, y S. J. Anthony-Cahill. 2012. **4th edition**. New Jersey, USA: Prentice Hall.
191. Mattimore, V., y J. R. Battista. 1996. *J Bact*. **178**(3): 633-637.
192. McCallister, E. L., E. Alm, y D. Baker. 2000. *Nat Struct Biol*. **7**(8): 669-673.
193. McCoy, A. J., R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni y R. J. Read. 2007. *J Appl Cryst*. **40**(4): 658-674.
194. McDonald, I., y J. Thornton. 1994. *J Mol Biol*. **238**(5): 777-793.
195. Milardi, D., C. la Rosa, S. Fasone, y D. Grasso. 1997. *Biophys Chem*. **69**(5): 43-51.
196. Minasov, G., L. Shuvalova, I. Dubrovskaya, K. Flores, S. Shatman, K. Kwon, W. F. Anderson, y Center for Structural Genomics of Infectious Diseases. 2016. **To be published**.
197. Minelli, A. 2009. **1st edition**. New Jersey, USA: Princeton University Press.
198. Mixcoha-Hernández, E., L. M. Moreno-Vargas, A. Rojo-Domínguez, y C. G. Benítez-Cardoza. 2007. *Protein J*. **26**(3): 491-498.



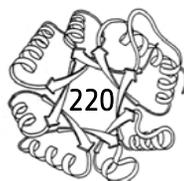
199. Moraes, J., R. Arreola, N. Cabrera, L. Saramago, D. Freitas, A. Masuda, I. Vaz da Silva, M. T. de Gómez-Puyou, R. Pérez-Montfort, A. Gómez-Puyou, y C. Logullo. 2011. *Insect Biochem Mol Biol.* **41(2)**: 400-409.
200. Morgan, C. J., D. K. Wilkins, L. J. Smith, Y. Kawata, y C. M. Dobson. 2000. *J Mol Biol.* **300(1)**: 11-16.
201. Muga, A., y F. Moro. 2008. *Curr Protein Pept Sci.* **9(3)**: 522-566.
202. Mukherjee, S., A. Roychowdhury, D. Dutta, B. Saha, y A. K. Das. 2011. **To be published.**
203. Muñoz-Trejo, R. 2007. Tesis de grado. Licenciatura en Biología. Universidad Nacional Autónoma de México. Ciudad de México.
204. Murphy, K. P., V. Bhakuni, D. Xie, y E. Freire. 1992. *J Mol Biol.* **227(3)**: 293-306.
205. Murphy, K. P., y E. Freire. 1992. *Adv Protein Chem.* **43(1)**: 313-361.
206. Murray, R. G. E., en: A. Balows, H. G. Trüper, M. Dworkin, W. Harder, y K. H. Schleifer. **2dn edition.** New York, USA: Springer-Verlag. 3732-3744.
207. Myers, J. K., C. N. Pace, y J. M. Scholtz. 1995. *Prot Sci.* **4(3)**: 2138-2148.
208. Myers, T. W., y D. H. Gelfand. 1991. *Biochemistry.* **30(31)**: 7661-7666.
209. Nagarajan, D., G. Deka, y M. Rao. 2015. *BMC Biochem.* **16(1)**: 18.
210. Nagano, N., C. A. Orengo, y J. M. Thornton. 2002. *J Mol Biol.* **321(5)**: 741-765.
211. Nájera, H., M. Costas, y D. A. Fernández-Velasco. 2003. *Biochem J.* **370(3)**: 785-792.
212. Nakamura, T., K. Makabe, K. Tomoyori, K. Maki, A. Mukaiyama, y K. Kuwajima. 2010. *J Mol Biol.* **396(8)**: 1361-1378.
213. Nelson, D. L., y M. M. Cox. 2000. **3rd edition.** New York, USA: Worth.
214. Nicholls, A., K. A. Sharp, y B. Honig. 1991. *Prot Struct Func Genet.* **11(4)**: 281-296.
215. Nishimura, C., S. Prytulla, H. J. Dyson., y P. E. Wright. 2000. *NatStruct Biol.* **7(8)**: 679-687.
216. Noble, M. E., J. P. Zeelen, y R. K. Wierenga. 1993. *Acta Crystallogr Sect D.* **49(4)**: 403-417.
217. Nomenclature Committee of the International Union of Biochemistry y Molecular Biology (NC-IUBMB). 1992. **1st edition.** California, USA: Academic Press.
218. Norledge, B.V., A. M. Lamberir, A. R. Abagyan, A. Rottman, A. M. Fernández, V. V. Filimonov, M. G. Peter, y R. W. Wierenga. 2001. *Proteins.* **42(3)**: 383-389.



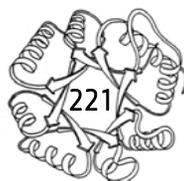
219. Offredi, F., F. Dubail, P. Kischel, K. Sarinski, A. S. Stern, C. Van de Weerd, J. C. Hoch, C. Prospero, J. M. Francois, S. L. Mayo, y J. A. Martial. 2003. *J Mol Biol.* **325(1)**: 163-174.
220. Olivares-Illana, V., H. Riveros-Rosas, N. Cabrera, M. Tuena de Gómez-Puyou, R. Pérez-Montfort, M. Costas, y A. Gómez-Puyou. 2017. *Proteins.* **85(7)**: 1190-1211.
221. Oshima, T., y K. Imahori. 1974. *Int J Syst Evol Microbiol.* **24(2)**: 102-112.
222. Pace, C. N., B. A. Shirley, M. McNutt, y K. Gajiwala. 1996. *FASEB J.* **10(1)**: 75-83.
223. Pace, N. C., y Ch. Tanford. 1968. *Biochemistry.* **7(3)**: 198-208.
224. Pace, C. N., F. Vadjos, L. Fee, G. Grimsley, y T. Gray. 1995. *Prot Sci.* **4(11)**: 2411-2423.
225. Pain, R. H. 2000. **2nd Edition**. New York, USA: Oxford University Press.
226. Pardo-Ávila, F. 2011. Tesis de grado. Licenciatura en Investigación Biomédica Básica. Universidad Nacional Autónoma de México. México, D.F.
227. Pareek, V., M. Samanta, N. V. Joshi, H. Balaram, M. R. Murthy, y P. Balaram. *Chembiochem.* **17(7)**: 620-629.
228. Park, S. H., H. S. Kim, M. S. Park, S. Moon, M. K. Song, H. S. Park, H. Hahn, S. J. Kim, E. Bae, H. J. Kim, y B. W. Han. 2015. *PLoS One.* **10(12)**: e0145331.
229. Pavankumar, T. L., A. K. Sinha, y M. K. Ray. 2010. *PLoS One.* **5(2)**: 9412-9431.
230. Pearson, T., P. Giffard, S. Beckstrom-Sternberg, R. Auerbach, H. Hornstra, A. Tuanyok, E. P. Price, M. B. Glass, B. Leadem, J. S. Beckstrom-Sternberg, G. J. Allan, J. T. Foster, D. M. Wagner, R. T. Okinaka, S. H. Sim, O. Pearson, Z. Wu, J. Chang, R. Kaul, A. R. Hoffmaster, T. S. Brettin, R. A. Robison, M. Mayo, J. E. Gee, P. Tan, B. J. Currie, y P. Keim. 2009. *BMC Biol.* **7(1)**: 78-92.
231. Peimbert, M., L. Domínguez-Domínguez, y D. A. Fernández-Velasco. 2008a. *Biochemistry.* **47(20)**: 5556-5564.
232. Peimbert, M., L. Domínguez-Ramírez, L. A. Téllez, L. M. Blancas, y D. A. Fernández-Velasco. 2008b., en: E. García-Hernández, y D. A. Fernández-Velasco. **1st edition**. India: Transworld Research Network. 193-217.
233. Pérez-Montfort, R., G. Garza-Ramos, G. H. Alcántara, H. Reyes-Vivas, X. G. Gao, E. Maldonado, M. T. de Gómez-Puyou, y A. Gómez-Puyou. 1999. *Biochemistry.* **38(13)**: 4114-4120.
234. Perl, D. C. Welker, T. Schindler, K. Schroder, M. A. Marahiel, R. Jaenicke, y F. X. Schmid. 1998. *Nat Struct Biol.* **5(1)**: 229-235.



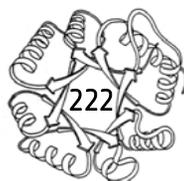
235. Pettersen, E. F., T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, y T. E. Ferrin. 2004. *J Comput Chem.* **25(13)**: 1605-1612.
236. Plaut, B., y J. R. Knowles. 1972. *Biochemistry.* **129(1)**: 311-320.
237. Plaza del Pino, I. M., B. Ibarra-Molero, y J. M. Sánchez-Ruiz. 2000. *Protein.* **40(8)**: 58-70.
238. Pompiliano, D.L., A. Peyman, y J. R. Knowles. 1990. *Biochemistry.* **29(13)**: 3186-3194.
239. Ponder, J. W., y F. M. Richards. *J Mol Biol.* **193(4)**: 775-791.
240. Prabhu, N. V., y K. A. Sharp. 2005. *Annu Rev Phys Chem.* **56(3)**: 521-548.
241. Privalov, P. L., 1980. *Pure Appl Chem.* **52(2)**: 479-497.
242. Privalov, P. L. 1990. *Crit Rev Biochem Mol Biol.* **25(3)**: 281-305.
243. Privalov, P. L., y G. I. Makhatadze. 1990. *J Mol Biol.* **213(2)**: 385-391.
244. Privalov, P. L., y N. N. Khechinashvili. 1974. *J Mol Biol.* **86(3)**: 665-684.
245. Puigbó, P., Y. I. Wolf, y E. V. Koonin. 2010. *Gen Biol Evol.* **2(1)**: 745-756.
246. Putman, S. J., A. F. W. Coulson, I. R. T. Farley, B. Riddleston, y J. R. Knowles. 1972. *Biochemistry.* **129(2)**: 301-310.
247. Quezada, A. G., A. J. Díaz-Salazar, N. Cabrera, R. Pérez-Montfort, Á. Piñeiro, y M. Costas. 2017. *Structure.* **25(1)**: 167-179.
248. Razvi, A., y J. M. Scholtz. 2006. *Prot Sci.* **15(2)**: 1569-1578.
249. Reardon, D., y G. K. Faber. 1995. *FASEBJ.* **9(2)**: 497-503.
250. Reche, P. 2013. SIAS, Universidad Complutense de Madrid, Madrid, España.
251. Rees, D. C., y A. D. Robertson. 2001. *Prot Sci.* **10(1)**: 1187-1194.
252. Regan, L., D. Caballero, M. R. Hinrichsen, A. Virrueta, D. M. Willian, C. S. O'Hern. 2015. *Biopolymers.* **104(4)**: 334-350.
253. Reid, K. L., H. M. Rodríguez, N. J. Hillier, y L. M. Gregoret. 2000. *Prot Sci.* **9(3)**: 1993-2000.
254. Reyes-Vivas, H., A. Diaz, J. Peón, G. Mendoza-Hernández, G. Hernández-Alcántara, I. de la Mora-de la Mora, S. Enríquez-Flores, L. Domínguez-Ramírez, y G. López-Velázquez. 2007. *J Mol Biol.* **365(3)**: 752-763.



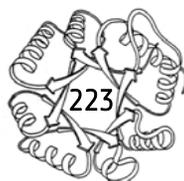
255. Reyes-Vivas, H., E. Martínez-Martínez, G. Mendoza-Hernández, G. López-Velázquez, R. Pérez-Monfort, M. T. de Gómez-Puyou, y A. Gómez-Puyou. 2001. *Proteins*. **48(3)**: 580-590.
256. Reyes-Vivas, H., G. Hernández-Alcántara, G. López-Velázquez, N. Cabrera, R. Pérez-Monfort, M. T. de Gómez-Puyou, y A. Gómez-Puyou. 2001. *Biochemistry*. **40(10)**: 3134-3140.
257. Rietveld, A. W., y S. T. Ferreira. 1996. *Biochemistry*. **35(24)**: 7743-7751.
258. Robertson, A.D., y K. P. Murphy. 1997. *Chem Rev*. **97(1)**: 1251-1267.
259. Roche, J., M. Dellarole, J. A. Caro, D. R. Norberto, A. E. Garcia, B. Garcia-Moreno, C. Roumestand, y C. A. Royer. 2013. *J Am Chem Soc*. **135(39)**: 14610-14618.
260. Rodríguez-Almazán, C., R. Arreola, D. Rodríguez-Larrea, B. Aguirre-López, M. T. de Gómez-Puyou, R. Pérez-Monfort, M. Costas, A. Gómez-Puyou, y A. Torres-Larios. 2008. *J Biol Chem*. **283(34)**: 23254-23263.
261. Rodríguez-Larrea, D., S. Minning, T. V. Borchert, y J. M. Sánchez-Ruiz. 2006. *J Mol Biol*. **360(8)**: 715-724.
262. Rodríguez-Romero, A., A. Hernández-Santoyo, L. Del Pozo-Yauner, A. Kornhauser, y D. A. Fernández-Velasco. 2002. *Structure. J Mol Biol*. **322(4)**: 669-675.
263. Rohl, C. A., C. E. Strauss, K. M. Misura, y D. Baker. 2004. *Methods Enzymol*. **383(1)**: 66-93.
264. Romero-Romero, S. 2013. Tesis de grado. Maestría en Ciencias Bioquímicas. Universidad Nacional Autónoma de México. México, D.F.
265. Romero-Romero, S. M. Costas, A. Rodríguez-Romero, y D. A. Fernández-Velasco. 2015. *Phys Chem Chem Phys*. **17(3)**: 20699-20714.
266. Ropson, I. J., J. A. Boyer, B. A. Schaeffer, y P. M. Dalessio. 2008. *Proteins*. **75(4)**: 799-806.
267. Rose, I. A., y E. O'Connell. 1961. *J Biol Chem*. **236(1)**: 3086-3092.
268. Röthlisberger, D., O. Khersonsky, A. M. Wollacott, L. Jiang, J. DeChancie, J. Betker, J. L. Gallaher, E. A. Althoff, A. Zanghellini, O. Dym, S. Albeck, K. N. Houk, D. S. Tawfik, y D. Baker. 2008. *Nature*. **453(7192)**: 190-195.
269. Rubin-Pitel, S. B., y H. Zhao. 2006. *Comb Chem High Throughput Screen*. **9(4)**: 247-257.
270. Rudolph, R., R. Siebendritt, y T. Kiefhaber. 1992. *Protein Sci*. **1(5)**: 654-66.
271. Ryabova, N. A., V. V. Marchekov, S. Y. Marchenkova, N. V. Kotova, y G. V. Semisotnov. 2013. *Biochemistry Mosc*. **78(5)**: 1405-1414.



272. Saab-Rincón, G., V. Rivelino-Juarez, J. Osuna, F. Sánchez, y J. Soberon. 2001. *Prot Engin.* **14(3)**: 149-155.
273. Salgado-Lugo, H., S. Lara-González., y L. G. Briebe. 2012. *To be published.*
274. Sánchez-Ruíz, J. M. 1992. *Biophys J.* **61(4)**: 921-935.
275. Sánchez-Ruíz, J. M. 2010. *Biophys Chem.* **148(1-3)**: 1-15.
276. Sánchez-Ruíz, J. M. 2011. *Annu Rev Phys Chem.* **62(2)**: 231-255.
277. Schindler, T., P. L. Graumann, D. Perl, S. Ma, F. X. Schmid, y M. A. Marahiel. 1999. *J Biol Chem.* **274(2)**: 3407-3413.
278. Schliebs, W., N. Thanki, R. Eritja, y R. K. Wierenga. 1996. *Prot Sci.* **5(2)**: 229-239.
279. Schliebs, W., N. Thanki, R. Jaenicke, y R. K. Wierenga. 1997. *Biochemistry.* **36(32)**: 9655-9962.
280. Schloss, P. D., y J. Handelsman. 2005. *Genome Biol.* **6(8)**: 229.
281. Scott, K. A., S. Batey, K. A. Hooton, y J. Clarke. 2004. *J Mol Biol.* **344(2)**: 195-205.
282. Segovia, L., y M. Peimbert. 2010. *Mensaje bioquímico XXXIV*: 135-141.
283. Seiki, E., N. Matsuda, S. Yokoyama, y T. Kigawa. 2008. *Anal Biochem.* **377(2)**: 156-161.
284. Serratos, I. N., G. Pérez-Hernández, G. Garza-Ramos, A. Hernández-Arana, E. González-Mondragón, y R. A. Zubillaga. 2011. *J Mol Biol.* **405(2)**: 158-172.
285. Sharma, P., y P. Guptasarma. 2015. *Biochem Biophys Res Commun.* **460(3)**: 753-758.
286. Shendure, J., R. D. Mitra, C. Varma, y G. M. Church. 2004. *Nat Rev Genet.* **5(5)**: 335-344.
287. Shi, Y., J. H. Liu, H. J. Zhang, y Y. Ding. 2008. *Prot Pept Lett.* **15(4)**: 365-370.
288. Shiao, D. F., R. Lymry, y J. Fahey. 1971. *J Amer Chem Soc.* **93(8)**: 2024-2035.
289. Shiraki, K., M. Kudou, S. Fujiwara, T. Imanaka, y M. Tagaki. 2002. *J Biochem.* **132(2)**: 591-595.
290. Shukla, A., y P. Guptasarma. 2004. *Proteins.* **55(3)**: 548-557.
291. Silva, J. L., y G. Weber. 1993. *Annu Rev Phys Chem.* **44(3)**: 89-113.
292. Silverman, H. A., R. Balakrishnan, y P. B. Harbury. 2001. *Proc Natl Acad Sci USA.* **98(6)**: 3092-3097.
293. Smadbeck, J., K. H. Chan, G. A. Khoury, B. Xue, R. C. Robinson, C. A. Hauser, y C. A. Floudas. 2014. *PLoS Comput Biol.* **10(7)**: e1003718.



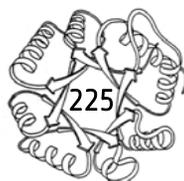
294. Smith, P. K., R. I. Krohn, G. T. Hermanson, A. K. Mallia, F. H. Gartner, M. D. Provenzano, E. K., Fujimoto, N. M. Goeke, B. J. Olson, y D. C. Klenk. 1985. *Anal Biochem.* **150**(1): 76-85.
295. Somero, G. 1995. *Annu Rev Physiol.* **57**(2): 43-68.
296. Spolar, R. S., y M. T. Record. 1994. *Science.* **263**(5148): 777-784.
297. Sreerama, N., y R. W. Woody. 2004. *Meth Enzymol.* **383**(4): 318-351.
298. Sterner, R., y B. Höcker. 2005. *Chem Rev.* **105**(11): 4038-4055.
299. Sturtevant, J. M. 1977. *Proc Natl Acad Sci USA.* **74**(4): 2236-2240.
300. Sun, A.Q., K. U. Yuksel, y R. W. Gracy. 1992a. *J Biol Chem.* **267**(3): 20168-20174.
301. Sun, A.Q., K. U. Yuksel, y R. W. Gracy. 1992b. *Arch Biochem Biophys.* **293**(2): 382: 390.
302. Sun, J., y N. S. Sampson. 1999. *Biochemistry.* **38**(35): 11474-11481.
303. Symersky, J., S. Li, M. Carson, y M. Luo. 2003. *Prot Struct Funct Genet.* **51**(3): 484-486.
304. Takahashi, K., y J. M. Sturtevant. 1981. *Biochemistry.* **20**(21): 6185-90
305. Tanaka, T., H. Kimura, M. Hayashi, Y. Fujiyoshi, K. Fukuhara, y H. Nakamura. 1994. *Protein Sci.* **3**(3): 419-427.
306. Taylor, W. R., V. Chelliah, S. M. Hollup, J. T. MacDonald, y I. Jonassen. 2009. *Structure.* **17**(8): 1244-1252.
307. Téllez L.A., L. M. Blancas-Mejía, E. Carrillo-Nava, G. Mendoza-Hernández, D. A. Cisneros, y D. A. Fernández-Velasco. 2008. *Biochemistry.* **47**(44): 11665-11673.
308. Thanki, N., J. Zeelen, M. Mathieu, R. Jaenicke, R. A. Abagyan, R. K. Wierenga, y W. Schliebs. 1997. *Prot Engin.* **10**(2): 159-167.
309. Trentham, D. R., C. H. McMurray, y C. I. Pogson. 1969. *Biochem J.* **114**(1): 19-24.
310. Tringe, S. G., y E. M. Rubin. 2005. *Nat Rev Genet.* **6**(11): 805-814.
311. Trivedi, V. D., B. Raman, C. M. Rao, y T. Ramakrishna. 1997. *FEBS Lett.* **418**(5): 363-366.
312. Vasserot, A. P., C. D. Dickinson, Y. Tang, W. D. Huse, K. S. Manchester, y J. D. Watkins. 2003. *Drug Discov Today.* **8**(3): 118-126.
313. Vázquez-Contreras, E., B. G. Sánchez-Rebollar, y M. E. Cháñez-Cárdenas. 2004. *Rev Soc Quim Mex.* **48**(2): 296-299.



314. Vázquez-Pérez, A. R., y D. A. Fernández-Velasco. 2007. *Biochemistry*. **46(29)**: 8624-8633.
315. Velanker, S. S., S. S. Ray, R. S. Gokhale, S. Suma, H. Balaram, P. Balaram, y M. R. N. Murphy. 1997. *Structure*. **5(6)**: 751-761.
316. Venter, J. C., K. Remington, J. F. Heidelberg, A. L. Halpern, D. Rusch, J. A. Eisen, D. Wu, I. Paulsen, K. E. Nelson, W. Nelson, D. E. Fouts, S. Levy, A. H. Knap, M. W. Lomas, K. Neelson, O. White, J. Peterson, J. Hoffman, R. Parsons, H. Baden-Tillson, C. Pfannkoch, Y. H. Rogers, y H. O. Smith. 2004. *Science*. **304(5667)**: 66-74.
317. Voet, D., J. Voet, y C. Pratt. 2006. 2nd edition. New York, USA: Wiley.
318. Voss, N. R., y M. Gerstein. 2010. *Nucleic Acids Res*. **38(4)**: W555-W562.
319. Walden, H., G. S. Bell, R. J. M. Rusell, B. Siebers, R. Hensel, y G. L. Taylor. 2001. *J Mol Biol*. **306(2)**: 745-754.
320. Walden, H., G. L. Taylor, E. Lorentzen, E. Pohl, H. Lilie, A. Schramm, T. Knura, K. Stubble, B. Tjaden, y R. Hensel. 2004. *J Mol Biol*. **342(3)**: 861-875.
321. Wallace, L. A., y R. Matthews. 2002. *J Mol Biol*. **315(2)**:193-211.
322. Waterhouse, A.M., J. B. Procter, D. M. A. Martin, y G. J. Barton. 2009. *Bioinformatics*. **25(9)**: 1189-1191.
323. Watve, M. G., R. Tickoo, M. M. Jog, y B. D. Bhole. 2001. *Arch microbiol*. **176(5)**: 386-390.
324. Wells, C.L., y T. D. Wilkins. 1996, en: S. Barron. 4th edition. Texas, USA: University of Texas Medical Branch. 28-31.
325. Wensley, B. G., M. Gärtner, W. X. Choo, S. Batey, y J. Clarke. 2009. *J Mol Biol*. **390(5)**: 1074-1085.
326. Whitmore, L., y B. A. Wallace. 2008. *Biopolymers*. **89(3)**: 392-400.
327. Wierenga, R. K. 2001. *FEBS Lett*. **492(3)**: 193-198.
328. Wierenga, R. K., K. H. Kalk, y W. G. J. Hol. 1987. *J Mol Biol*. **198(1)**: 109-121.
329. Wierenga, R. K., M. E. M. Noble, G. Vriend, y S. Nauche. 1991. *J Mol Biol*. **220(4)**: 995-1015.
330. Wierenga, R. K., M. E. Noble, y R. C. Davenport. 1992. *J Mol Biol*. **224(4)**: 1115-1126.
331. Willard, L., A. Ranjan, H. Zhang, H. Monzavi, R. F. Boyko, B. D. Sykes, y D. S. Wishart. 2003. *Nucleic Acids Res*. **31(3)**: 3316-3319.
332. Williams, J. C., J. P. Zeelen, G. Neubauer, G. Vriend, J. Backmann, P. A. Michels, A. M. Labeir, y R. K. Wierenga. 1999. *Prot Eng*. **12(3)**: 243-250.



333. Winn, M. D., C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin, y K. S. Wilson. 2011. *Acta Crystallogr Sect D*. **67(4)**: 235-242.
334. Wong, T. Y., y X. T. Yao. 1994. *Appl Env Microb*. **60(6)**: 2065-2068.
335. Woolfson, D. N., G. J. Bartlett, A. J. Burton, J. W. Heal, A. Niitsu, A. R. Thomson, y C. W. Wood. 2015. *Curr Opin Struct Biol*. **33(1)**: 16-26.
336. Yamamoto, H., y N. Kunishima. 2005. To be published.
337. Zabori, S., R. Rudolph, y R. Jaenicke. 1980. *Z Naturforsch*. **35(11-12)**: 999-1004.
338. Zaffagnini, M., L. Michelet, C. Sciabolini, N. Di Giacinto, S. Morisse, C. H. Marchand, P. Trost, S. Fermani, S. D. Lemaire. 2014. *Mol Plant*. **7(1)**: 101-120.
339. Zárate-Pérez, F., M. E. Chánez-Cárdenas, R. Arreola, A. Torres-Larios, y E. Vázquez-Contreras. 2009. *Biochem Biophys Commun*. **382(3)**: 626-630.
340. Zarrine-Afsar, A., S. M. Larson, y A. R. Davidson. 2005. *Curr Opin in Struct Biol*. **15(2)**: 42-49.
341. Zhang, Z., S. Sugio, E. A. Komives, K. D. Liu, J. R. Knowles, G. A. Petsko, y D. Ringe. 1994. *Biochemistry*. **33(10)**: 2830-2837.
342. Zhu, C., A. Anderson, Z. Zhang, S. Xiao, A. Schubart, H. B. Xiong, J. Imitola, S. Houry, X. X. Zheng, T. Strom, y V. Kuchroo. 2006. *Clin Immunol*. **119(2)**: S110.
343. Zomosa-Signoret, V., B. Aguirre-López, G. Hernández-Alcántara, R. Pérez-Montfort, M. T. de Gómez-Puyou, y A. Gómez-Puyou. 2007. *Proteins*. **67(1)**: 75-83.
344. Zomosa-Signoret, V., G. Hernández-Alcántara, H. Reyes-Vivas, E. Martínez-Martínez, G. Garza-Ramos, R. Pérez-Montfort, M. T. de Gómez-Puyou, y A. Gómez-Puyou. 2003. *Biochemistry*. **42(11)**: 3311-3318.



APÉNDICE A

Origen y evolución del dominio Bacteria

Las relaciones filogenéticas de los seres vivos son motivo de controversia y no hay un acuerdo general entre los diferentes autores. En la figura 67 se muestra un esquema de la historia evolutiva de los seres vivos basado en las ideas de Cavalier-Smith (Cavalier-Smith, 2006a-b). Según este autor, la raíz del árbol se situaría entre las bacterias gramnegativas, que serían los organismos más antiguos (existiendo desde hace 3, 500 millones de años), mientras que Archaea y Eukarya serían relativamente recientes (de hace sólo 900 millones años).

El siguiente cladograma muestra más en detalle los supertaxones a los que pertenecen las bacterias y señala además las relaciones entre los distintos grupos bacterianos, de acuerdo con las ideas de Cavalier-Smith (figura 67) (Cavalier-Smith, 2006).

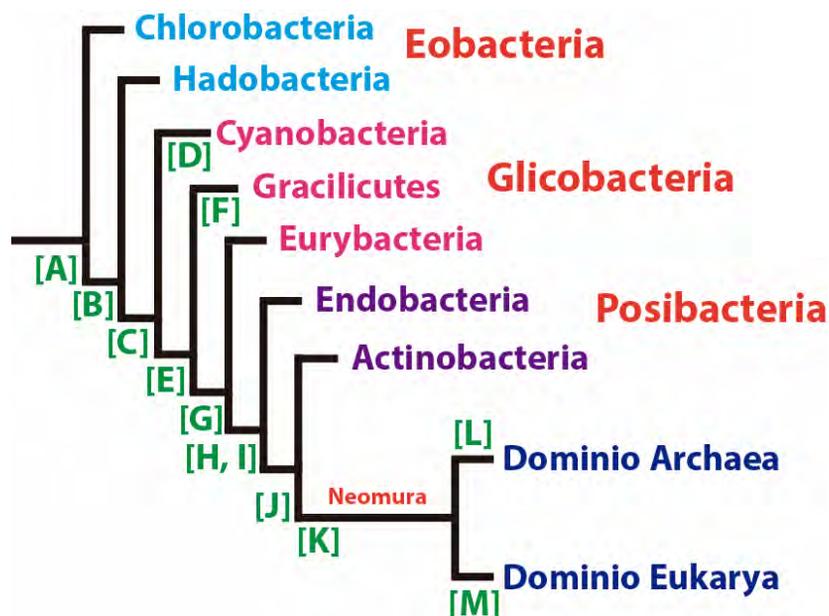


Figura 67. Cladograma que muestra la relación genética entre los distintos grupos bacterianos y su relación con los dominios Archaea y Eukarya. Se detallan los 3 supertaxones y los taxones incluidos en cada uno de ellos (modificado de: Cavalier-Smith, 2006a-b).



Eobacteria (Chlorobacteria + Hadobacteria): [A] Bacteria gramnegativa con pared de peptidoglicano; membrana externa carente de lipopolisacáridos; carencia de flagelos y endosporas; movilidad por deslizamiento bacterial; biología celular completamente desarrollada; presencia de citocromo c, clorosomas y fotosíntesis anoxigénica. [B] Surgimiento de Omp85 (un componente del mecanismo de inserción de proteínas en la membrana externa); desarrollo de cuatro nuevas catalasas; presencia de citocromo b; realización de fotosíntesis oxigénica, que podría haberse desarrollado en el antecesor común de Hadobacteria y Cyanobacteria, aunque los primeros son actualmente no fotosintéticos.

Glycobacteria (Cyanobacteria + Gracilicutes + Eurybacteria): [C] Surgimiento de la revolución glicobacteriana: bacteria gramnegativa con pared de peptidoglicano; membrana externa con inserción de moléculas complejas de lipopolisacáridos; presencia de hopanoides (agentes reforzantes de las membranas), ácido diaminopimélico, TolC y TonB en la pared de peptidoglicano. [D] Producción de ficobilisomas (estructuras de antena fotosintéticas presentes únicamente en cianobacterias y en ciertas algas). [E] Origen de los flagelos. [F] Cuatro inserciones en diversas proteínas: un aminoácido en Hsp60 y FtsZ y un dominio en las ARN polimerasas β y σ . [G] Formación de endosporas.

Posibacteria (Endobacteria + Actinobacteria): [H] Bacteria grampositiva: hipertrofia de la pared de peptidoglicano, pérdida de la membrana externa y origen de enzimas sortasas para enlazar las proteínas periplasmáticas a la pared celular y así evitar su pérdida. [I] Producción de glicerol 1-P deshidrogenasa (enzima que forma el glicerolfosfato de imagen especular al encontrado en los éter fosfolípidos bacterianos y eucariotas y característico de las arqueas). [J] Origen de los proteasomas; producción de fosfatidilinositol.

Neomura (Archaea + Eukarya): [K] Surgimiento de la revolución Neomura: el peptidoglicano y las lipoproteínas son sustituidos por glicoproteínas. [L] Producción de ADN girasa inversa (que induce un superenrollamiento positivo en el ADN para aumentar su estabilidad térmica); presencia

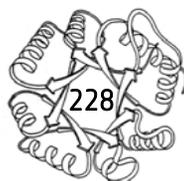


de lípidos éter isoprenoides en la membrana citoplasmática. [M] Surgimiento de la fagotrofia; adquisición de mitocondrias; cambio en la estructura general de la célula.

El esquema anterior se basa en la estructura celular de los distintos seres vivos enfatizando en la envoltura celular (membrana citoplasmática, pared celular y membrana externa). Según este criterio, el dominio Bacteria contiene organismos con dos tipos distintos de organización básica, gramnegativa y grampositiva, donde también podemos subdividir a las gramnegativas en dos subgrupos en función de la composición de la membrana externa (Pearson *et al.*, 2009).

Bajo esta clasificación, se agrupan las bacterias en tres grandes supertaxones: Eobacteria, Glycobacteria y Posibacteria. Las Eobacterias y Glycobacterias a su vez se clasifican en un supergrupo, llamado Negibacteria (bacterias gramnegativas) quienes presentan dos membranas lipídicas distintas, entre las que se localiza la pared celular, mientras que el resto de los organismos presentan una única membrana lipídica. La hipótesis del citoplasma fuera describe un posible modelo para la aparición de las dos membranas en estas primeras bacterias. Dentro de este grupo podemos distinguir dos subgrupos. Los subgrupos Eobacteria y Glycobacteria se distinguen por la composición de la membrana externa: los primeros presentan sólo simples fosfolípidos mientras que los segundos presentan inserción de moléculas complejas de lipopolisacáridos (Puigbó *et al.*, 2010).

El tercer supertaxón, las Posibacterias (bacterias grampositivas), presentan una única membrana y una pared de peptidoglicano mucho más gruesa. Se considera que las posibacterias proceden de las Negibacterias, y no al revés, porque las primeras presentan características moleculares y ultraestructurales más avanzadas. La pérdida de la membrana externa podría deberse a la hipertrofia de la pared celular, que aumenta la resistencia de estos organismos, pero impide la transferencia de lípidos para formar la membrana externa. Estos organismos fueron probablemente los primeros que colonizaron el suelo (Dagan *et al.*, 2010).

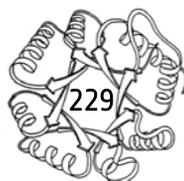


Archaea y Eukarya eventualmente tuvieron como origen una Posibacteria, a través de un organismo Neomura que sustituyó la pared celular de peptidoglicano por otra de glicoproteína. A continuación, y casi inmediatamente, las arqueas se adaptaron a ambientes calientes y ácidos, reemplazando los lípidos acilo éster de las bacterias por lípidos prenil éter, y usando las glicoproteínas como una nueva pared rígida. Los eucariontes, en cambio, usaron la nueva superficie de proteínas como una capa flexible para desarrollar la fagocitosis, lo que los llevó en última instancia, a profundos cambios en la estructura de la célula (Cavalier-Smith, 2006a; Lake *et al.*, 2009).

Filogenia y taxonomía del dominio Bacteria

Los filia bacterianos que se incluyen en el esquema evolutivo actualmente aceptado del dominio Bacteria, son los siguientes (Cavalier-Smith, 2006a-b; Cicarelli *et al.*, 2006):

- *Supertaxón Eobacteria*
 - **Taxón Chlorobacteria**
 - *Filo Chloroflexi* (bacterias verdes no del azufre). Pequeño filo de bacterias que realizan la fotosíntesis anoxigénica mediante bacterioclorofila, por lo que no producen oxígeno. Su vía de fijación del carbono también difiere de la de otras bacterias fotosintéticas. Son aerobias facultativas y típicamente filamentosas.
 - *Filo Thermomicrobia*. Pequeño filo de termófilos quimioheterótrofos.



○ **Taxón Hadobacteria**

- *Filo Deinococcus-Thermus*. Grupo de quimiorganotrofos extremófilos altamente resistentes. Unas especies soportan el calor y el frío extremo, mientras que otras son resistentes a la radiación y a un gran número de sustancias tóxicas.

• Supertaxón Glycobacteria

- **Taxón-Filo Cyanobacteria** (algas verde-azuladas). El grupo más importante de bacterias fotosintéticas. Presentan clorofila y realizan la fotosíntesis oxigénica. Son unicelulares o coloniales filamentosas.

○ **Taxón Gracilicutes**

- *Filo Spirochaetes*. Bacterias quimioheterótrofas con forma alargada típicamente enrollada en espiral que se desplazan mediante rotación. Muchas producen enfermedades.
- *Filo Chlorobi* (bacterias verdes del azufre). Es un pequeño filo de bacterias fototrófas mediante bacterioclorofila y anaerobias obligadas. Una especie es termófila y vive en fuentes hidrotermales.
- *Filo Bacteroidetes*. Un extenso filo de bacterias con amplia distribución en el medio ambiente. Es un grupo heterogéneo que incluye aerobios obligados o anaerobios obligados, comensales, parásitos y formas de vida libre.
- *Filo Fibrobacteres*. Pequeño filo que incluye muchas de las bacterias estomacales que permiten la degradación de la celulosa en los rumiantes.



- *Filo Proteobacteria*. Es un grupo muy diverso y el segundo más extenso entre las bacterias. Casi todas son heterótrofas y muchas causantes de enfermedades, pero los rizobios son simbioses al realizar la fijación de nitrógeno y las bacterias púrpuras son fototrofas con bacterioclorofila.
- *Filo Aquificae*. Un pequeño grupo de bacterias quimiolitotrofas, termófilas o hipertermófilas. Se les encuentra en manantiales calientes, pozos sulfurosos y fuentes hidrotermales oceánicas.
- *Filo Deferribacteres*. Grupo de bacterias acuáticas anaerobias.
- *Filo Chrysiogenetes*. Comprende una sola especie de quimiolitotrofo. Tiene una bioquímica y una forma de vida únicas: en vez de respirar oxígeno, respira arseniato.
- *Filo Acidobacteria*. Pequeño filo de bacterias acidófilas comunes en el suelo. Incluye una bacteria fototrofa usando bacterioclorofila.
- *Filo Planctomycetes*. Bacterias principalmente acuáticas aerobias encontradas en agua dulce, salobre y marina. Su ciclo biológico implica la alternancia entre células sésiles y flageladas. Se reproducen por gemación.
- *Filo Chlamydiae*. Un pequeño grupo de parásitos intracelulares obligados de las células eucariotas.
- *Filo Lentisphaerae*. Pequeño grupo de bacterias recientemente descubiertas en aguas marinas y hábitats terrestres anaerobios.
- *Filo Verrucomicrobia*. Comprende bacterias terrestres, acuáticas y algunas asociadas con huéspedes eucariotas.



○ **Taxón Eurybacteria**

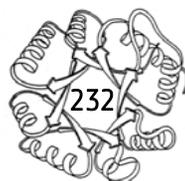
- *Filo Fusobacteria*. Comprende un sólo género de bacterias heterótrofas anaerobias causantes de infecciones en humanos. Constituyen uno de los principales tipos de flora del aparato digestivo.
- *Filo Thermotogae*. Un filo de hipertermófilos, anaerobios obligados, heterótrofos fermentativos.

• Supertaxón Posibacteria

○ **Taxón Endobacteria**

- *Filo Dictyoglomi*. Comprende una sola especie de hipertermófilo, quimioorganotrofo y aerobio.
- *Filo Firmicutes*. Es el grupo más extenso y comprende a las bacterias Gram positivas con contenido GC bajo. Se encuentran en diversos hábitats, incluyendo algunos patógenos notables. Una de las familias, Heliobacteria, obtiene su energía a través de la fotosíntesis.

- **Taxón-filo Actinobacteria**. Un extenso filo de bacterias Gram positivas de contenido GC alto. Son comunes en el suelo aunque algunas habitan en plantas y animales, incluyendo algunos patógenos.



APÉNDICE B

Algunas características de las especies seleccionadas para analizar su TIM se describen a continuación:

1. *Deinococcus radiodurans*

Deinococcus radiodurans (antes *Micrococcus radiodurans*) es un microorganismo quimiorganotrofo extremófilo altamente resistente a la radiación, sólo es superado en resistencia a la radiación por la arquea *Thermococcus gammatolerans*.

Esta bacteria pertenece al supertaxón Eobacteria, filo: Deinococcus-Thermus, orden: Deinococcales, familia: Deinococcaceae y género: *Deinococcus* (Brenner *et al.*, 2005). *D. radiodurans* está relacionada directamente con miembros de Chloroflexi, que en conjunto a este filo, se consideran las bacterias más primitivas (Cavalier-Smith, 2006a).

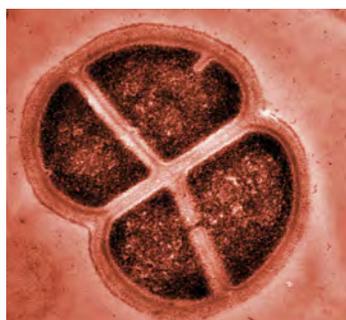
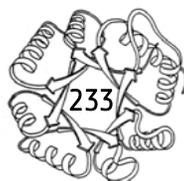


Figura 68. Micrografía de la bacteria *Deinococcus radiodurans*. Técnica: microscopía electrónica de transmisión con modificación de color por computadora. Fuente: <http://www.emsl.pnl.gov>

D. radiodurans tiñe como grampositiva, pero la composición de la pared (y el número de membranas) es del tipo de gramnegativas; la razón de esta tinción “anormal” se debe a que la



composición química de su pared es distinta, lo que provoca la captación del primer colorante, y aunque la capa de peptidoglicano es delgada, su composición permite la fijación de dicho colorante y su observación como grampositiva, a pesar de que morfológicamente parezca más una bacteria gramnegativa (figura 68).

Algunas propiedades estructurales y fisiológicas de este microorganismo son: su membrana externa carece de lipopolisacáridos, posee la proteína Omp85 (un componente clave en la evolución bacteriana, ya que indica que esta bacteria posee el mecanismo de inserción de proteínas en la membrana externa. La bacteria carece de flagelos y endosporas y su movilidad es realizada por deslizamiento (tipo gliding).

D. radiodurans presenta citocromo *c*, clorosomas y realiza fotosíntesis anoxigénica. Otra característica relevante del microorganismo es su alta resistencia a la radiación: mientras que una dosis de 10 Gy es suficiente para matar a un ser humano, y una dosis de 60 Gy es capaz de matar todas células en una colonia de *E. coli*, *D. radiodurans* puede resistir una dosis instantánea de hasta 5000 Gy sin pérdida de viabilidad, y dosis de hasta 15000 Gy con un 37% de pérdida de viabilidad. Además, puede sobrevivir en condiciones de calor, frío, deshidratación, vacío y ácido (Mattimore y Battista, 1996). La bacteria se ha utilizado en biorremediación para consumir y digerir disolventes y metales pesados, incluso en espacios altamente radiactivos (Brim *et al.*, 2000).



2. *Nostoc punctiforme*

Es una bacteria gramnegativa cocoide de color verdoso; forma colonias esféricas compuestas de filamentos (tricomas). Pertenece al supertaxón Glicobacteria, filo: Cianobacteria, orden: Nostocales, familia: Nostocaceae y género: *Nostoc* (Brenner *et al.*, 2005). Sus colonias son esféricas de superficie lisa sobre el agar, de aspecto mucilaginoso y de color desde verde oliva a verde oscuro casi negro. Al microscopio se observan masas densas de filamentos envueltos en una vaina. Células casi esféricas con un diámetro medio $4.3\ \mu\text{m}$, con heterocistos sobresaliendo de esas masas, semiesféricos y con un diámetro de $4.2\ \mu\text{m}$ y una longitud de $3\ \mu\text{m}$ (figura 69) (Abbot *et al.*, 1989).

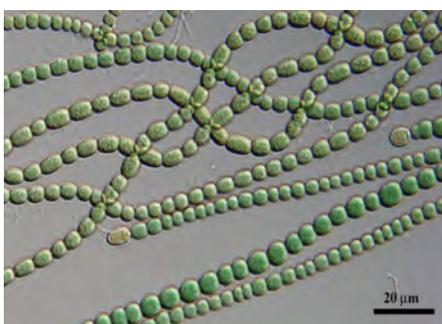


Figura 69. Micrografía de un filamento de la bacteria *Nostoc punctiforme*. Técnica: microscopía de contraste de fases. Fuente: <http://www.ccala.butbn.cas.cz>

Algunas características de esta bacteria son: contiene clorofilas tipo a y b, así como ficobilisomas, es capaz de realizar fotosíntesis oxigénica, lo cual permitió la formación de una atmósfera oxidante. Presenta además los heterocistos, que contienen la maquinaria de fijación del nitrógeno. Evolutivamente, son los primeros organismos después de la revolución glicobacteriana que dio surgimiento a las bacterias gramnegativas con pared de peptidoglicano, además se cree que dio origen por endosimbiosis a los plastos. Su membrana externa presenta inserciones de moléculas complejas de lipopolisacáridos además de hopanoides (agentes reforzantes de las membranas) (Cavalier-Smith, 2006a-b).



3. *Gemmata obscuriglobus*

Es una bacteria perteneciente al supertaxón Glicobacteria, filo: Planctomycetes, orden: Planctomycetacia, familia: Planctomycetaceae y género: Gemmata. Tienen forma de ovoide que se une a una especie de tallo y se reproduce por gemación. El ciclo biológico implica la alternancia entre células sésiles y flageladas (Brenner *et al.*, 2005).

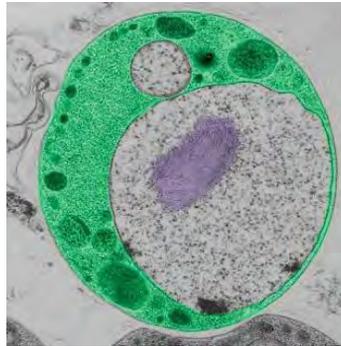


Figura 70. Micrografía de la bacteria *Gemmata obscuriglobus*. Técnica: microscopía electrónica de transmisión con modificación de color por computadora –en verde se muestran algunos compartimentos asociados a membranas internas y en morado se indica el ADN bacteriano-. Fuente: <http://www.emsl.pnl.gov>

Entre sus características metabólicas y estructurales relevantes se encuentran: estructuralmente contiene una región del ADN que está separado del resto de la célula por dos membranas nucleares (figura 70), además hay otros dos compartimientos separados por membranas: riboplasma (o pirelulosoma) que contiene ribosomas con las proteínas asociadas y parifoplasma que no contiene ribosomas. Carece de mureína en su pared celular, en su lugar, las paredes se componen de una glicoproteína rica en glutamato y lípidos llamados "ladderanos". Algunos genes esenciales de la bacteria no se organizan en operones, posee genes que son similares a los que poseen los organismos eucariontes como la integrina alfa-V. Todas estas características la hacen interesante en cuanto a su relevancia evolutiva (Fuerst y Weeb, 1991).



4. *Azotobacter vinelandii*

Azotobacter vinelandii es una bacteria gramnegativa quimiorganotrófica. Pertenece al supertaxón Glicobacteria, filo: Proteobacteria, clase: Gammaproteobacteria, orden: Pseudomonadales, familia: Pseudomonadaceae, género: *Azotobacter* (Brenner *et al.*, 2005).

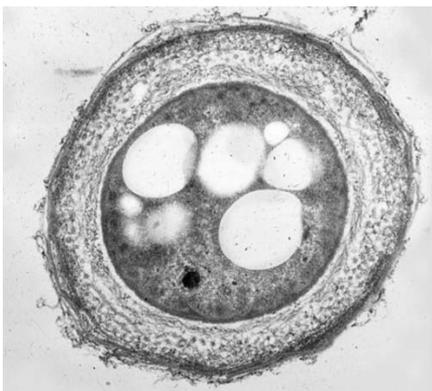
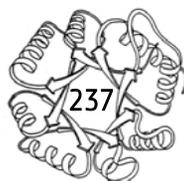


Figura 71. Micrografía de la bacteria *Azotobacter vinelandii* en estado quiste. Técnica: microscopía electrónica de transmisión. Fuente: <http://www.emsl.pnl.gov>

Se reproduce por fisión binaria, vive en suelos y en sistemas de agua dulce y salada, presenta células ovoides y grandes de 1,5 a 2 μm de diámetro, pleomórficas, variando su morfología desde bacilos hasta cocos. *A. vinelandii* es poliploide (posee varias copias de su cromosoma) (Wong y Yao, 1994).

Las capacidades metabólicas y genéticas por las que *A. vinelandii* ha sido y es objeto de estudio son: fija nitrógeno atmosférico en presencia de oxígeno por tres sistemas diferentes de nitrogenasas. Posee mecanismos de protección de la nitrogenasa. Posee una alta capacidad respiratoria que en condiciones diazotróficas o de fijación de nitrógeno es hasta 10 veces más alta que la de *E. coli*. Produce dos polímeros de uso industrial: el polisacárido extracelular alginato y el poliéster intracelular polihidroxibutirato. Sufre un proceso de diferenciación morfológica para formar quistes resistentes a la desecación (figura 71). Evolutivamente es de las primeras bacterias con flagelos y proteínas transportadoras en membrana (<http://www.azotobacter.org/>).



5. *Clostridium perfringens*

Clostridium perfringens es una bacteria anaeróbica grampositiva, inmóvil y formadora de esporas (figura 72) que se encuentra en los intestinos de los seres humanos y de varios animales homeotermos, en el suelo, en el agua, en los alimentos (sobre todo en las carnes que no están bien cocinadas), entre otros. Pertenece al supertaxón Posibacteria, filo: Firmicutes, clase: Clostridia, orden: Costridiales, familia: Clostridiaceae y género: Clostridium (Brenner *et al.*, 2005).

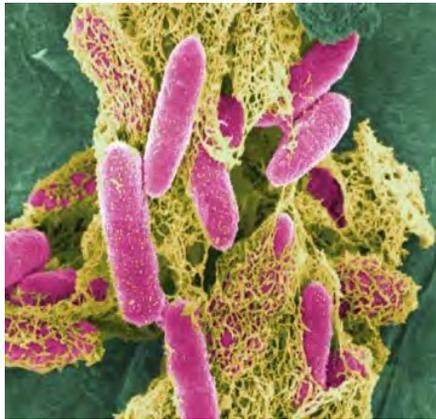


Figura 72. Micrografía de la bacteria *Clostridium perfringens*. Técnica: microscopía electrónica de barrido con modificación de color por computadora –en rosa se indican los bacilos bacterianos y en amarillo se indica el biofilm producido por las bacterias-. Fuente: www.emsl.pnl.gov

Esta bacteria produce toxinas que pueden causar enfermedades como la enteritis necrótica o la gangrena gaseosa. En la gangrena gaseosa, *Clostridium perfringens* provoca destrucción en los tejidos infectados si persiste. Esto es provocado por la liberación de exoenzimas específicos que atacan a las moléculas constituyentes de los tejidos de animales: fosfolipasas, hemolisinas, colagenasas, proteasas, que provocan la putrefacción del tejido acompañada de una producción de gas, y de ahí su nombre ("gaseosa") (Brook, 2007). Evolutivamente es de las primeras bacterias donde existe hipertrofia de la pared de peptidoglicano y pérdida de la membrana externa, generando al grupo de las grampositivas, además forma endosporas (Wells y Wilkins, 1996).

6. *Streptomyces coelicolor*

Streptomyces coelicolor es una bacteria perteneciente al género más extenso de actinobacterias. Se encuentra predominantemente en suelos y en la vegetación descompuesta. Se distinguen por el olor a "tierra húmeda" que desprenden, resultado de la producción de un metabolito volátil, la geosmina. Pertenece al supertaxón Posibacteria, filo: Actinobacteria, orden: Actinomycetales, suborden: Streptomycineae, familia: Streptomycetaceae y género: *Streptomyces* (Brenner *et al.*, 2005). El genoma completo de *S. coelicolor* A3, fue publicado en 2002. En esa fecha era el genoma bacteriano más grande conocido (Bentley *et al.*, 2002).

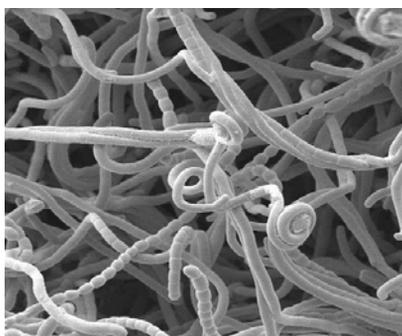
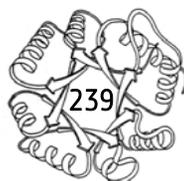


Figura 73. Micrografía de la bacteria *Streptomyces coelicolor*. Técnica: microscopía electrónica de barrido –se puede identificar la morfología cocácea característica de la bacteria y los filamentos producidos por las mismas-. Fuente: <http://www.jic.ac.uk>

Esta bacteria presenta características de coco grampositivo, aerobio, no móvil, no productor de esporas. Su crecimiento es en colonias filamentosas y ciclo celular lento (figura 73; Brenner *et al.*, 2005). Estructuralmente su ADN tiene contenido de G-C alto y un único cromosoma lineal (Bentley *et al.*, 2002). Evolutivamente, son las primeras bacterias que presentan proteosomas y síntesis de fosfatidilinositol. Pertenece al grupo de bacterias que son ancestros del organismo Neomura, origen de los eucariontes y arqueas. Biotecnológicamente se utiliza para la producción de metabolitos secundarios como antibióticos e inmunosupresores, además se ha utilizado como un organismo para la expresión de proteínas en sistemas heterólogos (Watve *et al.*, 2001).



7. *Moritella marina*

Moritella marina, antes llamada *Vibrio marinus*, es un anaeróbico facultativo gramnegativo con una morfología de bacilos curvados o rectos. La bacteria es móvil por flagelos polares, es convexa, opaca, de color crema, halófila y no pigmentada (figura 74). Su crecimiento óptimo es a bajas temperaturas, inferiores a los 17 grados (Felter *et al.*, 1969). El nombre *Vibrio* deriva de los descubrimientos del médico italiano Filippo Pacini, que en 1854 aisló microorganismos del intestino de pacientes con cólera, los que llamó "vibriones", por su motilidad.

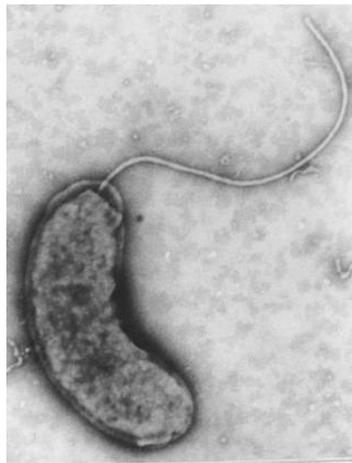
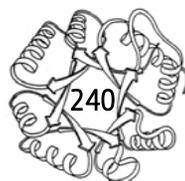


Figura 74. Micrografía de la bacteria *Moritella marina*. Se puede notar la presencia de su flagelo polar, lo cual le permite movilidad. Fuente: <http://deskuennis.nic>

Bioquímicamente se caracterizan por dar positivo en las pruebas de la catalasa y de la oxidasa, también dan negativo en la adenina dihidrolasa, y positivo en la ornitina descarboxilasa. *Vibrio cholerae* concretamente es sacarosa y manitol positivo y nitrato reductasa positivo (Brenner *et al.*, 2005). Su metabolismo es fermentativo; pueden fermentar, entre otros sustratos, la glucosa y además no forma esporas en su ciclo de vida. Soporta bien los medios alcalinos, así como las concentraciones salinas (Brook, 2007).



8. *Thermus thermophilus*

Thermus thermophilus es una eubacteria gramnegativa utilizada en una gran variedad de aplicaciones en biotecnología, incluyéndolo como organismo modelo en la manipulación genética, genómica estructural y biología de sistemas. La bacteria es un termófilo extremo, con una temperatura de crecimiento óptimo a una temperatura de unos 65 °C. *Thermus thermophilus* fue aislado originalmente en un ambiente de fumarolas hidrotermales en Japón por Tairo Oshima y Kazutomo Imahori (figura 75; Oshima e Imahori, 1974). También se ha visto que el organismo es importante en la degradación de materia orgánica en la fase termogénica del compostaje (Beffa *et al.*, 1996).

La maquinaria biológica de estos organismos tiene una estabilidad superior a otros organismos debido al medio ambiente en que tienen que vivir. En general, los termófilos son anaerobios que pueden vivir en el ambiente caliente con baja solubilidad de oxígeno debido a la temperatura, con la excepción de *Thermus*, que son quimioorganotróficos aeróbicos (Brenner *et al.*, 2005). Una de las aplicaciones de este microorganismo es que su rTth ADN polimerasa es una polimerasa recombinante de ADN termoestable con una actividad óptima a 70-80 °C, utilizada en algunas aplicaciones de la PCR. La enzima posee una eficiente actividad como transcriptasa inversa en presencia de manganeso (Myers y Gelfand, 1991).



Figura 75. Micrografía de la bacteria *Thermus thermophilus*. Se puede notar la forma alargada de sus células, característica de esta especie. Fuente: <http://www.thermus.org>



APÉNDICE C

A continuación se presentan las secuencias de las proteínas estudiadas en el presente trabajo:

1. Barriles TIM naturales:

>Azotobacter vinelandii

MKRKQVIGNWKMNGDLAANRRLLDELLPSLAALDGVELAVCPPFPYLAQVAAALLEGSGVALGAQNLNVAAKGAFTEVSAGMLG
ELGCRYVLVGHSERRSLYGECDETVAEKFAVALRAGLVPVLCVGETLAQRROGRTEALVSAQLQAVLRRSGIEGLAAGVIAYEPVWAI
NGNETASPEQAQAVHRHIRDLLAQHDPARGVATPILYGGSVKADNAAALFAQDPDIDGGGIGGASLEAAAFQAICQAAAR

>Clostridium perfringens

MRTPIIAGNWKMHYTIIDEAVKLVEELKPLVKDAKCEVWCPTFVCLDAVKKAVEGTNIKVGAQNMHFEEKGAFTEIAPRMLEAMNI
DYVIIGHSERREYFNETDETCNKKVKAFAHNLTPILCCGETLEQRENGTTNDVIKAQITADLEGLTKEQAEKVVIAYEPIWAIGTGKTA
TSDQANETIAAIRAMVAEMFGQEVADKVRIOYGGSVKPNTIAEQMAKSDIDGALVGGASLVAADFAQIVNY

>Deinococcus radiodurans

MQTLALLNWKMNKTPTEARSWAEELTTKYAPAEGVDLAVLAPALDLSALAANLPAGIAFGGQDVSAHESGAYTGEISAAMLKDAG
ASCVVVGHSERRYHDESDATVAAKARQAQANGLLPIVCVGENLDVREGERHPQTLAQLRGSLEGVADVVWAYEPVWAIGTGKT
ATADDAEELAAAIRGALREQYGARAEIRVLYGGSVKPNIAEICGKPNVNGALVGGASLKVDPDLGMLDALR

>Entamoeba histolytica

MSARKFVGGNWKCNGLASIELTKGVAASVDAELAKKVEVIVGVPFIYIPKVOQJILAGEANGANILVSAENAWTKSGAYTGEVHV
GMLVDCQVPYVILGHSERRQIFHESNEQVAEKVVAIDAGLKVACIGETEARIANQTEEVAAQLKAINNAISKEAWKNIILAYEPV
WAIGTGKTATPDQAQEVHQYIRKWMTENISKEVAEATRIQYGGSVNPANCNELAKKADIDGFLVGGASLDAAKFKTIINSVSEKF

>Escherichia coli

MRHPLVMGNWKLNGSRHMVHELVSNLKELAGVAGCAVAIAPPEMYIDMAKREAGSHIMLGAQNVLDLNLGAFTEGTSAAAML
KDIGAQYIIIGHSERRYHKSDELIAKKFVAVLKEQGLTPVLCIGETEAEENEAGKTEEVCAEQIDAVLKTQGAAAFEGAVIAYEPVWAIG
TGKSATPAQAQAVHKFIRDHIAKVDANIAEQVIIQYGGSVNASNAELFAQDPDIDGALVGGASLKADAFVIVKAAEAAKQA

>Gallus gallus

MAPRKFFVGGNWKMNKSLGELIHTLNGAKLSADTEVWCGAPSIYLDFAEQKLDKIGVAAQNCYKVPKGAFTEISPAKIDIG
AAWVILGHSERRHVFGEDELIGQKVAHALAEGLVIAICIGEKLDEREAGITEKVVFEQTKAIADNVKDWKSVLAYEPVWAIGTGKT
ATPQQAQEVHEKLRGWLKSHVSDAVAQSTRIYGGSVTGGNCKELASQHDVDGFLVGGASLKPEFVDIINAKH



>**Gemmataobscuriglobus**

MPTRKKFVAGNWKMNTTLAEAKALGA AVAKGVTDDRVTAVFPPYPWLTAVGEVLK GSPVALGAQDVSSEKKGAFTEVSPAML
LETGCKYALIGHSEERRHIIGSETFINHKVHTALEEGLSVLWCMGETLAERERGLQERVFORQVYAACAGLTDEQFGRIVIAYEPVWAI
GTGKVATPEQAQEAHAFVRSKLRLLYGDKIADSTPIVYGGSVTPDNTVGLMSQPDVDGALVGGASLKADSF LAIVKAAG

>**Geobacillusstearotherophilus**

MRKPIIAGNWKMHKTLAEAVQFVEDVKGHVPPADEVISVWCAPFLFLDRLVQAADGTDLKGIAQTMHFADQ GAYTGEVSPVMLKDL
GVTYVILGHSERRQMF AETDETVNKKVLA AFTRGLIPIICCGESLEEREAGQTN AVVASQVEKALAGLTPEQVKQAVIAYEPIWAIGTG
KSSTPEDANSVCGHIRSVSRLFGPEAAEAI RIQYGGSVKPDNIRDFLAQQQIDGPLVGGASLEPASFLQLVEAGRHE

>**Giardiaintestinalis**

MPARRPFIGGNFKNGSLDFIKSHVAAIAAHKIPDSVDWVIAPSAVHLSTAI AANTSQKLRIAAQNVYLEGNGAWTGETSVEMLQD
MGLKHVIVGHSEERRIMGETDEQS AKKAKRALEKGMTVIFCVGETLDERKANRTMEVNIAQLEALGKELGESKMLWKEWVIAYEPV
WSIGTG VATPEQAEVHVGLR KWFAEKVCAEGAQHIRIYGG SANGSNCEKLGQCPNIDGFLVGGASLKPEFMTMIDILTKTRT

>**Helicobacterpylori**

MTKIAMANFKSAMPFIKSHAYLKELEKTLKPQHFD RVFVFPDFGLLPNSFLHFTLG VQNAVPRDCGAFTEITSKHLEELKIHTLLIG
HSERRTLKESPSFLKEKFDFFKSNFKIVYCIGEELTTREKGFKAVKEFLSEQLENI DLNYPNLWVAYEPIWAIGTKKSASLEDIYLTHG
FLKQILNQKTPLLYGGSVNTQNAKEILGIDSVDGLLIGSASWELENFKTIISFL

>**Homo sapiens**

MAPSRKFFVGGNWKMNGR KQSLGELIGTLNAAKVPADTEVVCAPPTAYIDFARQKLDPKI AVAAQNCYKVTNGAFTGEISPGMIKD
CGATWVVLGHSERRHVFGESEDELIGQV AHALAEGLVIAICIGEKLDEREAGITEKVVFEQTKVIADNVKDW SKVWLAYEPVWAIGTG
KTATPQQAQEVHEKLRGWLKSNVSDAVAQSTRIIYGGSVTGATCKELASQPDVDGFLVGGASLKPEFVDIINAKQ

>**Leishmaniamexicana**

MSAKPQPIAAANWKMNGT TASIEKLVQVFNEHTISHDVQCWVAPT FVHIPLVQAKLRNPKYVISAENAI AKSGAFTGEVSMPIKDIG
VHWVILGHSERRTYGETDEIVAQK VSEACKQGF MVIACIGETLQOREANQTAKVLSQTS AIAAKLTKD AWNQVWLAYEPVWAIGT
GKVATPEQAQEVHLLLRKWVSENI GTDVA AKLRILYGGSVNAANATLYAKPDINGFLVGGASLKPEFRDIIDATR

>**Moritellamarina**

MRHPVVMGNWKLNGSKEMVVDLLNGLNAELEGVTGVDVAVAPPALFVDLAERTLTEAGS AIIILGAQNTDLNNSGAFTGDMSPAM
LKEFGATHIIIIGHSE RREYHAESDEFVAKKFALKENGLTPVLCIGESDAQNEAGETMAV CARQLDAVINTQGVEALEGAI IAYEPIWA
IGTGKAATAEDAQR IHAQIRAHIAEKSEAVAKNVVIQYGGSVKPENAAAYFAQPDIDGALVGGAA LDAKSF AAIKAAAEAKA

>**Mycobacteriumtuberculosis**

MSRKPLIAGNWKMNLNHYEIALVQKIAFSLPDKYDRVDVAVIP PFTDLRSVQTLVDGDKLR LTYGAQDLSPHDSGAYTGDVSGAF
LAKLGC SYVVGHSE RRTYHNEDDALVA AKAATALKHGLTPIVCIGEHLDVREAGNHVAH NIEQLRGS LVGLLAEQIGSVVIAYEPVW
AIGTGRVASAADAQEVCAAIRKELASLASPRIADTVRVLYGGSVNAKNVGDIVAQDDVDGGLVGGASLDGEHFATLAAI AAGGGLP

>**Nostocpunctiforme**

MRKVIAGNWKMFKTQ AETQEF LQGFPHLEETPQGREVILCPPFTDLSVLSKTLHGSLIQLGAQNIHWEEFGAYTGEISGPMLTESG
VRFVIVGHSE RRYQFGETDATVNLRLRTAQRFG LTPILCVGETKQQRDAGETESLIALQLDKGLVDIDQNNLVIAYEPIWAIGTGETCE
AVEANRIIGLIRSQLSNPNVSIQYGGSVKPNNIDEI MAQPEIDGVLVGGASLEPESFARIVNFHLV



>**Oryctolagus cuniculus**

MAPSRKFFVGGNWKMNKRKKNLGE LITTLNAAKVPADTEVVCAPPTAYIDFARQKLDPKIAVAAQNCYKVTNGAFTGEISPGMIKD
CGATWVVLGHSERRRHVFGESDELIGQVAHALSEGLGVIACIGEKLDEREAGITEKVVFEQTKVIADNVKDWKWLAYEPVWAI GTG
KTATPQQAQEVHEKLRGWLKSNVSDAVAQSTRIIYGGSVTGATCKELASQPDVDGFLVGGASLKPEFVDIINAKQ

>**Plasmodium falciparum**

MARKYFVAANWKCNGTLESIKSLTNSFNLD FDFPSKLDVWVFPVSVHYDHTRKLQSKFSTGIQNVSKFNGSYTGEVSAEIAKDLNI
EYVIIGHFERRKYFHETDEDEVREKLOASLKNLKA VVCFGESLEQREQNKTEVITKQVKAFVDLIDNFDNVILAYEPLWAI GTGKTATP
EQAQLVHKEIRKIVKDTCGEKQANQIRILYGGSVNTENCSSLIQOQEDIDGFLVGNASLKESFVDIISAM

>**Saccharomyces cerevisiae**

MARTFFVGGNFKLNKSGKQSIKEIVERLNTASIPENVEVICPPATYLDYSVSLVKKPQVTVGAQNAYLKASGAF TGENSVDQIKDVG
KAWILGHSERRSYFHEDDKFIADKTKFALGQGVG VILCIGETLEEKKAGKTLDVVERQLNAVLEEVKDWNTN VVAYEPVWAI GTGLAA
TPEDAQDIHASIRKFLASKLGDKAASELRILYGG SANGSNAVTFKDKADVDGFLVGGASLKPEFVDIINSRN

>**Streptomyces coelicolor**

MTRTRPLMAGNWKMNLNHLEAIAHVQKLAFA LADKDYDAVEVAVLAPFTDLRSVQTLVDGDKLKIKYGAQDISAHDGGAYTGEIS
GPMLAKLKCTYVAVGHSERRQYHAETDEIVNAK VKAAYKHGLTPILCVGEELDVREAGNHVEHTLAQVEGGLKDLAAEQAESVVIAY
EPVWAI GTGKVCGADDAQEVCAAIRGKLAELYSQELADKVRIOYGGSVKSGNVAEIMAKPDIDGALVGGASLDSDEFVKIVFRDQ

>**Susscrofa**

MAPARKFFVGGNWKMNKRKNNLGE LINTLNAAKLPADTEVVCAPPTAYIDFARQKLDPKIAVAAQNCYKVANGAFTGEIGPGMIK
DLGATWVVLGHSERRRHVFGESDELIGQVAHALA EGLGVIACIGEKLDEREAGITEKVVFEQTKVIADNVKDWKWLAYEPVWAI GT
GKTATPQQAQEVHEKLRGWLKTHVPEAVAHSTR IYGGSVTGATCKELASQPDVDGFRVSGASLKPEFVDIINAK

>**Taeniasolium**

MTRKLFVGGNWKMNKNGSYSHINTFFDTLQKAD TDPNADIVIGVPACYLKYAQDKAPKGIKIAENCYKVGSGAFTGEISTEMIKDCGC
EWWILGHSERRRHIFGESNELIGEKVHALDSGLN VIPCIGELLSEREAGKTNVCFQAQMDAIAKNVPSKEAWDKVIAYEPVWAI GTG
KTATPAQAEVHKVVRDWIRKHKVDAGIADKVRIL YGGSVTASNAKDLGTQPDVDGFLVGGASLKPDFITIINARR

>**Thermusthermophilus**

MRRVLVAGNWKMHKTPSEARVWF AELKRLLPPLQSEAAVLPAPFIPILVAVEVLAETQVGYGAQDVSAHKEGAYTGEVSARMLSDL
GCRYAIVGHSERRRYHGETDALVAEKAKRLEEGITPILCVGEPLVREKGEAVPYTLRQLRGSLEGVEPPGPEALVIAYEPVWAI GTG
KNATPEDAEAMHQEIRKALSERYGEAFASRVRI LYGGSVNPKNFADLLSMPNVDGGLVGGASLELESFLALLRIAG

>**Trichomonas vaginalis (Ile45var)**

MRTFFVGGNWKANPKTVEEAEKLIEMLN GAKVEGNVEVWAAPFIFLPTLQQLRDKDWKVAENVFTKPNGAFTGEVTVPMIKSFGI
EWTILGHSERRDILKEDDEFLAAKAKFALENGM KIIYCCGEHLSEREAGKASEFVSAQIEKMIPAIPAGKWDDVVIAYEPIWAI GTGK
ASTQDAQEMCKVIRDILAAKVGADIANKVRILY GGSVKPNNCNELAACPDVDGFLVGGASLEPGFINIVNSNVHSK

>**Trichomonas vaginalis (Val45var)**

MRTFFVGGNWKANPKTVQEAELVEMLN GAKVEGNVEVWAAPFVFLPTLQQLRDKDWKVAENVFTKPNGAFTGEVTVPMIKSF
GIEWTILGHSERRDILKEDDEFLAAKAKFALENGM KIIYCCGEHLSEREAGKASEFVSAQIEKMIPAIPAGKWDDVVIAYEPIWAI GTG
KVA STQDAQEMCKVIRDILAAKVGADIANKVRILY GGSVKPNNCNELAACPDVDGFLVGGASLEAGFINIVNSNVHSK



>Trypanosomabrucei

MSKPQPIAAANWKCNGSQSLSELIDLFNSTSINHVDVQCVASTFVHLAMTKERLSHPKFVIAAQNAIAKSGAFTGEVSLPILKDFG
VNWIVLGHSERRAYYGETNEIVADKVAASVSGFMVIACIGETLQERESGRTAVVLTQIAAIAKLLKADWAKVVIAYEPVWAIGTG
KVATPQQAQEAHALIRSWSSKIGADVAGELRILYGGSVNGKNARTLYQQRDVNGFLVGGASLKPEFVDIIKATQ

>Trypanosomacruzi

MASKPQPIAAANWKCNGSESLVPLIETLNAATFDHDVQCVWAPFLHIPMTKARLTNPKFQIAAQNAITRSGAFTGEVSLQILKDY
GIKWVVLGHSERRLYYGETNEIVAEKVAQACAAGFHVIVCVGETNEEREAGRTAAVLTQLAAVAQKLSKEAWAHVVIAYEPVWAIG
TGKVATPQQAQEVHELLRRWVRSKLGADIAAQLRILYGGSVTAKNARTLYQMRDINGFLVGGASLKPEFVEIIEATK

2. Barriles TIM diseñados de novo:

>sTIM11

DKDEAWKQVEQLRREGATQIAYRSDDWRDLKEAWKKGADILIVDATDKDEAWKQVEQLRREGATQIAYRSDDWRDLKEAWKKGGA
DILIVDATDKDEAWKQVEQLRREGATQIAYRSDDWRDLKEAWKKGADILIVDATDKDEAWKQVEQLRREGATQIAYRSDDWRDLKE
AWKKGADILICDAT

>sTIM11-SS

DKDEAWKQVEQLRREGATQIAYRSDDWRDLKEAWKKGADILIVDATDKDEAWKQVEQLRREGATQIAYRSDDWRDLKEAWKKGGA
DILIVDATDKDEAWKQVEQLRREGATQIAYRSDDWRDLKEAWKKGADILIVDATDKDEAWKQVEQLRREGATQIAYRSDDWRDLKE
AWKKGADILIVDAT

>ReTIM1

DKDEAWKQVEQLRREGATQIVYISDDWRDLKEAVKKGDDILIVDATDKDEAWKQVEQLRREGATQIWIYISDDWRDLKEAVKKGDDV
LIVDATDKDEAWKQVEQLRREGATQIVYISDDWRDLKEAVKKGDDILIVDATDKDEAWKQVEQLRREGATQIWIYISDDWRDLKEAVK
KGGDVLIVDAT

>ReTIM2

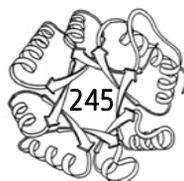
DKDEAWKQVEQLRREGATQIVYQSDWRDLKEAVKKGDDILMV DATDKDEAWKQVEQLRREGATQIMYISDDWRDLKEAVKKGDD
FLV DATDKDEAWKQVEQLRREGATQIVYQSDWRDLKEAVKKGDDILMV DATDKDEAWKQVEQLRREGATQIMYISDDWRDLKE
AVKKGDDFLV DAT

>ReTIM3

DKDEAWKQVEQLRREGATQIMYVSDWRDLKEAVKKGDDSLV DATDKDEAWKQVEQLRREGATQILYKSDDWRDLKEAVKKGDD
TLV DATDKDEAWKQVEQLRREGATQIMYVSDWRDLKEAVKKGDDSLV DATDKDEAWKQVEQLRREGATQILYKSDDWRDLKE
AVKKGDDTLV DAT

>ReTIM4

DKDEAWKQVEQLRREGATQIMYFSDWRDLKEAVKKGDDLLV DATDKDEAWKQVEQLRREGATQILYISDDWRDLKEAVKKGDDI
LLV DATDKDEAWKQVEQLRREGATQIMYFSDWRDLKEAVKKGDDLLV DATDKDEAWKQVEQLRREGATQILYISDDWRDLKEAV
KKGDDILLV DAT



>ReTIM5

DKDEAWKQVEQLRRLGATQIAYRSDDWRDLREAVKKGDDILIVDATDKDEAWKQVEQLRRLGATQIAYRSDDWRDLREAVKKGDDILIVDATDKDEAWKQVEQLRRLGATQIAYRSDDWRDLREAVKKGDDILIVDATDKDEAWKQVEQLRRLGATQIAYRSDDWRDLREAVKKGDDILIVDAT

>ReTIM6

DKDEAWKQVEILRRLGAKQIAYRSDDWRDLQEALKKGADILIVDATDKDEAWKQVEILRRLGAKQIAYRSDDWRDLQEALKKGADILIVDATDKDEAWKQVEILRRLGAKQIAYRSDDWRDLQEALKKGADILIVDATDKDEAWKQVEILRRLGAKQIAYRSDDWRDLQEALKKGADILIVDAT

>ReTIM7

DKDEAWKQVEILRRLGAKQIAYRSDDWRDLDEARKKGADILIVDATDKDEAWKQVEILRRLGAKQIAYRSDDWRDLDEARKKGADILIVDATDKDEAWKQVEILRRLGAKQIAYRSDDWRDLDEARKKGADILIVDATDKDEAWKQVEILRRLGAKQIAYRSDDWRDLDEARKKGADILIVDAT

>ReTIM8

DVDEMLKQVEQLRREGATQIAVRSDDWRILKEAVKKGDDILIVDATDVDEMLKQVEQLRREGATQIAVRSDDWRILKEAVKKGDDILIVDATDVDEMLKQVEQLRREGATQIAVRSDDWRILKEAVKKGDDILIVDATDVDEMLKQVEQLRREGATQIAVRSDDWRILKEAVKKGDDILIVDAT

>ReTIM9

DPDEAQKQVEQLRREGATQIAIRSDDWRYLKEAVKKGDDILIVDATDPDEAQKQVEQLRREGATQIAIRSDDWRYLKEAVKKGDDILIVDATDPDEAQKQVEQLRREGATQIAIRSDDWRYLKEAVKKGDDILIVDATDPDEAQKQVEQLRREGATQIAIRSDDWRYLKEAVKKGDDILIVDAT

>ReTIM10

DSDEAIKQVEQLRREGATQIAVRMDDWRKLKEAVKKGDDILIVDATDSDEAIKQVEQLRREGATQIAVRMDDWRKLKEAVKKGDDILIVDATDSDEAIKQVEQLRREGATQIAVRMDDWRKLKEAVKKGDDILIVDATDSDEAIKQVEQLRREGATQIAVRMDDWRKLKEAVKKGDDILIVDAT

>ReTIM11

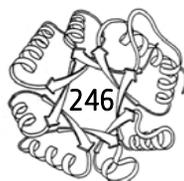
DDKDEAWKMVEQLRREGATQIAYRSDDWRDLKEAWKKGADILIVDATDKDEAWKMVEQLRREGATQIAYRSDDWRDLKEAWKKGADILIVDATDKDEAWKMVEQLRREGATQIAYRSDDWRDLKEAWKKGADILIVDATDKDEAWKMVEQLRREGATQIAYRSDDWRDLKEAWKKGADILIVDAT

>ReTIM12

DPDEAWKQVEQLRREGATQIAYRSDDWRDLKEAWKKGADILIVDATDPDEAWKQVEQLRREGATQIAYRSDDWRDLKEAWKKGADILIVDATDPDEAWKQVEQLRREGATQIAYRSDDWRDLKEAWKKGADILIVDATDPDEAWKQVEQLRREGATQIAYRSDDWRDLKEAWKKGADILIVDAT

>ReTIM13

DKDEAWKQVEQLRREGATQIAYRSDDPRDLKEAVKKGDDILIVDATDKDEAWKQVEQLRREGATQIAYRSDDPRDLKEAVKKGDDILIVDATDKDEAWKQVEQLRREGATQIAYRSDDPRDLKEAVKKGDDILIVDATDKDEAWKQVEQLRREGATQIAYRSDDPRDLKEAVKKGDDILIVDAT



>ReTIM23

DKDEAWKQVEILRRLGAKQIVYISDDWRDLQEALKKGGDILIVDATDKDEAWKQVEILRRLGAKQIWIYISDDWRDLQEALKKGGDVLI
VDATDKDEAWKQVEILRRLGAKQIVYISDDWRDLQEALKKGGDILIVDATDKDEAWKQVEILRRLGAKQIWIYISDDWRDLQEALKKG
GDVLI

>ReTIM24

DVDEMLKQVEILRRLGAKQIWIYISDDWRILQEALKKGGDILIVDATDVDEMLKQVEILRRLGAKQIWIYISDDWRILQEALKKGGDVLI
VDATDVDEMLKQVEILRRLGAKQIWIYISDDWRILQEALKKGGDILIVDATDVDEMLKQVEILRRLGAKQIWIYISDDWRILQEALKKGGDVLI
VDAT

>ReTIM25

DVDEMLKQVEILRRLGAKQIWIYISDDWRILQEALKKGGDILIVDRTDVDEMLKQVEILRRLGAKQIWIYISDDWRILQEALKKGGDVLI
VDATDVDEMLKQVEILRRLGAKQIWIYISDDWRILQEALKKGGDILIVDRTDVDEMLKQVEILRRLGAKQIWIYISDDWRILQEALKKGGDVLI
VDAT

>ReTIM26

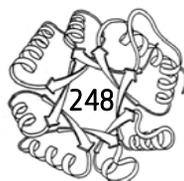
DVDEMLKQVEQLRREGATQIWIYISDDWRILKEAVKGGDILIVDRTDVDEMLKQVEQLRREGATQIWIYISDDWRILKEAVKGGDVLI
VDATDVDEMLKQVEQLRREGATQIWIYISDDWRILKEAVKGGDILIVDRTDVDEMLKQVEQLRREGATQIWIYISDDWRILKEAVKGGDILIV
DRT

>ReTIM27

DVDEMLKQVEILRRLGAKQIYISDDWRILQEALKKGGDILIVDRTDVDEMLKQVEILRRLGAKQIYISDDWRILQEALKKGGDILIV
DRTDVDEMLKQVEILRRLGAKQIYISDDWRILQEALKKGGDILIVDRTDVDEMLKQVEILRRLGAKQIYISDDWRILQEALKKGGDILIV
DRT

>ReTIM28

DKDEAWKQVEILRRLGAKQIVYISDDWRDLQEALKKGGDILIVDRTDKDEAWKQVEILRRLGAKQIWIYISDDWRDLQEALKKGGDVLI
VDATDKDEAWKQVEILRRLGAKQIVYISDDWRDLQEALKKGGDILIVDRTDKDEAWKQVEILRRLGAKQIWIYISDDWRDLQEALKKG
GDVLI





Proteína	-4	-3	-2	-1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
sTIM11					D	K	D	E	A	W	K	C	V	E	Q	L	R	R	E	G	A	T	Q	I	A	Y	R	S	D	D	W	R
sTIM11-SS (ReTIM0)					D	K	D	E	A	W	K	Q	V	E	Q	L	R	R	E	G	A	T	Q	I	A	Y	R	S	D	D	W	R
ReTIM1					D	K	D	E	A	W	K	Q	V	E	Q	L	R	R	E	G	A	T	Q	I	V	Y	I	S	D	D	W	R
ReTIM2					D	K	D	E	A	W	K	Q	V	E	Q	L	R	R	E	G	A	T	Q	I	V	Y	Q	S	D	D	W	R
ReTIM3					D	K	D	E	A	W	K	Q	V	E	Q	L	R	R	E	G	A	T	Q	I	M	Y	V	S	D	D	W	R
ReTIM4					D	K	D	E	A	W	K	Q	V	E	Q	L	R	R	E	G	A	T	Q	I	M	Y	F	S	D	D	W	R
ReTIM5					D	K	D	E	A	W	K	Q	V	E	Q	L	R	R	L	G	A	T	Q	I	A	Y	R	S	D	D	W	R
ReTIM6					D	K	D	E	A	W	K	Q	V	E	I	L	R	R	L	G	A	K	Q	I	A	Y	R	S	D	D	W	R
ReTIM7					D	K	D	E	A	W	K	Q	V	E	I	L	R	R	L	G	A	K	Q	I	A	Y	R	S	D	D	W	R
ReTIM8					D	V	D	E	M	L	K	Q	V	E	Q	L	R	R	E	G	A	T	Q	I	A	V	R	S	D	D	W	R
ReTIM9					D	P	D	E	A	Q	K	Q	V	E	Q	L	R	R	E	G	A	T	Q	I	A	I	R	S	D	D	W	R
ReTIM10					D	S	D	E	A	I	K	Q	V	E	Q	L	R	R	E	G	A	T	Q	I	A	V	R	M	D	D	W	R
ReTIM11					D	K	D	E	A	W	K	M	V	E	Q	L	R	R	E	G	A	T	Q	I	A	Y	R	S	D	D	W	R
ReTIM12					D	P	D	E	A	W	K	Q	V	E	Q	L	R	R	E	G	A	T	Q	I	A	Y	R	S	D	D	W	R
ReTIM13					D	K	D	E	A	W	K	Q	V	E	Q	L	R	R	E	G	A	T	Q	I	A	Y	R	S	D	D	P	R
ReTIM14					D	K	D	E	A	W	K	Q	V	E	Q	L	R	R	E	G	A	T	Q	I	A	Y	R	S	D	D	W	R
ReTIM15				D	C	K	D	E	A	W	K	Q	V	E	Q	L	R	R	E	G	A	T	Q	I	A	Y	R	S	D	D	W	R
ReTIM16	D	C	A	A	A	K	D	E	A	W	K	Q	V	E	Q	L	R	R	E	G	A	T	Q	I	A	Y	R	S	D	D	W	R
ReTIM17 (ReTIM1 + ReTIM8)					D	V	D	E	M	L	K	Q	V	E	Q	L	R	R	E	G	A	T	Q	I	V	V	I	S	D	D	W	R
ReTIM18 (ReTIM6 + ReTIM8)					D	V	D	E	M	L	K	Q	V	E	I	L	R	R	L	G	A	K	Q	I	A	V	R	S	D	D	W	R
ReTIM19 (ReTIM7 + ReTIM8)					D	V	D	E	M	L	K	Q	V	E	I	L	R	R	L	G	A	K	Q	I	A	V	R	S	D	D	W	R
ReTIM20 (ReTIM11 + ReTIM8)					D	V	D	E	M	L	K	M	V	E	Q	L	R	R	E	G	A	T	Q	I	A	V	R	S	D	D	W	R
ReTIM21 (ReTIM12 + ReTIM8)					D	P	D	E	M	L	K	Q	V	E	Q	L	R	R	E	G	A	T	Q	I	A	V	R	S	D	D	W	R
ReTIM22 (ReTIM14 + ReTIM8)					D	V	D	E	M	L	K	Q	V	E	Q	L	R	R	E	G	A	T	Q	I	A	V	R	S	D	D	W	R
ReTIM23 (ReTIM1 + ReTIM6)					D	K	D	E	A	W	K	Q	V	E	I	L	R	R	L	G	A	K	Q	I	V	Y	I	S	D	D	W	R
ReTIM24 (ReTIM1 + ReTIM6 + ReTIM8)					D	V	D	E	M	L	K	Q	V	E	I	L	R	R	L	G	A	K	Q	I	V	V	I	S	D	D	W	R
ReTIM25 (ReTIM24 + ReTIM14)					D	V	D	E	M	L	K	Q	V	E	I	L	R	R	L	G	A	K	Q	I	V	V	I	S	D	D	W	R
ReTIM26 (ReTIM17 + ReTIM14)					D	V	D	E	M	L	K	Q	V	E	Q	L	R	R	E	G	A	T	Q	I	V	V	I	S	D	D	W	R
ReTIM27 (ReTIM18 + ReTIM14)					D	V	D	E	M	L	K	Q	V	E	I	L	R	R	L	G	A	K	Q	I	A	V	R	S	D	D	W	R
ReTIM28 (ReTIM23 + ReTIM14)					D	K	D	E	A	W	K	Q	V	E	I	L	R	R	L	G	A	K	Q	I	V	Y	I	S	D	D	W	R

	Hélices alfa
	Estructura no repetitiva
	Hebras beta



Proteína	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
sTIM11	D	L	K	E	A	W	K	K	G	A	D	I	L	I	V	D	A	T	D	K	D	E	A	W	K	Q	V	E	Q	L	R	R
sTIM11-SS (ReTIM0)	D	L	K	E	A	W	K	K	G	A	D	I	L	I	V	D	A	T	D	K	D	E	A	W	K	Q	V	E	Q	L	R	R
ReTIM1	D	L	K	E	A	V	K	K	G	G	D	I	L	I	V	D	A	T	D	K	D	E	A	W	K	Q	V	E	Q	L	R	R
ReTIM2	D	L	K	E	A	V	K	K	G	G	D	I	L	M	V	D	A	T	D	K	D	E	A	W	K	Q	V	E	Q	L	R	R
ReTIM3	D	L	K	E	A	V	K	K	G	G	D	S	L	L	V	D	A	T	D	K	D	E	A	W	K	Q	V	E	Q	L	R	R
ReTIM4	D	L	K	E	A	V	K	K	G	G	D	L	L	V	V	D	A	T	D	K	D	E	A	W	K	Q	V	E	Q	L	R	R
ReTIM5	D	L	R	E	A	V	K	K	G	G	D	I	L	I	V	D	A	T	D	K	D	E	A	W	K	Q	V	E	Q	L	R	R
ReTIM6	D	L	Q	E	A	L	K	K	G	G	D	I	L	I	V	D	A	T	D	K	D	E	A	W	K	Q	V	E	I	L	R	R
ReTIM7	D	L	D	E	A	R	K	K	G	G	D	I	L	I	V	D	A	T	D	K	D	E	A	W	K	Q	V	E	I	L	R	R
ReTIM8	I	L	K	E	A	V	K	K	G	G	D	I	L	I	V	D	A	T	D	V	D	E	M	L	K	Q	V	E	Q	L	R	R
ReTIM9	Y	L	K	E	A	V	K	K	G	G	D	I	L	I	V	D	A	T	D	P	D	E	A	Q	K	Q	V	E	Q	L	R	R
ReTIM10	K	L	K	E	A	V	K	K	G	G	D	I	L	I	V	D	A	T	D	S	D	E	A	I	K	Q	V	E	Q	L	R	R
ReTIM11	D	L	K	E	A	V	K	K	G	G	D	I	L	I	V	D	A	T	D	K	D	E	A	W	K	M	V	E	Q	L	R	R
ReTIM12	D	L	K	E	A	V	K	K	G	G	D	I	L	I	V	D	A	T	D	P	D	E	A	W	K	Q	V	E	Q	L	R	R
ReTIM13	D	L	K	E	A	V	K	K	G	G	D	I	L	I	V	D	A	T	D	K	D	E	A	W	K	Q	V	E	Q	L	R	R
ReTIM14	D	L	K	E	A	V	K	K	G	G	D	I	L	I	V	D	R	T	D	K	D	E	A	W	K	Q	V	E	Q	L	R	R
ReTIM15	D	L	K	E	A	V	K	K	G	G	D	I	L	I	V	D	A	T	D	K	D	E	A	W	K	Q	V	E	Q	L	R	R
ReTIM16	D	L	K	E	A	V	K	K	G	G	D	I	L	I	V	D	A	T	D	K	D	E	A	W	K	Q	V	E	Q	L	R	R
ReTIM17 (ReTIM1 + ReTIM8)	I	L	K	E	A	V	K	K	G	G	D	I	L	I	V	D	A	T	D	V	D	E	M	L	K	Q	V	E	Q	L	R	R
ReTIM18 (ReTIM6 + ReTIM8)	I	L	Q	E	A	L	K	K	G	G	D	I	L	I	V	D	A	T	D	V	D	E	M	L	K	Q	V	E	I	L	R	R
ReTIM19 (ReTIM7 + ReTIM8)	I	L	D	E	A	R	K	K	G	G	D	I	L	I	V	D	A	T	D	V	D	E	M	L	K	Q	V	E	I	L	R	R
ReTIM20 (ReTIM11 + ReTIM8)	I	L	K	E	A	V	K	K	G	G	D	I	L	I	V	D	R	T	D	V	D	E	M	L	K	M	V	E	Q	L	R	R
ReTIM21 (ReTIM12 + ReTIM8)	I	L	K	E	A	V	K	K	G	G	D	I	L	I	V	D	A	T	D	P	D	E	M	L	K	Q	V	E	Q	L	R	R
ReTIM22 (ReTIM14 + ReTIM8)	I	L	K	E	A	V	K	K	G	G	D	I	L	I	V	D	R	T	D	V	D	E	M	L	K	Q	V	E	Q	L	R	R
ReTIM23 (ReTIM1 + ReTIM6)	D	L	Q	E	A	L	K	K	G	G	D	I	L	I	V	D	A	T	D	K	D	E	A	W	K	Q	V	E	I	L	R	R
ReTIM24 (ReTIM1 + ReTIM6 + ReTIM8)	I	L	Q	E	A	L	K	K	G	G	D	I	L	I	V	D	A	T	D	V	D	E	M	L	K	Q	V	E	I	L	R	R
ReTIM25 (ReTIM24 + ReTIM14)	I	L	Q	E	A	L	K	K	G	G	D	I	L	I	V	D	R	T	D	V	D	E	M	L	K	Q	V	E	I	L	R	R
ReTIM26 (ReTIM17 + ReTIM14)	I	L	K	E	A	V	K	K	G	G	D	I	L	I	V	D	R	T	D	V	D	E	M	L	K	Q	V	E	Q	L	R	R
ReTIM27 (ReTIM18 + ReTIM14)	I	L	Q	E	A	L	K	K	G	G	D	I	L	I	V	D	R	T	D	V	D	E	M	L	K	Q	V	E	I	L	R	R
ReTIM28 (ReTIM23 + ReTIM14)	D	L	Q	E	A	L	K	K	G	G	D	I	L	I	V	D	R	T	D	K	D	E	A	W	K	Q	V	E	I	L	R	R

	Hélices alfa
	Estructura no repetitiva
	Hebras beta



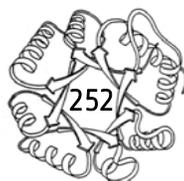
Nota: sólo se muestra la secuencia de la mitad del barril ya que, como las proteínas diseñadas son simétricas, con esa parte se puede reconstruir la secuencia total para cada una de las ReTIMs.

Proteína	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92
sTIM11	E	G	A	T	Q	I	A	Y	R	S	D	D	W	R	D	L	K	E	A	W	K	K	G	A	D	I	L	I	V	D	A	T
sTIM11-SS (ReTIM0)	E	G	A	T	Q	I	A	Y	R	S	D	D	W	R	D	L	K	E	A	W	K	K	G	A	D	I	L	I	V	D	A	T
ReTIM1	E	G	A	T	Q	I	W	Y	I	S	D	D	W	R	D	L	K	E	A	V	K	K	G	G	D	V	L	I	V	D	A	T
ReTIM2	E	G	A	T	Q	I	M	Y	I	S	D	D	W	R	D	L	K	E	A	V	K	K	G	G	D	F	L	V	V	D	A	T
ReTIM3	E	G	A	T	Q	I	L	Y	K	S	D	D	W	R	D	L	K	E	A	V	K	K	G	G	D	T	L	V	V	D	A	T
ReTIM4	E	G	A	T	Q	I	L	Y	I	S	D	D	W	R	D	L	K	E	A	V	K	K	G	G	D	I	L	L	V	D	A	T
ReTIM5	L	G	A	T	Q	I	A	Y	R	S	D	D	W	R	D	L	R	E	A	V	K	K	G	G	D	I	L	I	V	D	A	T
ReTIM6	L	G	A	K	Q	I	A	Y	R	S	D	D	W	R	D	L	Q	E	A	L	K	K	G	G	D	I	L	I	V	D	A	T
ReTIM7	L	G	A	K	Q	I	A	Y	R	S	D	D	W	R	D	L	Q	E	A	R	K	K	G	G	D	I	L	I	V	D	A	T
ReTIM8	E	G	A	T	Q	I	A	V	R	S	D	D	W	R	I	L	K	E	A	V	K	K	G	G	D	I	L	I	V	D	A	T
ReTIM9	E	G	A	T	Q	I	A	I	R	S	D	D	W	R	Y	L	K	E	A	V	K	K	G	G	D	I	L	I	V	D	A	T
ReTIM10	E	G	A	T	Q	I	A	V	R	M	D	D	W	R	K	L	K	E	A	V	K	K	G	G	D	I	L	I	V	D	A	T
ReTIM11	E	G	A	T	Q	I	A	Y	R	S	D	D	W	R	D	L	K	E	A	V	K	K	G	G	D	I	L	I	V	D	A	T
ReTIM12	E	G	A	T	Q	I	A	Y	R	S	D	D	W	R	D	L	K	E	A	V	K	K	G	G	D	I	L	I	V	D	A	T
ReTIM13	E	G	A	T	Q	I	A	Y	R	S	D	D	P	R	D	L	K	E	A	V	K	K	G	G	D	I	L	I	V	D	A	T
ReTIM14	E	G	A	T	Q	I	A	Y	R	S	D	D	W	R	D	L	K	E	A	V	K	K	G	G	D	I	L	I	V	D	R	T
ReTIM15	E	G	A	T	Q	I	A	Y	R	S	D	D	W	R	D	L	K	E	A	V	K	K	G	G	D	I	L	I	V	D	A	T
ReTIM16	E	G	A	T	Q	I	A	Y	R	S	D	D	W	R	D	L	K	E	A	V	K	K	G	G	D	I	L	I	V	D	A	T
ReTIM17 (ReTIM1 + ReTIM8)	E	G	A	T	Q	I	W	V	I	S	D	D	W	R	I	L	K	E	A	V	K	K	G	G	D	V	L	I	V	D	A	T
ReTIM18 (ReTIM6 + ReTIM8)	L	G	A	K	Q	I	A	V	R	S	D	D	W	R	I	L	Q	E	A	L	K	K	G	G	D	I	L	I	V	D	A	T
ReTIM19 (ReTIM7 + ReTIM8)	L	G	A	K	Q	I	A	V	R	S	D	D	W	R	I	L	D	E	A	R	K	K	G	G	D	I	L	I	V	D	A	T
ReTIM20 (ReTIM11 + ReTIM8)	E	G	A	T	Q	I	A	V	R	S	D	D	W	R	I	L	K	E	A	V	K	K	G	G	D	I	L	I	V	D	R	T
ReTIM21 (ReTIM12 + ReTIM8)	E	G	A	T	Q	I	A	V	R	S	D	D	W	R	I	L	K	E	A	V	K	K	G	G	D	I	L	I	V	D	A	T
ReTIM22 (ReTIM14 + ReTIM8)	E	G	A	T	Q	I	A	V	R	S	D	D	W	R	I	L	K	E	A	V	K	K	G	G	D	I	L	I	V	D	R	T
ReTIM23 (ReTIM1 + ReTIM6)	L	G	A	K	Q	I	W	Y	I	S	D	D	W	R	D	L	Q	E	A	L	K	K	G	G	D	V	L	I	V	D	A	T
ReTIM24 (ReTIM1 + ReTIM6 + ReTIM8)	L	G	A	K	Q	I	W	V	I	S	D	D	W	R	I	L	Q	E	A	L	K	K	G	G	D	V	L	I	V	D	A	T
ReTIM25 (ReTIM24 + ReTIM14)	L	G	A	K	Q	I	W	V	I	S	D	D	W	R	I	L	Q	E	A	L	K	K	G	G	D	V	L	I	V	D	R	T
ReTIM26 (ReTIM17 + ReTIM14)	E	G	A	T	Q	I	W	V	I	S	D	D	W	R	I	L	K	E	A	V	K	K	G	G	D	V	L	I	V	D	R	T
ReTIM27 (ReTIM18 + ReTIM14)	L	G	A	K	Q	I	A	V	R	S	D	D	W	R	I	L	Q	E	A	L	K	K	G	G	D	I	L	I	V	D	R	T
ReTIM28 (ReTIM23 + ReTIM14)	L	G	A	K	Q	I	W	Y	I	S	D	D	W	R	D	L	Q	E	A	L	K	K	G	G	D	V	L	I	V	D	R	T

	Hélices alfa
	Estructura no repetitiva
	Hebras beta

APÉNDICE E

Producción científica





Cite this: *Phys. Chem. Chem. Phys.*,
2015, 17, 20699

Reversibility and two state behaviour in the thermal unfolding of oligomeric TIM barrel proteins†‡

Sergio Romero-Romero,^a Miguel Costas,^b Adela Rodríguez-Romero^c and D. Alejandro Fernández-Velasco^{*a}

Temperature is one of the main variables that modulate protein function and stability. Thermodynamic studies of oligomeric proteins, the dominant protein natural form, have been often hampered because irreversible aggregation and/or slow reactions are common. There are no reports on the reversible equilibrium thermal unfolding of proteins composed of $(\beta/\alpha)_8$ barrel subunits, albeit this “TIM barrel” topology is one of the most abundant and versatile in nature. We studied the eponymous TIM barrel, triosephosphate isomerase (TIM), belonging to five species of different bacterial taxa. All of them were found to be catalytically efficient dimers. The three-dimensional structure of four enzymes was solved at high/medium resolution. Irreversibility and kinetic control were observed in the thermal unfolding of two TIMs, while for the other three the thermal unfolding was found to follow a two-state equilibrium reversible process. Shifts in the global stability curves of these three proteins are related to the organismal temperature range of optimal growth and modulated by variations in maximum stability temperature and in the enthalpy change at that temperature. Reversibility appears to correlate with the low isoelectric point, the absence of a residual structure in the unfolded state, small cavity volume in the native state, low conformational stability and a low melting temperature. Furthermore, the strong coupling between dimer dissociation and monomer unfolding may reduce aggregation and favour reversibility. It is therefore very thought-provoking to find that a common topological ensemble, such as the TIM barrel, can unfold/refold in the Anfinsen way, *i.e.* without the help of the cellular machinery.

Received 19th March 2015,
Accepted 7th July 2015

DOI: 10.1039/c5cp01599e

www.rsc.org/pccp

1. Introduction

The study of protein-folding has evolved during the last 50 years to become a mature research field.¹ Although the folding pathway of several proteins is being studied at atomic resolution and several *de novo* proteins have been designed,^{2–4} our current understanding of the sequence/structure/stability relationship of proteins is limited by the information available in these different levels. The current number of protein sequences deposited in the GenBank⁵ ($\approx 10^8$) is several orders of magnitude higher than the number of structures

deposited in the Protein Data Bank (PDB)⁶ ($\approx 10^5$). Likewise, the number of protein three-dimensional structures in the PDB is much bigger than their number in the Protein Stability Database⁷ ($\approx 10^3$).

The evaluation of protein stability requires the use of a perturbant. Temperature and chaotropes such as guanidine hydrochloride (GndHCl) and urea are the main perturbants employed.⁸ The characterization of the thermal unfolding transition gives a full description of the thermodynamic parameters that govern protein stability.^{9,10} Irrespective of the perturbant used to promote unfolding, the two prerequisites for the estimation of thermodynamic parameters are the reversibility of the process and that the system reaches equilibrium. These experimental requirements have limited the number of available data, particularly for medium and large monomeric and oligomeric proteins.¹¹ Although our current understanding of protein folding is based on small, single domain, monomeric and marginally stable systems,¹² oligomers account for at least 80% of the proteins in living cells.¹³ One of the most common protein architectures, found mainly in oligomers but also in monomeric proteins, is the TIM barrel fold, observed in 10% of the proteins deposited in the PDB.¹⁴ This $(\beta/\alpha)_8$ fold is composed of eight β strands that form a

^a Laboratorio de Físicoquímica e Ingeniería de Proteínas, Departamento de Bioquímica, Facultad de Medicina, Universidad Nacional Autónoma de México, 04510 Ciudad de México, Distrito Federal, Mexico. E-mail: fdaniel@unam.mx

^b Laboratorio de Biofísicoquímica, Departamento de Físicoquímica, Facultad de Química, Universidad Nacional Autónoma de México, 04510 Ciudad de México, Distrito Federal, Mexico

^c Laboratorio de Química de Biomacromoléculas 3, Departamento de Química de Biomacromoléculas, Instituto de Química, Universidad Nacional Autónoma de México, 04510 Ciudad de México, Distrito Federal, Mexico

† In memoriam Prof. Armando Gómez-Puyou.

‡ Electronic supplementary information (ESI) available. See DOI: 10.1039/c5cp01599e

central β barrel surrounded by eight α helices. This common architecture has been found in many different proteins with various functions.^{15,16} The Gnd-HCl and urea-induced unfolding of several monomeric TIM barrels has been studied in detail,^{17–19} in contrast, there are only three reports on the reversible temperature-induced unfolding of monomeric TIM barrels^{20–22} and none describing an oligomeric TIM barrel. The only oligomeric TIM barrel with a reversible chaotrope-induced unfolding transition so far characterized is triosephosphate isomerase, the first protein where the TIM barrel fold was found.^{23–38}

Triosephosphate isomerase (E. C. 5.3.1.1) (TIM) is a key glycolytic enzyme that catalyzes the isomerization of glyceraldehyde-3-phosphate (GAP) and dihydroxyacetone phosphate (DHAP).^{39,40} TIM is an obligate homooligomer, with monomers structured in the canonical $(\beta/\alpha)_8$ barrel. The three-dimensional structures and catalytic properties of TIM from different species are quite similar.^{41,42} In contrast, the unfolding pathway of TIM from different organisms is very diverse. Chaotrope-induced unfolding transitions have been described by models which increase in complexity from two-state to multistate reactions that involve monomeric and/or dimeric intermediates.^{23–38} Regarding thermal unfolding, eighteen TIMs, mainly from eukaryotes, as diverse as Amoebozoa, Euglenozoa, Ascomycota and Chordata, have been studied. Even though full thermodynamic characterization has been hampered by irreversible aggregation and/or the presence of hysteresis in all of them,^{36,38,43–53} the activation parameters that describe the kinetic control of five eukaryotic TIMs have been reported.^{52,53}

Although species from the bacterial domain are abundant and widely distributed under a variety of environmental conditions, bacterial TIMs have been much less studied than eukaryotic TIMs despite their biomedical or biotechnological importance. In this work, we characterized the structure, catalytic properties, association state and temperature-induced unfolding of five bacterial TIMs representatives of different supertaxa or clades that comprise different stages of bacterial evolution.⁵⁴ They were the TIMs of *Deinococcus radiodurans* (*Dr*TIM) from Hadobacteria, of *Nostoc punctiforme* (*Np*TIM) from Cyanobacteria, of *Gemmata obscuriglobus* (*Go*TIM) from Gracilicutes, of *Clostridium perfringens* (*Cp*TIM) from Endobacteria and of *Streptomyces coelicolor* (*Sc*TIM) from Actinobacteria (Fig. S1, ESI \ddagger). Here, we studied their thermal unfolding finding that two of them present kinetic control whereas for the other three, the presence of reversible transitions allowed an equilibrium thermodynamic analysis. Furthermore, the three-dimensional structures of four of these enzymes were solved at high/medium resolution and a structural analysis was performed to explore the possible relation between molecular level properties and reversibility.

2. Results and discussion

2.1 Biochemical characterization

The five recombinant proteins characterized in this work were purified to homogeneity with high yields (see Materials and methods). Their far UV-circular dichroism (CD) native spectra

showed minima at 210 and 220 nm (Fig. S2, ESI \ddagger), a characteristic feature of β/α proteins. λ_{max} and SCM of their intrinsic fluorescence spectra are (Table 1 and Fig. S3, ESI \ddagger) in the range found in many native proteins.⁵⁵ As previously observed in other wild type TIMs, catalytic activity measurements were well-fitted by the Michaelis–Menten equation and catalytic efficiency values were near the diffusion-limited behaviour (Table 1). Size Exclusion Chromatography (SEC) experiments indicate that BacTIMs show Stoke's radius expected for a compact dimer (Table 1); however, at a lower protein concentration (in the 0.5–5.0 μM range), catalytic activity decreased in a concentration dependent way, the signature of a dissociation process to inactive monomers (Fig. 1). For *Dr*TIM, *Sc*TIM and *Cp*TIM dissociation was a reversible process; in contrast, for *Go*TIM and *Np*TIM activity was not recovered after the samples were concentrated (Fig. 1). From these experiments the dimer dissociation constants were calculated. The obtained values (30 to 180 nM; Table 1) are similar to those reported for eukaryotic TIMs (EukTIMs).^{56–60}

2.2 Kinetic control in the thermal unfolding of *Np*TIM and *Go*TIM

The temperature-induced unfolding of BacTIMs was then studied. The native CD spectra of *Np*TIM and *Go*TIM were not recovered after heating the samples to 75 $^{\circ}\text{C}$ and cooling back to 25 $^{\circ}\text{C}$, protein aggregation in the CD cell was observed, and there was no recovery of enzyme activity. Hence, the temperature-induced unfolding of these two proteins is irreversible. Differential scanning calorimetry (DSC) was then used to determine the activation parameters that describe the thermal unfolding process. DSC traces showed an increase in melting temperature (T_m) with a scan rate (Fig. 2a, and Fig. S4, ESI \ddagger) and no endotherm was observed in a second heating run, showing that *Np*TIM and *Go*TIM display kinetic control. The following evidence indicates that the temperature-induced unfolding of these proteins can be described by the two-state irreversible model: (i) at all scanning rates, the thermograms were well-fitted (Fig. 2a), giving similar activation energy (E_A) values (Table S1, ESI \ddagger), (ii) the Arrhenius plots were linear (Fig. 2b), and (iii) the data consistency test was successful, *i.e.* eqn (5) was fulfilled (Fig. 2c). This evidence also implies that the kinetically relevant transition state is dimeric. The average E_A values obtained from the three methods are $360 \pm 25 \text{ kJ mol}^{-1}$ for *Np*TIM and $715 \pm 46 \text{ kJ mol}^{-1}$ for *Go*TIM (Table S1, ESI \ddagger). These values are similar to the lower and higher values reported for E_A of EukTIMs ($325\text{--}774 \text{ kJ mol}^{-1}$; Table S1, ESI \ddagger). A consequence of E_A for *Go*TIM being twice that of *Np*TIM is that the denaturation rate under physiological conditions for *Go*TIM is many orders of magnitude slower than for *Np*TIM, indicating a much higher kinetic stability for the former. The fractional degree of exposure of the proteins to the solvent in the transition state ($m^{\ddagger}/m_{\text{eq}}$) was calculated from DSC experiments using different concentrations of urea (Fig. 2d). The values obtained for *Np*TIM and *Go*TIM (0.06 and 0.29 respectively) indicate that the transition state for thermal unfolding is native-like in both proteins (Table S1, ESI \ddagger).

Table 1 Physicochemical characterization of bacterial TIM

TIM	Catalytic parameters for GAP isomerization ^a			Hydrodynamic and spectroscopic properties			Predicted secondary structure content ^c (%)					T_m^e (°C)
	K_m (mM)	K_{cat} (s ⁻¹)	Catalytic efficiency ^b (M ⁻¹ s ⁻¹)	Stokes radius ^d (Å)	λ_{max} (nm)	SCM (nm)	α -helix	β -strand	β -turn	Random coil	K_{diss}^d (M)	
<i>NpTIM</i>	0.51 ± 0.04	1736 ± 42	(8.63 ± 0.29) × 10 ⁷	31.4 ± 0.7	331	342.1	36.1 ± 0.6	15.1 ± 0.5	16.3 ± 0.2	32.6 ± 0.8	(5.6 ± 0.7) × 10 ⁻⁸	55.2 ± 0.4
<i>GoTIM</i>	0.34 ± 0.02	1306 ± 27	(9.51 ± 0.14) × 10 ⁷	31.7 ± 0.8	325	339.5	31.4 ± 1.7	16.8 ± 0.8	15.8 ± 0.5	35.1 ± 0.5	(1.1 ± 0.2) × 10 ⁻⁷	65.6 ± 0.5
<i>DrTIM</i>	0.69 ± 0.03	5678 ± 84	(2.06 ± 0.18) × 10 ⁸	31.6 ± 0.9	324	337.7	33.3 ± 0.9	16.3 ± 0.7	16.2 ± 0.3	34.3 ± 1.1	(1.4 ± 0.5) × 10 ⁻⁷	44.6 ± 0.1
<i>ScTIM</i>	0.74 ± 0.02	5668 ± 62	(1.93 ± 0.06) × 10 ⁸	31.9 ± 0.6	325	339.5	41.3 ± 1.0	13.3 ± 0.9	15.7 ± 0.4	29.7 ± 0.7	(1.8 ± 0.5) × 10 ⁻⁷	45.3 ± 0.2
<i>CpTIM</i>	0.72 ± 0.04	6348 ± 127	(2.24 ± 0.10) × 10 ⁸	32.1 ± 0.6	326	339.4	33.9 ± 1.9	16.5 ± 0.8	16.8 ± 0.4	32.8 ± 1.4	(2.9 ± 0.3) × 10 ⁻⁸	58.3 ± 0.1

^a Errors are the standard deviation of three different experiments. ^b Catalytic efficiency was calculated from k_{cat}/K_m with the K_m value having been adjusted by the appropriate ratio (1:29) considering that only 4% of GAP in aqueous solution is the unhydrated aldehyde substrate, the only form that TIM is able to catalyze. ^c CD spectra were analyzed with CDNN, ⁶² CDPro⁶³ and DichroWeb⁶⁴ to calculate the contributions of the various components to the protein secondary structure. Reported errors are an average of the deviations obtained using the three different software packages. ^d Fitted to the dilution experiments at 25 °C using eqn (1) (Fig. 1). For *NpTIM* and *GoTIM* K_{diss} are apparent K_{diss} since the process is irreversible. Errors are derived from data fitting. ^e Average value of the CD and DSC experiments with a protein concentration of 15 μ M and a scan rate of 1.0 K min⁻¹. Errors are the standard deviation of CD and DSC experiments.

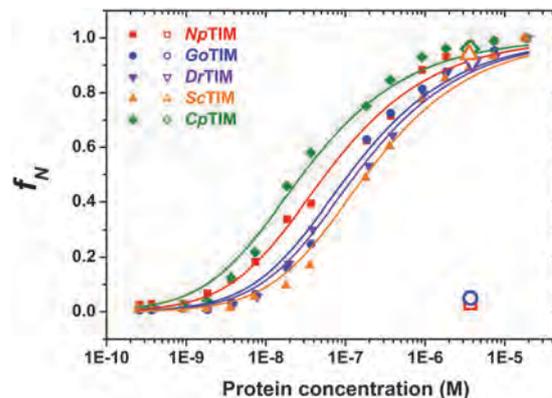


Fig. 1 Fraction of active TIM molecules as a function of protein concentration at 25 °C. Solid curves show the best fit of experimental data (closed symbols) to a dissociation model (eqn (1)). For the reversibility test (open symbols), samples were incubated at 2×10^{-9} M for 48 hours and then concentrated to 4×10^{-6} M prior to catalytic activity measurements.

2.3 The thermal unfolding of *DrTIM*, *ScTIM* and *CpTIM* is a two-state equilibrium reversible process

For *DrTIM*, *ScTIM* and *CpTIM*, the following experimental behaviour was found: (i) their CD spectra were very similar before and after a heating cycle (Fig. 3a–c inset) and unfolding and refolding transitions were super-imposable (Fig. 3a–c); the high temperature CD spectra of *DrTIM*, *ScTIM* and *CpTIM* display the common spectral shape observed in thermally unfolded proteins with no residual structure, characterized by a broad shoulder at around 220 nm and a large negative peak at *ca.* 200 nm (Fig. S5, ESI[†]);^{10,65,66} (ii) T_m was independent of the scan rate (Fig. S4, ESI[†]); (iii) DSC endotherms obtained in the first and second scans were almost identical (Fig. 4a–c) and (iv) catalytic activity of the refolded enzymes was between 84 and 96% of the native samples. All the evidence shows that the temperature-induced unfolding of *DrTIM*, *ScTIM* and *CpTIM* is reversible and at equilibrium under the experimental conditions tested. All these contrast with the thermal unfolding of *NpTIM*, *GoTIM* and all previously studied TIMs which are irreversible or present hysteresis^{36,38,43–53} and their high temperature CD spectra show the shape and a decrease in the spectral signal observed in unfolded and aggregated proteins (Fig. S5, ESI[†]).

The thermal unfolding of *DrTIM*, *ScTIM* and *CpTIM* can be well-described by a two state dissociation/unfolding process; the unfolding melts followed by CD (Fig. 3a–c), as well as the DSC traces (Fig. 4a–c) were well-fitted to this model. In addition, the enthalpy change from the native state to the unfolded state (ΔH_{UH}) calculated from the van't Hoff analysis of CD data (Fig. 3d) and those obtained from DSC experiments were similar (Table 2). Likewise, T_m increased with protein concentration, as expected for the coupled dissociation/unfolding of a dimeric protein (Fig. 4a–c inset) and the Takahashi–Sturtevant plot constructed with CD and DSC data was linear (Fig. 4d). Finally, the ratio of the van't Hoff and calorimetric enthalpies obtained (Table 2) was in excellent agreement with that expected (1.33) for a coupled dissociation/unfolding process.⁶⁷

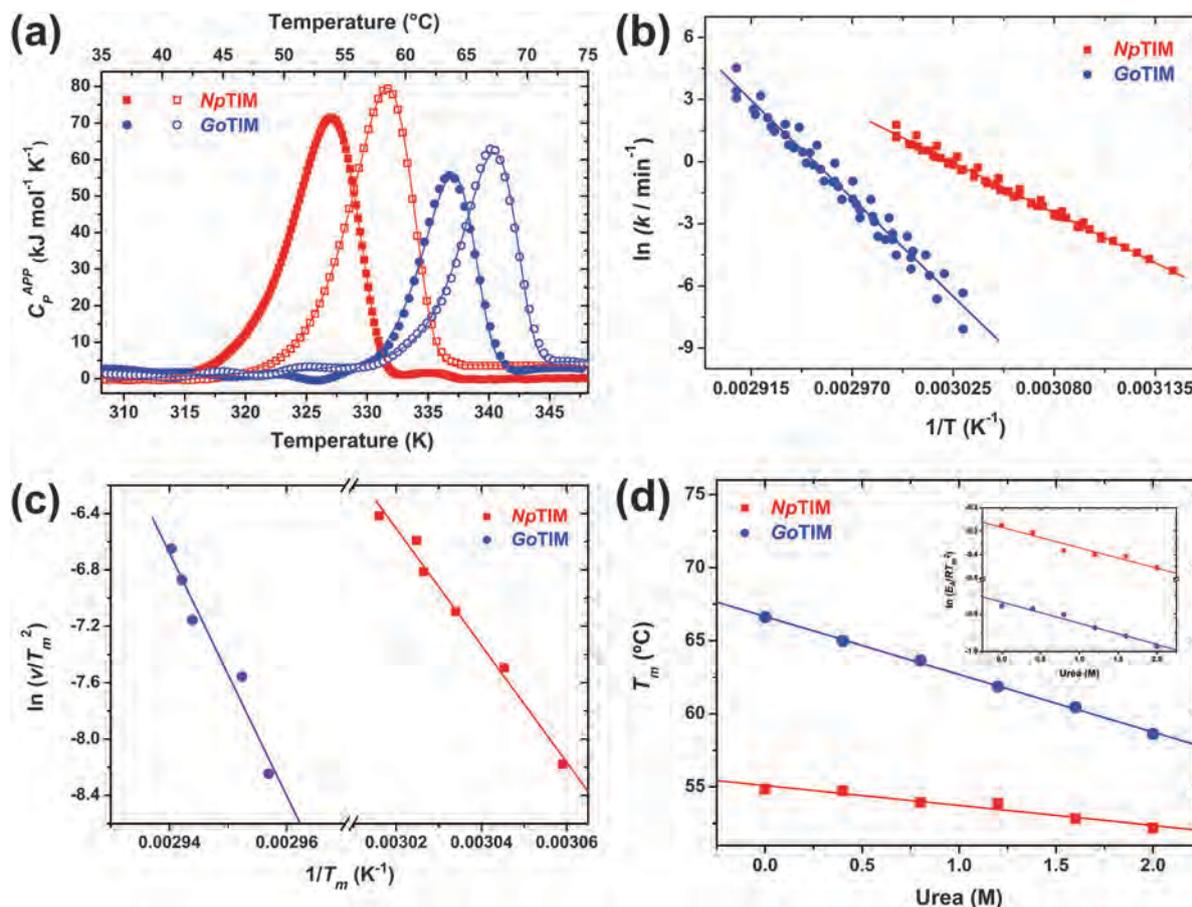


Fig. 2 Irreversible thermal unfolding of *NpTIM* and *GoTIM*. (a) DSC thermograms at 0.5 and 3.0 K min^{-1} (closed and open symbols, respectively). Lines show the best fit to the two-state irreversible model (eqn (3) and (4)). Other scan rates (1.0, 1.5, 2.0 and 2.5 K min^{-1}) were also assayed but are not shown for clarity. (b) Arrhenius plot from data obtained at six scan rates varying from 0.5 to 3.0 K min^{-1} . The lines are the best fit to the Arrhenius equation (eqn (2)). (c) Effect of the scan rate on T_m for the unfolding of *NpTIM* and *GoTIM*. Lines represent the best fits to eqn (5). (d) Effect of urea concentration on T_m . The slopes of the linear fits provide the value of $dT_m/d[\text{urea}]$ used in eqn (6). The inset illustrates the calculation of the derivative $d\ln(E_a/RT_m)/d[\text{urea}]$ that is also involved in the calculation of kinetic m^\ddagger values. Protein concentration in all panels was 15 μM .

2.4 The global stability curve of *DrTIM*, *ScTIM* and *CpTIM*

The reversibility and equilibrium conditions found in the temperature-induced unfolding of *DrTIM*, *ScTIM* and *CpTIM* allowed the estimation of ΔH , T_m and change in heat capacity of unfolding (ΔC_p) from DSC measurements (Tables 1 and 2). ΔC_p was obtained from the DSC thermograms at different TIM concentrations. To this end, two different procedures were employed: (i) C_p for the native and unfolded state was calculated from the difference of absolute heat capacity in the native and unfolded state (see Material and methods) and (ii) DSC traces were fitted to an equation that includes ΔC_p as a fitting parameter (eqn (8)). The average ΔC_p values obtained using these methods were 27.6 ± 1.7 for *DrTIM*, 26.8 ± 2.5 for *ScTIM* and 28.5 ± 2.1 kJ mol^{-1} for *CpTIM* (Table 2). These values are in accordance with the ΔC_p values predicted from the number of amino acids of these TIMs;⁶⁹ they are, however, much lower than the ΔC_p values predicted from other parameterizations, as previously reported for other dimeric systems.¹¹

Using eqn (10), the global stability curves, *i.e.* $\Delta G(T)$ for the coupled dissociation/unfolding (ΔG_{tot}) as a function of temperature, were calculated and are shown in Fig. 5. ΔG_{tot} values near T_m , obtained from CD experiments (solid symbols in Fig. 5), are in excellent agreement with the stability curve generated using DSC data. ΔG_{diss} values obtained from dilution experiments carried out at different temperatures are also shown (open symbols in Fig. 5). *D. radiodurans*, *S. coelicolor* and *C. perfringens* are mesophiles, however, when the stability curves of their TIMs are compared, a 14 degree difference in T_m and a 17 kJ mol^{-1} difference in ΔG_{tot} were found. Changes in T_S (the temperature where $\Delta S = 0$), ΔH_{TS} (enthalpy change at T_S), and/or ΔC_p modulate the shape of the stability curve.⁷⁵ In the present case, these differences are mainly due to changes in ΔH_{TS} and T_S (Fig. 5). ΔH_{TS} can be changed by mutation and has been related to an increase in stabilizing interactions in the native state; T_S was originally not considered a source of stabilization, however, recently it has been reported that changes in both ΔH_{TS} and T_S can increase T_m .⁷⁶ A comparison of the temperature range of optimal growth (TOG)^{77–79} in relation to the stability

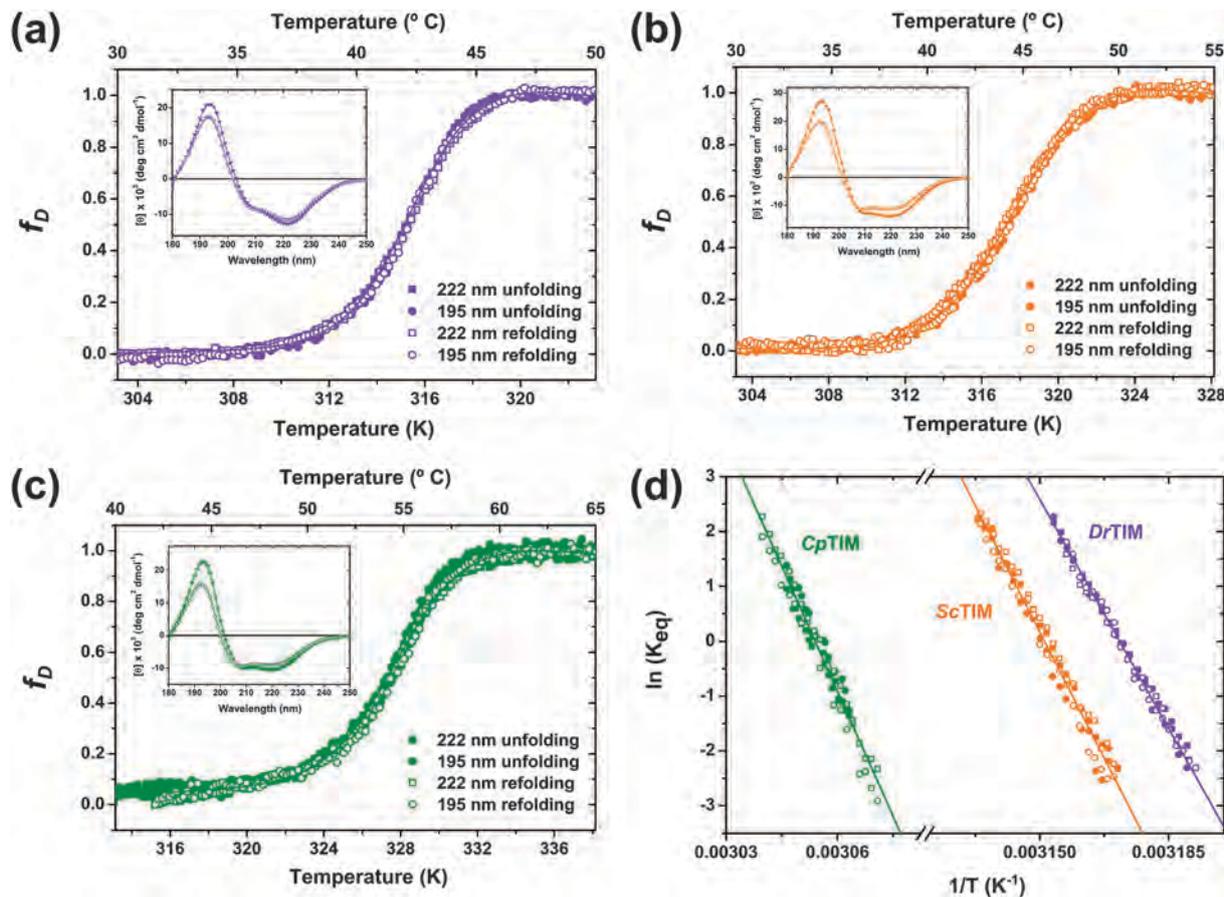


Fig. 3 Reversible thermal unfolding of *DrTIM*, *ScTIM* and *CpTIM* followed by CD. Unfolding/refolding data (closed and open symbols, respectively) for *DrTIM* (a), *ScTIM* (b), and *CpTIM* (c) were followed at 220 nm (squares) and 195 nm (circles) at 1.0 K min^{-1} . Lines show the best fit to a two-state dimer dissociation model ($N_2 \rightleftharpoons 2U$; eqn (7)). The inset shows the CD spectra of native TIM at 25°C (closed symbols) and refolded TIM after a heating cycle (open symbols). (d) van't Hoff plots of unfolding and refolding data (closed and open symbols, respectively) followed at 220 nm (squares) and 195 nm (circles). Lines are the best fits to eqn (7). ΔH_{vH} values derived from these fittings are reported in Table 2. In all panels, protein concentration was $15 \mu\text{M}$.

curve indicates that these proteins function at temperatures higher than their maximal stability (purple, orange and green stripes in Fig. 5). It is interesting to note that an increase in T_m is accompanied by an increase in TOG, in such a way that conformational stability within the TOG fluctuates in a narrow range (24.3 to 39.8 kJ mol^{-1}).

2.5 Three-dimensional structure of BacTIMs

In order to correlate the thermodynamic and structural properties of BacTIMs, crystallization assays were performed with all of them. Good quality crystals were obtained for *GoTIM*, *DrTIM*, *ScTIM* and *CpTIM*, and their structure was therefore solved at high/medium resolution by X-ray crystallography. Data collection and refinement statistics are shown in Table S2, ESI.† Crystallographic data show a dimeric association state, in agreement with SEC experiments (Table 1). The canonical $(\beta\alpha)_8$ topology of the barrel is very similar in all the structure of the BacTIMs determined in this work (Fig. 6a). The orientation of the monomers with respect to each other in the dimer is also similar; in fact, it is fairly conserved among TIMs from the three domains of life. The so-called “catalytic side” of TIM barrels is formed by eight loops that connect the carboxyl-terminal

end of the β -strands with the amino-terminal end of the α -helix. In BacTIMs, as well as in all previously reported TIM structures, these loops provide the catalytic residues and most of the interactions that stabilize the monomer–monomer interface. The other side of the barrel is formed by the “stability loops” that connect the carboxyl-terminal end of the α -helices with the amino-terminal end of the β -strands (Fig. 6b). The correlation between the structure and thermodynamics will be discussed below in the context of the reversibility in the folding process.

2.6 On the possible relation between molecular properties and reversibility

The molecular origin of the reversibility of protein folding is not known. Hence, as a first step to advance towards that knowledge, it is interesting to establish reasonably supported relations between reversibility and the molecular level properties of the protein. Here, we explored such possible relations at three levels: (i) sequence-related properties and their phylogenetic relations, (ii) the structural properties of the stable states, and (iii) the physicochemical properties of the unfolding/refolding transition between them. To this end, the three reversible TIMs reported

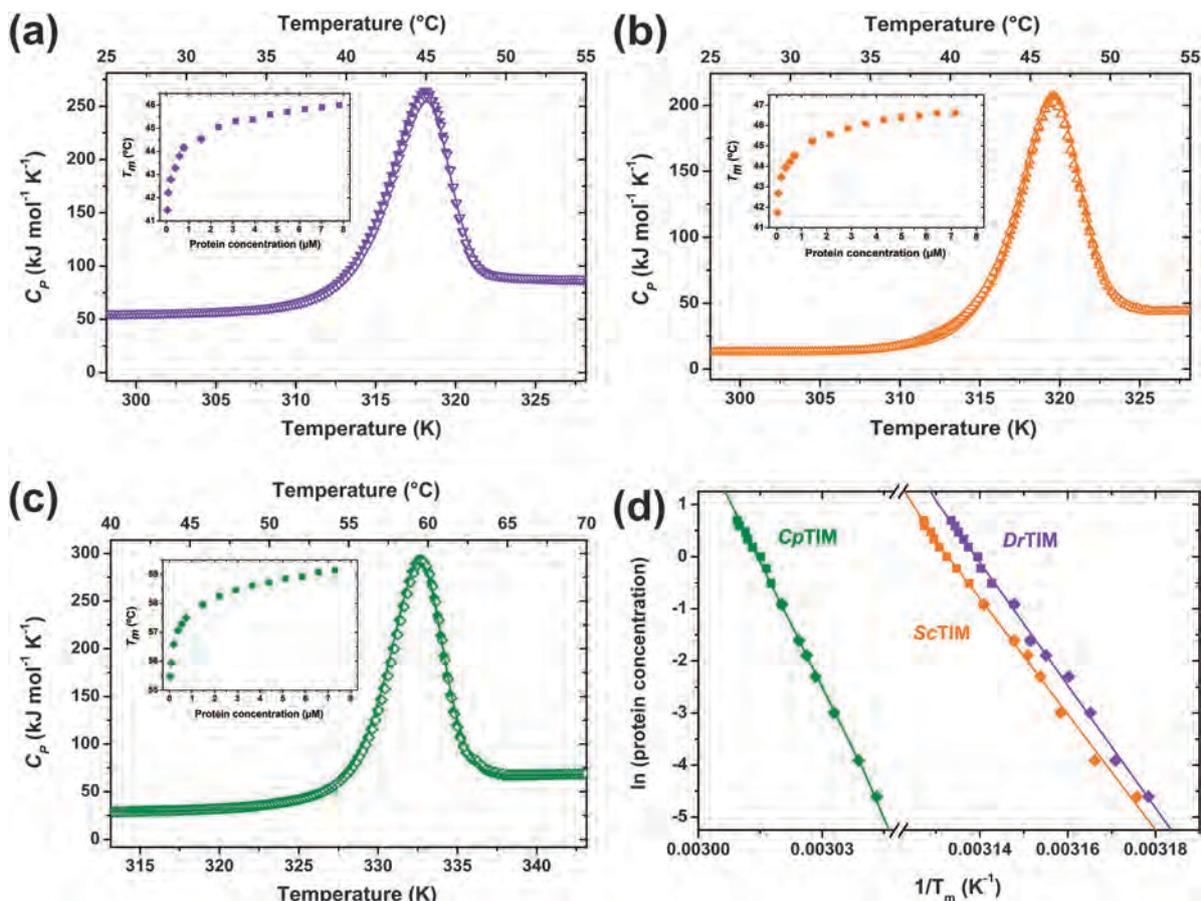


Fig. 4 Reversible thermal unfolding of *DrTIM*, *ScTIM* and *CpTIM* followed by DSC. The first and second endotherms (closed and open symbols, respectively) are shown for *DrTIM* (a), *ScTIM* (b), and *CpTIM* (c). Continuous curves show the best fit to a two-state dimer dissociation model ($N_2 \rightleftharpoons 2U$; eqn (8)). In all panels, the scan rate was 1.5 K min^{-1} and protein concentration is $15 \mu\text{M}$. The inset shows the T_m values obtained by CD (diamonds) and DSC (squares) as a function of TIM concentration at a scan rate of 1.5 K min^{-1} . (d) Takahashi–Sturtevant plot for CD and DSC data (diamonds and squares, respectively). The lines are the best fit to eqn (9).

in this work, herein named “RevTIMs”, were compared with the 19 dimeric TIMs that show irreversible thermal unfolding, henceforth called “IrrevTIMs”. Note that for three IrrevTIMs there is no structure available.

(i) Sequence-related properties and phylogenetic relations.

The identity among the sequence of the three RevTIMs is between 40% and 52%. The sequence identity among the three RevTIMs is similar to those between RevTIMs and IrrevTIMs (Table S3, ESI \ddagger). There are no significant differences in the abundance of a particular group of amino acids between both groups (Table S4, ESI \ddagger). There is also no particular region in the sequence that is different amongst RevTIMs and IrrevTIMs (Fig. S6, ESI \ddagger). Regarding the global properties of the sequence, RevTIMs tend to have a lower pI than IrrevTIMs (Table 3); in accordance with this, it is known that reversibility increases when the experimental pH is far away from the pI of the protein.^{81–83} It should be noted, however, that some IrrevTIMs such as TIM from *Moritella marina* (*MmTIM*), *NpTIM* and TIM from *Geobacillus stearothermophilus* (*GsTIM*) show a pI similar to that of RevTIMs. Nevertheless, the three RevTIMs present pI values outside the $\mu \pm \sigma$ range determined for the IrrevTIMs (Table 3). Concerning the

phylogenetic relation between the species where RevTIMs come from, *S. coelicolor* and *C. perfringens* belong to the Actinobacteria and Firmicutes, respectively. Both phyla form the Posibacteria group, whereas *D. radiodurans* belongs to the Deinococcus-Thermus phylum that forms part of the distant Eobacteria group (Fig. S1, ESI \ddagger). Therefore reversibility does not seem to be restricted to a particular phylogenetic group.

(ii) **Structural properties of stable states.** The properties of the unfolded state may be involved in reversibility. *NpTIM* and *GoTIM*, the two IrrevTIMs reported here, show the spectra observed when thermal unfolding is associated with irreversible aggregation and is consequently not suitable for an estimation of the residual structure; *YtTIM*, which presents under some conditions a reversible thermal unfolding transition with a marked hysteresis, shows high temperature CD spectra with evidence of the residual structure.⁴⁷ In contrast, the CD spectra of the thermally-induced unfolded states of RevTIMs do not show evidence of the secondary structure (Fig. S5, ESI \ddagger). Therefore, it is possible that the absence of a residual structure may be related to reversibility. There are no significant differences in the backbone trace of the native state of RevTIMs and IrrevTIMs (Fig. 6a). The length of the secondary

Table 2 Thermodynamic parameters of reversible thermal unfolding of *Dr*TIM, *Sc*TIM and *Cp*TIM

Physicochemical parameter	<i>Dr</i> TIM	<i>Sc</i> TIM	<i>Cp</i> TIM
Enthalpy change (ΔH_{vH}) (kJ mol ⁻¹)			
ΔH_{vH} by DC	1071 ± 29	1178 ± 38	1205 ± 46
ΔH_{vH} by DSC	1121 ± 42	1255 ± 96	1230 ± 67
ΔH_{cal} by DSC	811 ± 34	941 ± 33	895 ± 50
Calorimetric criterion ^a ($\Delta H_{\text{vH}}/\Delta H_{\text{cal}}$)	1.38 ± 0.06	1.33 ± 0.09	1.37 ± 0.08
$\Delta H'$ from a Takahashi–Sturtevant plot	1108 ± 46	1150 ± 33	1247 ± 63
Heat capacity change (ΔC_p) (kJ mol ⁻¹ K ⁻¹)			
Experimental ΔC_p	27.6 ± 0.9	25.9 ± 3.4	28.5 ± 3.4
Fitted ΔC_p ^b	27.2 ± 0.9	27.6 ± 1.7	28.5 ± 1.7
Average ΔC_p	27.6 ± 1.7	26.8 ± 2.5	28.5 ± 2.1
Predicted ΔC_p based on amino acidic sequence			
ΔC_p by ref. 68	32.2	35.6	35.2
ΔC_p by ref. 69	28.9	30.1	29.3
Predicted ΔC_p based on Δ ASA correlations			
ΔC_p by ref. 70	57.8	57.8	59.4
ΔC_p by ref. 71	43.9	43.9	45.2
ΔC_p by ref. 72	71.1	71.1	73.2
ΔC_p by ref. 73	40.2	40.2	41.8
ΔC_p by ref. 69	33.1	33.1	35.2
ΔC_p by ref. 74	37.1	38.1	40.2

^a Errors are the standard deviation of different experiments at different protein concentrations and/or scan rates. Calculated with DSC values.

^b ΔC_p obtained from DSC experiments fitted to eqn (8).

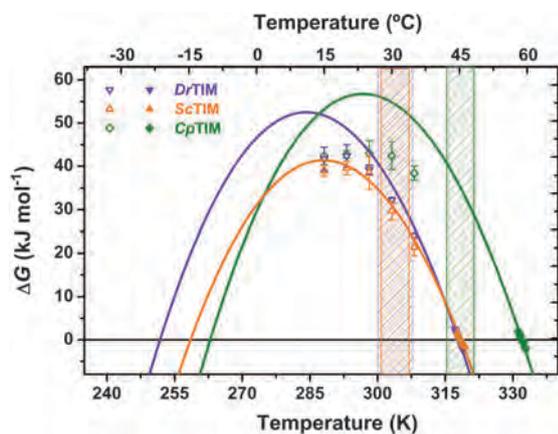


Fig. 5 Stability curves of *Dr*TIM, *Sc*TIM and *Cp*TIM. Solid lines are the stability curves constructed with data obtained by DSC and the Gibbs–Helmholtz equation (eqn (10)) (Data used: *Dr*TIM (purple): $\Delta H = 811$ kJ mol⁻¹, $\Delta C_p = 27.6$ kJ mol⁻¹ K⁻¹ and $T_m = 317.4$ K; *Sc*TIM (orange): $\Delta H = 941$ kJ mol⁻¹, $\Delta C_p = 26.8$ kJ mol⁻¹ K⁻¹ and $T_m = 318.5$ K; *Cp*TIM (green): $\Delta H = 895$ kJ mol⁻¹, $\Delta C_p = 28.5$ kJ mol⁻¹ K⁻¹ and $T_m = 331$ K). Open symbols indicate experimental ΔG values obtained by CD in the transition (Fig. 3a–c) and closed symbols show the dissociation ΔG obtained from dilution experiments (Fig. 1) carried out at different temperatures (288.15 to 308.15 K). Purple, orange and green stripes are the temperature ranges of optimal growth (TOG) of *D. radiodurans*, *S. coelicolor* and *C. perfringens*, respectively.^{77–79}

structure elements and the conformation of loops are very similar (Table S5, ESI[†]), and the average RMSD values among RevTIMs and IrrevTIMs are also indistinguishable (0.2–1.5 Å; Table S3, ESI[†]). It should be emphasized that, the four structures determined in this work, as well as most of the fifteen structures used for comparison, were crystallized without ligands; consequently, most of them show the catalytic loop 6 in the “open”

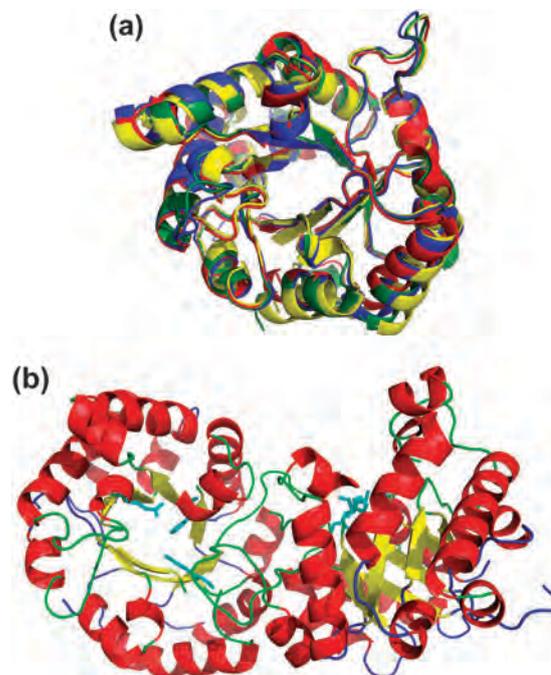


Fig. 6 The three-dimensional structure of TIM. (a) Structural superposition of the monomer of *Go*TIM (red), *Dr*TIM (yellow), *Sc*TIM (green), and *Cp*TIM (blue). (b) Ribbon representation of the 3D structure of the *Cp*TIM dimer. Loops from the catalytic side are shown in green whereas stability loops are shown in blue. Catalytic residues are shown in cyan sticks.

conformation (Table S6, ESI[†]). Regarding the structural features of the monomers, no differences were found when the interactions within the monomer were compared (Table S5, ESI[†]). Likewise, when the dimer interface of RevTIMs and IrrevTIMs

Table 3 Isoelectric point, cavity volume and melting temperature for IrrevTIMs and RevTIMs

TIM	pI ^a	Cavity volume ^b (Å ³)	T _m ^c (°C)	Ref. ^d
IrrevTIMs				
<i>Eh</i> TIM ^e	6.1	3777	40.5/56.3	51
<i>Mm</i> TIM	4.8	3927	41.0	44
<i>Ec</i> TIM	5.6	4667	54.0	44
<i>Mt</i> TIM	5.9	3087	55.0	49
<i>Np</i> TIM	5.0	ND	55.2	This work
<i>Tb</i> TIM	9.1	3780	55.8	52
<i>Oc</i> TIM	7.1	3485	56.0	43
<i>Gl</i> TIM	7.1	3920	57.5	50
<i>Yt</i> TIM	5.8	3699	59.1	47
<i>Tc</i> TIM	8.6	3414	59.2	52
<i>Ts</i> TIM	6.6	ND	62.3	48
<i>Lm</i> TIM	8.2	3531	62.5	52
<i>Hs</i> TIM	6.5	3259	63.1	53
<i>Ss</i> TIM	7.1	ND	63.9	48
<i>Pf</i> TIM	6.1	4452	65.0	46
<i>Go</i> TIM	6.3	4010	65.6	This work
<i>Tv</i> TIM ^f	5.6	3383	67.1	38
<i>Gg</i> TIM	6.7	3518	67.2	36
<i>Gs</i> TIM	5.2	2503	76.0	45
Mean (μ)	6.5	3651	59.3	
Standard deviation (σ)	1.2	514	8.5	
$\mu + \sigma$	7.7	4165	67.8	
$\mu - \sigma$	5.3	3137	50.7	
RevTIMs				
<i>Dr</i> TIM	4.9	2134	44.6	This work
<i>Sc</i> TIM	5.2	2921	45.3	This work
<i>Cp</i> TIM	5.0	2241	58.3	This work

^a Determined by ProtParam software.⁸⁰ ^b Cavity volume in the native state was calculated without water molecules as described in Materials and methods. ^c Value reported in experiments at pH 7.4 with protein concentration of ≈ 10 – $15 \mu\text{M}$ and scan rate of 1.0 K min^{-1} , except for *Tv*TIM in which the scan rate used was 2.0 K min^{-1} . ^d T_m values were obtained from these references. ^e The values reported correspond to dimer dissociation and monomer unfolding. ^f Reported T_m of the dimeric Ile45 variant. Abbreviations: *Cp*TIM, TIM from *Clostridium perfringens*; *Dr*TIM, TIM from *Deinococcus radiodurans*; *Ec*TIM, TIM from *Escherichia coli*; *Eh*TIM, TIM from *Entamoeba histolytica*; *Gg*TIM, TIM from *Gallus gallus*; *Go*TIM, TIM from *Gemmata obscuriglobus*; *Gl*TIM, TIM from *Giardia lamblia*; *Gs*TIM, TIM from *Bacillus stearothermophilus*; *Hs*TIM, TIM from *Homo sapiens*; *Lm*TIM, TIM from *Leishmania mexicana*; *Mm*TIM, TIM from *Moritella marina*; *Mt*TIM, TIM from *Mycobacterium tuberculosis*; *Np*TIM, TIM from *Nostoc punctiforme*; *Oc*TIM, TIM from *Oryctolagus cuniculus*; *Pf*TIM, TIM from *Plasmodium falciparum*; *Sc*TIM, TIM from *Streptomyces coelicolor*; *Ss*TIM, TIM from *Sus scrofa*; *Tb*TIM, TIM from *Trypanosoma brucei*; *Tc*TIM, TIM from *Trypanosoma cruzi*; *Ts*TIM, TIM from *Taenia solium*; *Tv*TIM, TIM from *Trichomonas vaginalis*; *Yt*TIM, TIM from *Saccharomyces cerevisiae*.

was compared, no differences in the number of hydrogen bonds or salt bridges between the monomers were found (Table S5, ESI†). The total surface buried upon association and its hydrophobic content were also similar (Table 4).

The only structural difference between the native state of RevTIMs and IrrevTIMs so far detected relates to cavities. RevTIMs show a lower number of cavities; their volume is approximately 40% lower than that of IrrevTIMs (Tables 3 and 4). The native structure of the three RevTIMs has cavity volume values outside the $\mu \pm \sigma$ interval determined for the IrrevTIMs (Table 3). The size of the biggest cavity, located at the interface between β -strands 5–7 and α -helices 5–6 (comprising residues 90–120, 145–160 and 195–200; Fig. S6, ESI†), is the main responsible of the cavity

volume difference between RevTIMs and IrrevTIMs (Table 4 and Fig. 7). There are no previous reports on the relation between cavity volume in the native state and reversibility in unfolding transitions; however, it has been suggested that internal cavity volume influences the stability and the unfolding kinetics of proteins.^{75,84,85}

(iii) Physicochemical properties of the unfolding/refolding transition. It is more likely to find reversible thermal unfolding in proteins that exhibit a low T_m because high temperatures promote the chemical modification of amino acid side chains. In agreement, *Dr*TIM and *Sc*TIM present some of the lower T_ms reported to date. It is worth noting that, however, T_m of *Cp*TIM is higher than eight TIMs that show irreversible thermal unfolding (Table 3). In fact, *Cp*TIM is the only RevTIM whose T_m is within the $\mu \pm \sigma$ interval determined for IrrevTIMs (Table 3). Regarding thermodynamic stability, ΔG_{diss} obtained from dilution experiments is very similar in both RevTIMs and IrrevTIMs (white bars in Fig. 8), and this suggests that the stability of the interface does not correlate with reversibility. Concerning the overall stability of the dimer, ΔG_{tot} for IrrevTIMs is higher than the ΔG_{tot} values of RevTIMs (compare orange and red bars in Fig. 8) and consequently, it is possible that reversibility is related to a lower ΔG_{tot} . It should be noted, however, that ΔG_{tot} for IrrevTIMs was obtained from chemical unfolding experiments whereas ΔG_{tot} for the reversible ones was obtained using temperature as a perturbant. Despite this, it is reasonable to consider a correlation between reversibility and low conformation stability.

The equilibrium thermal unfolding of all RevTIMs can be well described with a two state model (Fig. 3a–c and 4a–c), and most of the studied IrrevTIMs show a single cooperative transition described by a single T_m value (Table 3); hence, it seems that there are no stable and/or populated equilibrium intermediates involved. Nevertheless, it is likely that the aggregation properties of transient intermediates play a role in determining the reversibility of folding transitions. In this regard, the monomeric intermediate observed in the thermal unfolding of *Eh*TIM unfolds irreversibly.⁵¹ Likewise, evidence for kinetic intermediates was found in the highly irreversible temperature-induced unfolding of *Tc*TIM.⁸⁶ Consequently, it is worthwhile to explore a possible relation between two state unfolding and reversibility. At 25 °C, *Dr*TIM and *Sc*TIM show that $\Delta G_{\text{tot}} \approx \Delta G_{\text{diss}}$ whereas for *Cp*TIM ΔG_{diss} is $\approx 75\%$ of ΔG_{tot} (compare white and red bars in Fig. 8). ΔG_{tot} from temperature-induced unfolding experiments has not been obtained previously for IrrevTIMs; however when ΔG_{tot} has been obtained using urea or GndHCl as a perturbant, ΔG_{diss} was also found to be a major component of ΔG_{tot} (compare blue and orange bars in Fig. 8). It is worth stressing that in previous reports,^{56–60} reversibility has not been tested in the dilution experiments used to obtain ΔG_{diss} . This, and the different nature of chemical unfolding and dilution, may explain why these values are not always similar to ΔG_{diss} obtained from urea and GndHCl-induced unfolding experiments (compare white and blue bars in Fig. 8). However, it seems that irrespectively of the perturbant, the dissociation event is the main component of the conformational stability of TIM.^{24–38}

Table 4 Structural features used for comparison between IrrevTIMs and RevTIMs

Structural property	<i>Dr</i> TIM	<i>Sc</i> TIM	<i>Cp</i> TIM	Average RevTTMs	Average IrrevTTMs
Total ASA of the folded oligomer (\AA^2)	18 107	19 475	19 947	19 176 \pm 780	19 040 \pm 629
Hydrophobic ASA of the folded oligomer (\AA^2)	10 389	11 367	11 481	11 079 \pm 490	11 498 \pm 600
% Hydrophobicity ^a	57.4	58.4	57.6	57.8 \pm 0.4	60.4 \pm 1.3
Total Δ ASA buried on folding of the monomer (\AA^2)	25 222	26 550	26 679	26 151 \pm 658	27 056 \pm 740
Hydrophobic Δ ASA buried on folding of the monomer (\AA^2)	18 367	19 465	19 269	19 034 \pm 478	19 552 \pm 375
% Hydrophobicity ^b	72.8	73.3	72.2	72.8 \pm 0.5	72.3 \pm 1.9
Total Δ ASA of the folded monomer buried on assembly (\AA^2)	3113	3425	3467	3335 \pm 158	3275 \pm 207
Hydrophobic Δ ASA of the folded monomer buried on assembly (\AA^2)	2131	2286	2339	2252 \pm 89	2005 \pm 354
% Hydrophobicity ^c	68.5	66.7	67.5	67.6 \pm 0.7	61.2 \pm 2.8
Total Δ ASA of the unfolded monomer on assembly (\AA^2)	53 558	53 524	56 826	55 636 \pm 1474	57 388 \pm 1365
Hydrophobic Δ ASA of the unfolded monomer on assembly (\AA^2)	38 864	41 217	40 878	40 320 \pm 1038	41 110 \pm 835
% Hydrophobicity ^d	72.6	77.0	71.9	73.8 \pm 2.3	71.6 \pm 2.1
Δ ASA _{dissoc} / Δ ASA _{tot} (%)	5.8	6.4	6.1	6.1 \pm 0.3	5.7 \pm 0.4
Δ ASA _{dissoc hydrophob} / Δ ASA _{tot hydrophob} (%)	5.5	5.5	5.7	5.6 \pm 0.1	4.8 \pm 0.7
Molecular volume of the oligomer (\AA^3)	62 052	67 188	69 885	66 375 \pm 3249	67 339 \pm 2221
Number of cavities in oligomer (without solvent)	10	11	9	10 \pm 1	16 \pm 3
Total volume of cavities in oligomer (\AA^3) (without solvent)	2134	2921	2241	2432 \pm 349	3651 \pm 498
Total volume of the biggest cavity in the oligomer (\AA^3)	324	683	290	432 \pm 178	1001 \pm 209
Total volume of cavities in oligomer (\AA^3) (with solvent)	928	1335	1008	1090 \pm 176	1908 \pm 366

^a Errors are the standard deviation from averaged values of all structures in each group. % Hydrophobicity = (hydrophobic ASA of the oligomer/total ASA of the oligomer) \times 100%. ^b % Hydrophobicity = (hydrophobic Δ ASA buried on folding of the monomer/total Δ ASA buried on folding of the monomer) \times 100%. ^c % Hydrophobicity = (hydrophobic Δ ASA of the folded monomer buried on assembly/total Δ ASA of the folded monomer buried on assembly) \times 100%. ^d % Hydrophobicity = (hydrophobic Δ ASA of the unfolded monomer on assembly/total Δ ASA of the unfolded monomer on assembly) \times 100%.

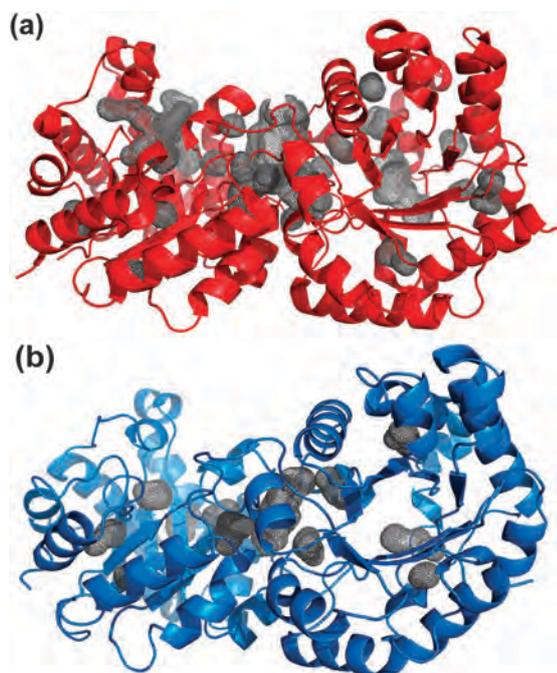


Fig. 7 Internal cavities in the three-dimensional structure of TIM. The structure and cavities are shown for: (a) GoTIM, an example of an IrrevTIM, and (b) CpTIM, an example of a RevTIM. Total volume of cavities in oligomers of GoTIM and CpTIM is 4010 \AA^3 and 2241 \AA^3 , respectively.

There are two extreme scenarios for a dissociation process; in the rigid body case, subunit dissociation takes place without the unfolding of the monomers; on the other extreme, dissociation is coupled to complete monomer unfolding (two state behaviour). Reversibility is likely related to the cooperativity of the folding process. For example, if TIM monomers were isolated from the

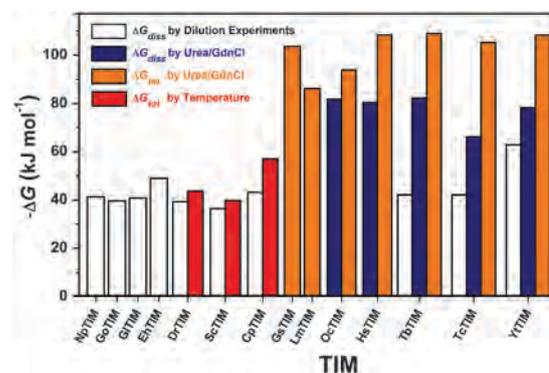


Fig. 8 Comparison of the reported free energies for different TIMs. Red and orange bars indicate ΔG_{tot} at 25 $^{\circ}\text{C}$, blue bars are ΔG_{diss} at 25 $^{\circ}\text{C}$ and white bars represent ΔG_{diss} values derived from dilution experiments previously reported^{56–60} and determined in this work. Data for HsTIM, GsTIM, OcTIM, LmTIM and YtTIM were obtained from urea or guanidinium hydrochloride experiments,^{23–37} whereas data for DrTIM, ScTIM and CpTIM are from the DSC experiments in this work.

native dimer in a rigid body scenario, each of them would expose a surface area of 1645 \AA^2 (Table 4), nearly 70% of this area is hydrophobic, increasing the chances of aggregation; therefore, the presence of stable monomeric intermediates would likely reduce reversibility. Analysis of the structural data for TIM from several species shows that the association of the monomers to form the native dimer accounts for \approx 5–6% of the overall Δ ASA calculated in going from the unfolded monomers to the folded dimer (Table 4). From structure-based parameterizations, we estimate that for the rigid body case, dissociation would contribute a similar 5–6% of the total ΔC_p (Table 2). In contrast, the available experimental data for the chaotrope-induced three state unfolding of TIM from several organisms

show that the dissociation step accounts for $\approx 75\%$ of the overall conformational stability of the dimer.³² Likewise, dissociation represents 55% of the total enthalpy change; this has led to the conclusion that dimer dissociation leads to extensive unfolding of the monomers, *i.e.* dissociation is not a rigid body process.⁵¹ In accordance, folding and association are strongly coupled in thermally unfolded *YrTIM*.⁴⁷ *DrTIM*, *ScTIM* and *CpTIM* show the extreme cooperative behaviour, that is, their thermal unfolding is a two state equilibrium process where isolated monomers are not stable at all. In agreement, ΔG_{diss} and ΔG_{tot} are similar for these proteins (Fig. 5 and 8). It is then tempting to correlate reversibility with two state behaviour. Nevertheless, it should be pointed out that the temperature-induced unfolding of many IrrevTIMs shows a single transition; therefore, extreme cooperativity does not by itself determine reversibility.

From the above discussion about the possible correlations between reversibility and the molecular level properties of TIM, it appears that there is a delicate balance between several contributions whose concerted interplay is necessary to achieve thermal reversibility in oligomeric enzymes. All the correlations between molecular level properties and reversibility proposed in this work were obtained from the analysis of the stability and structure of a limited number of proteins. Clearly, more quantitative data, such as the physicochemical characterization of TIMs from yet unexplored branches of the tree of life, are needed to further assess the findings in this work. A clear avenue to obtain a more detailed atomic description of the unfolding and dissociation of TIM is the use of molecular dynamics. This methodology is a valuable tool to evaluate intermolecular interactions at the finest level of detail.⁸⁷ Steered molecular dynamics have been successfully used for the estimation of protein–ligand⁸⁸ and protein–protein energetics;⁸⁹ the optimization of this and other methodologies for the analysis of protein–protein interactions is a current focus of research.⁹⁰ It is likely that the molecular dynamics approach, applied to the *in silico* unfolding of RevTIMs and IrrevTIMs, will be able to dissect the location of conformational changes taking place upon dimer dissociation and monomer unfolding, as well as the role of their coupling in reversibility. In addition, the kinetic characterization of the thermal unfolding of RevTIMs should bring more information regarding the unfolding pathway followed by these proteins.

3. Conclusions

The results presented in this work show that the thermal unfolding of a given protein architecture, such as the TIM barrel, can exhibit kinetic or thermodynamic control. *NpTIM* and *GoTIM* show kinetic control and their unfolding activation energies are in the low and high range of those determined for EukTIMs. It has been suggested that kinetic stability is more strongly selected by evolution than thermodynamic stability because the former facilitates irreversible alteration processes even when a high thermodynamic stability is maintained.⁹¹ Notably, *DrTIM*, *ScTIM* and *CpTIM* presented thermodynamic

control and reversible thermal unfolding. Stability and reversibility are biotechnologically relevant;^{92,93} however, the reasons underlying the reversibility of the unfolding process are not understood. In this work we found no correlation between this behaviour and the composition or the gross structural properties of TIM; however, low pI, the absence of residual structure in the unfolded state, a small cavity volume in the native state, low conformational stability and low T_m may be involved. It is very possible that reversibility is encoded in the folding pathway. In this respect, one of the main conclusions derived from the characterization of the folding mechanism of monomeric TIM barrels is the presence of intermediates that have been related to the folding of particular β/α modules within the barrel. In contrast, the dimeric TIM barrels studied in this work show two state thermal unfolding. It is therefore possible that extreme cooperativity and reversibility are related.

For all living organisms from bacteria to mammals, temperature is probably the most common and natural stress agent at the molecular level. In fact, life at “constant temperature” is a recent evolutionary event.^{94,95} Molecular mechanisms such as the chaperonin system have evolved to help the cells to deal with temperature stress.^{96–98} TIM has not been identified as a GroEL substrate, albeit TIM barrels are common substrates of the chaperonin system in *E. coli*.⁹⁹ Because oligomeric TIM barrels are frequent in cells, reversibility in their folding/unfolding process is an evolutive advantage because it reduces the chance of non-productive irreversible aggregation and the load of the cellular chaperonin machinery. That such a common topological ensemble can unfold and refold without the help of the cellular machinery is indeed biologically relevant. Furthermore, the finding that the three reversible proteins come from organisms from different phyla suggests that reversibility in thermal unfolding may be more common than what is currently known.

4. Materials and methods

4.1 Enzymes and biochemicals

α -Glycerol-phosphate dehydrogenase (GDH) was purchased from Roche. All other reagents were of analytical grade from Sigma-Aldrich. The water used was distilled and deionized.

4.2 Selection of species and sequences

From the phylogenetic tree published by Ciccarelli and cols¹⁰⁰ and the phylogenetic classification proposed by Cavalier-Smith,⁵⁴ we selected six representative species with evolutionary significance and biomedical or biotechnological importance belonging to six different phyla (different bacteria paraphyletic supertaxa; Fig. S1, ESI \ddagger): *Deinococcus radiodurans* (*DrTIM*; phylum: Deinococcus-Thermus), *Nostoc punctiforme* (*NpTIM*; phylum: Cyanobacteria), *Gemmata obscuriglobus* (*GoTIM*; phylum: Planctomycetes), *Azotobacter vinelandii* (*AvTIM*; phylum: Proteobacteria), *Clostridium perfringens* (*CpTIM*; phylum: Firmicutes), and *Streptomyces coelicolor* (*ScTIM*; phylum: Actinobacteria). TIM sequences of these species were extracted from GenBank⁶ and annotated with their identification number: *DrTIM* (GI: 653293168), *NpTIM*

(GI: 186684444), *Go*TIM (GI: 497733394), *Av*TIM (GI: 226946590), *Cp*TIM (GI: 18310284), and *Sc*TIM (GI: 490074698).

4.3 Cloning, expression, and purification of BacTIMs

The nucleotide sequence of the gene coding for BacTIMs was optimized for their expression in *E. coli*, synthesized and cloned in the pBluescript-II-SK(-) plasmid (Agilent Technologies) by Epoch Life Science (Missouri City, TX-USA). BacTIM genes were subcloned into the pET28b(+) vector (Novagen) using the terminal restriction sites for *Nde*I and *Xho*I. The genes were expressed in *E. coli* strain BL21(DE3)p*Lys*S (Invitrogen) transformed with the BacTIM-pET28b(+) plasmids. To this end, 1 L of the Luria broth (LB) medium supplemented with 30 mg mL⁻¹ kanamycin was inoculated with a 1 mL preculture and incubated at 37 °C and 250 rpm. After an OD₆₀₀ of 0.6–0.8 was reached, overexpression was induced by adding 0.8 mM IPTG; growth was continued by 16 hours. After incubation, cells were harvested by centrifugation (Thermo/SLA-3000, 15 min, 8000 rpm, 4 °C), suspended in buffer A (35 mM NaH₂PO₄, 300 mM NaCl and 5 mM imidazole, pH 8.0), lysed by sonication (Cole Parmer Ultrasonic Processor, 10 cycles in 45 s intervals, 30% pulse, 0 °C) and centrifuged again (Sorvall/SS-34, 20 min, 13 000 rpm, 4 °C). Affinity chromatography was then performed using a Waters AP-2 column containing 30 mL of Protino-NiTED resin (Macherey-Nagel), coupled to an Akta system (GE Healthcare). The supernatant was loaded into the column previously equilibrated with buffer A and the nonbound fraction was washed with 110 mL of buffer A. Bound protein was eluted with a linear gradient of 5–500 mM Imidazole. BacTIM-containing fractions were pooled and dialyzed against 50 mM tris (hydroxymethyl) aminomethane (TRIS), 10 mM CaCl₂ (pH 8.0). His-tag was removed with the Thrombin CleanCleave kit (Sigma-Aldrich) using the manufacturer's protocol. Cleaved and not cleaved proteins were separated by a second affinity chromatography step. Unbound fractions were pooled, dialyzed against buffer B (10 mM triethanolamine (TEA), 1 mM ethylenediaminetetraacetate (EDTA), and 1 mM dithiothreitol (DTT), pH 8.0), and finally purified using a MonoQ column HR 10/10 (GE Healthcare). The sample was eluted with a linear gradient of 0–500 mM NaCl in buffer B. BacTIMs, 99% pure according to SDS-PAGE (12.5% acrylamide), were stored at 4 °C until use. After His-tag removal, ≈35 milligrams of protein were obtained per liter of culture. We were unable to express *Av*TIM in pET28b(+). Although several vectors were assayed, expression was only observed in pET22a. Irrespectively of the expression conditions, the protein was always found in inclusion bodies, numerous refolding protocols were assayed without success. Interestingly, of the six sequences studied in this work, the *Av*TIM sequence was the only predicted by the stability index as an unstable protein.¹⁰¹ Some measurements were carried out in 10 mM phosphate pH 8.0 (buffer C). Protein concentration, estimated from the absorbance at 280 nm (and the absorption coefficient calculated from the amino acid sequence) or using the Bicinchoninic acid assay, gave similar results.

4.4 Activity assays

Catalytic activity in the GAP to DHAP direction was determined by a coupled assay with GDH.⁵¹ The reaction cell contained

1 mL of buffer D (100 mM TEA, 10 mM EDTA and 1 mM DTT, pH 8.0) with 2 mM GAP (except when catalytic constants were determined), 20 µg of GDH, and 0.2 mM NADH. Catalytic assays started with 0.27 nM BacTIM. NADH absorbance changes at 340 nm were followed in a Beckman DU7500 spectrophotometer at 25 °C. For the determination of catalytic constants, the concentration of GAP ranged between 0.1 and 3 mM. Changes in NADH absorbance were linear with time, indicating that no dimer dissociation took place during activity assays.

4.5 Hydrodynamic measurements

Size-exclusion-chromatography experiments were performed on a Superdex 75 HR 10/30 column, equilibrated with buffer B, coupled to an Akta System (GE Healthcare). The elution volume was independent of protein concentration in the tested range (0.4–20 µM). Stokes-radii were calculated from elution volumes and a calibration curve constructed using six different proteins (albumin, ovalbumin, carbonic anhydrase, lysozyme, ribonuclease and cytochrome C).¹⁰²

4.6 Stability to dilution experiments

The stability of the dimer was estimated from dilution experiments. BacTIMs were incubated for 24 hours at several temperatures (15 to 35 °C) and different concentrations (0.4 nM to 18 µM) in buffer D. Subsequently activity was determined at the incubation temperature with 0.27 nM BacTIM. For reversibility tests 1.5 L of 2 nM BacTIM were incubated at 25 °C for 24 hours and then concentrated to 4 µM (≈0.75 mL) using Millipore Centricon centrifugal filter devices (10 000 MWCO). Thereafter catalytic activity was measured and this was normalized (f_N) with the bigger activity value obtained.

4.7 Spectroscopic properties

Circular dichroism (CD) experiments were carried out in buffer C in a Chirascan Spectropolarimeter (Applied Photophysics; Leatherhead, Surrey-UK) equipped with a Peltier device. CD spectra (180–250 nm wavelength range) were obtained with 20 µM BacTIM at 25 °C. Fluorescence measurements were made on a PC1 ISS Spectrofluorometer (Champaign, IL-USA) equipped with a Peltier and a water-jacketed cell holder for temperature control. The raw data were converted to molar ellipticity ($[\theta]$) using the formula:¹⁰³ $[\theta] = \theta/(l \times C \times N_r)$ where θ is ellipticity in millidegrees, l is the cell path length in millimeters, C is the molar concentration of protein and N_r is the number of residues. Intrinsic fluorescence spectra of 8 µM BacTIM were determined in buffer C at 25 °C with 2 nm bandwidth slits with excitation at 295 nm and emission in the 310–410 nm range. The wavelength of maximal emission (λ_{max}) and fluorescence spectral center of mass (SCM) were calculated from intensity data (I_i) obtained at different wavelengths: $\text{SCM} = \sum \lambda_i I_i / \sum I_i$.

4.8 Thermal transitions monitored by circular dichroism (CD) spectroscopy

Thermal unfolding and refolding transitions were followed using buffer C by monitoring ellipticity at 195 and 222 nm as a function of temperature at a heating rate of 1.0 K min⁻¹.

Actual temperatures within the cell were registered with the external probe of the cell holder. Protein concentration ranged from 0.4 to 15 μM . 1.0 or 0.1 cm path-length cells were used accordingly. The changes in the CD signal were analyzed after normalization of the transition curves to the fraction of unfolded molecules (f_D) by: $f_D = [y_{\text{obs}} - (y_N + m_N T)] / [(y_D + m_D T) - (y_N + m_N T)]$, where y_{obs} is the experimental observed DC signal at a given temperature and $(y_N + m_N T)$ and $(y_D + m_D T)$ are the fitting equations to straight lines representing the pre and post transition regions (native and unfolded states), respectively.

4.9 Thermal transitions monitored by differential scanning calorimetry (DSC)

DSC experiments were done in a VP-Capillary DSC system (MicroCal, GE Healthcare). Protein concentration varied from 8 to 80 μM , and scan rates from 0.5 to 3.0 K min^{-1} . Protein solutions were prepared by exhaustive dialysis in buffer C and then degassed at room temperature. Buffer–buffer traces were subtracted from sample endotherms. For all proteins, a reheating run was carried out to determine the reversibility or irreversibility of the process. To verify that irreversibility was not the result of a too high final scanning temperature, first scans were also performed heating near the T_m . When irreversibility was observed, DSC traces were also obtained in the presence of urea (up to 2 M, where activity assays indicated the proteins remain in its dimer native state). Urea concentrations were determined from refractive index measurements. The Origin software package (MicroCal) was used for data analysis, calculation of unfolding enthalpies and ΔC_p determination.

4.10 Crystallization and data collection

All proteins were dialyzed in buffer B containing 100 mM NaCl. The hanging-drop vapor diffusion method was used to screen the crystallization conditions contained in the HR Crystal Screen I and II kits (Hampton Research). Crystals were obtained at 18 °C after incubating drops of proteins with the precipitant solutions in a 1 : 1 ratio. Crystallization conditions for the four BacTIMs were: for *Go*TIM 0.2 M potassium sodium tartrate, 0.1 M sodium citrate tribasic pH 5.6 and 2 M ammonium sulfate; for *Dr*TIM 1.8 M ammonium sulfate, 0.1 M sodium citrate tribasic pH 5.6 and 0.2 M potassium sodium tartrate; for *Sc*TIM 0.1 M TRIS hydrochloride pH 8.5 and 2.0 M ammonium sulfate; and for *Cp*TIM 0.2 M ammonium acetate, 0.1 M sodium acetate pH 4.6 and 30% w/v polyethylene glycol (PEG) 4000. For crystals obtained under conditions without PEG, a 30% w/v glycerol solution was used as a cryoprotectant. Diffraction data were collected at 100 K with an oscillation of 0.5° per frame on an R-Axis IV++ image plate detector (Rigaku, The Woodlands, TX) with X-rays (wavelength: 1.5418 Å) generated by a Rigaku MicroMax-007 HF rotating anode (Rigaku, The Woodlands, TX). All data sets were integrated using XDS¹⁰⁴ and scaled with SCALA in the CCP4 program suite v.6.1.2.¹⁰⁵

4.11 Structure determination and refinement

All structures were solved by the molecular replacement method with PHASER in the PHENIX software suite.¹⁰⁶ *Thermotoga maritima*

(PDB ID: 1B9B), *Thermus thermophilus* TIM (PDB ID: 1YYA), *Mycobacterium tuberculosis* TIM (PDB ID: 3TA6) and *Geobacillus stearothermophilus* TIM (PDB ID: 1BTM) were used as the starting model for molecular replacement phasing of *Go*TIM (TFZ: 42.3 and log-likelihood gain: 1360.2), *Dr*TIM (TFZ: 32.8 and LLG: 1104.3), *Sc*TIM (TFZ: 45.6 and LLG: 3064.96) and *Cp*TIM (TFZ: 21.3 and LLG: 755.8), respectively. For all the structures, refinement was done with *phenix.refine*¹⁰⁶ and the models were improved by iteratively model rebuilding in COOT.¹⁰⁷ The final tridimensional structures were validated with MolProbity¹⁰⁸ and using the Protein Data Bank (PDB) validation server.⁷ The coordinates and structure factors of *Go*TIM, *Dr*TIM, *Sc*TIM and *Cp*TIM have been deposited in the PDB under accession codes 4Y96, 4Y90, 4Y9A and 4Y8F, respectively. Table S2, ESI† summarizes data collection and refinement statistics. The figures were produced using the PyMOL Molecular Graphics System, v1.7.2.¹⁰⁹

4.12 Sequence and structural analysis

The sequence alignment was performed with MAFFT¹¹⁰ and visually modified with Jalview.¹¹¹ Sequence identity percentage was calculated with SIAS server.¹¹² Nineteen dimeric structures for which their thermal unfolding have been studied (including those studied in this work) were selected for a structural comparison; the characteristics of these three-dimensional structures are summarized in Table S6, ESI†. Structural alignments were performed using UCSF Chimera¹¹³ and PyMOL Molecular Graphics System, v1.7.2.¹⁰⁹ RMSD values were calculated with a distance matrix using GRASP.¹¹⁴ The assignment and composition of secondary structure elements were determined with STRIDE.¹¹⁵ Molecular volumes and cavities were calculated with the 3V software suite¹¹⁶ and MOLE v2.13.9.6 software¹¹⁷ using a random exploration path and a probe radius of 1.5 Å. A cavity was defined as an empty space buried inside the protein structure. The accessible surface area (ASA) and the difference in accessible surface area upon protein folding and assembly were calculated with PDBEPIA v1.51,¹¹⁸ Travel Depth¹¹⁹ and NACCESS software¹²⁰ using a probe radius of 1.5 Å. Changes in ASA upon dissociation were calculated for a rigid body dissociation. For the unfolded state, accessibility was calculated using an extended Gly-X-Gly peptide with VADAR.¹²¹ Hydrogen bonds between protein atoms were calculated using the HBPLUS¹²² and VADAR routines with their default parameters for distances and angles. When two atoms of opposite charge were observed within 4 Å they were assigned to a salt bridge.

4.13 Data treatment and fitting

(i) **Dissociation constant from the stability to dilution experiments.** To obtain the dissociation constant for the dimer, K_{diss} , the normalized catalytic activity, $f_N(T)$, was plotted against protein concentration, P_c , and fitted to:

$$f_N(T) = \frac{K_{\text{diss}}(T) + 4P_c - \sqrt{K_{\text{diss}}(T)^2 + 8K_{\text{diss}}(T)P_c}}{4P_c} \quad (1)$$

(ii) **Activation energy from the irreversible DSC transitions and kinetic and equilibrium urea m values.** The shape of the calorimetric transitions is adequately described by the two-state

irreversible model ($N \rightarrow F$)^{123,124} where N is the dimeric native protein, F is the final state (unable to fold back to the native protein). The kinetic conversion from N to F is described by a first-order rate constant (k) that changes with temperature according to the Arrhenius equation:

$$k = \exp\left[-\frac{E_A}{R}\left(\frac{1}{T} - \frac{1}{T^*}\right)\right] \quad (2)$$

where T^* is the temperature at which $k = 1 \text{ min}^{-1}$, and E_A is the activation energy from the native state to the transition state. The apparent heat capacity (corrected for the chemical baseline) in the two-state irreversible model is given by:

$$C_p^{\text{APP}} = \frac{\Delta H E_A}{RT_m^2} \exp(x) \times \exp[-\exp(x)] \quad (3)$$

$$x = \frac{E_A}{RT_m^2}(T - T_m) \quad (4)$$

where T_m is the temperature corresponding to the maximum of the transition, T is the temperature at each point of the endotherm and ΔH is the unfolding enthalpy (consider a constant). The activation energies were obtained following three procedures. In the first, the DSC traces were fitted using eqn (3) and (4) at each of the employed scan rates, producing individual E_A values (one for each scan rate). In the second procedure, E_A is obtained from the slope of Arrhenius plots, *i.e.* $\ln k$ vs. $1/T$, employing all scan rates. Finally, E_A can also be derived from a data consistency test, evaluating the effect of the scanning rate (ν) on T_m :⁵²

$$\ln\left(\frac{\nu}{T_m^2}\right) = \text{constant} - \frac{E_A}{RT_m} \quad (5)$$

Kinetic urea m values (m^\ddagger) were used as a measure of exposure to the solvent (and, consequently, degree of unfolding) in the transition state.¹²⁵ They were obtained using:

$$m^\ddagger = -\frac{E_A}{T_m} \left(\frac{dT_m}{d[\text{urea}]}\right) - RT_m \left(\frac{d \ln\left(\frac{E_A}{RT_m}\right)}{d[\text{urea}]}\right) \quad (6)$$

The equilibrium m values (m_{eq}) were calculated using a correlation between m_{eq} and ΔASA .⁷³

(iii) Enthalpy changes from the reversible CD and DSC transitions. The enthalpies from the native state to unfolded state, ΔH_{vH} , were obtained from a fit of the temperature dependence of the equilibrium constant within the two-state reversible model,¹²⁶ $K_{\text{eq}} = (2f_{\text{D}}2P_{\text{c}})/(1 - f_{\text{D}})$, using the protein unfolded fraction (f_{D}) from CD data:

$$f_{\text{D}} = \frac{(y_{\text{N}} + m_{\text{N}}T) + (y_{\text{D}} + m_{\text{D}}T) \exp\left[\frac{\Delta H_{\text{vH}}}{R}\left(\frac{1}{T} - \frac{1}{T_m}\right)\right]}{1 + \exp\left[\frac{\Delta H_{\text{vH}}}{R}\left(\frac{1}{T} - \frac{1}{T_m}\right)\right]} \quad (7)$$

where T_m is the temperature corresponding to the 50% native and unfolded states.

The DSC endotherms were fitted to the equilibrium two-state model with dimer dissociation using (Microcal DSC data analysis tutorial guide v5.0):

$$C_p(T) = B_0 + B_1T + f(T)\Delta C_p + \frac{\Delta H_{\text{cal}}(T)}{RT_m^2} \left[\frac{1 - f(T)}{-1 + \frac{2}{f(T)}} \right] \quad (8)$$

where B_0 and B_1 are constants and $f(T)$ is the protein fraction in the monomer state, producing ΔH_{cal} , ΔC_p , and T_m . The van't Hoff enthalpy $\Delta H_{\text{vH}} = 4RT_{1/2}^2 C_{p1/2} / \Delta H_{\text{cal}}$ with $T_{1/2}$ being the temperature where the fraction of native and unfolded proteins is 50%, and $C_{p1/2}$ the experimental heat capacity value at that temperature.

The enthalpy change $\Delta H'$ was also calculated from the dependence of T_m on protein concentration according to the Takahashi-Sturtevant equation:¹²⁷

$$\ln P_c = b - \frac{\Delta H'}{RT_m} \quad (9)$$

with b being a constant, and T_m is the temperature corresponding to the 50% native and unfolded states.

(iv) Experimental, fitted and predicted heat capacity changes from the reversible DSC transitions. The experimental heat capacities for each protein concentration (ten of them) were converted into absolute heat capacities (Microcal DSC data analysis tutorial guide v5.0). The absolute heat capacity for the native state (C_{pN}) and unfolded state (C_{pD}) were fitted to straight lines and $\Delta C_p = C_{\text{pD}} - C_{\text{pN}}$ evaluated at T_m (and then averaged). ΔC_p was also obtained as a fitted parameter using the equilibrium two-state model with dimer dissociation (eqn (8)). Finally, predicted ΔC_p values can be obtained from reported parametric equations, namely correlations based on the amino acidic sequence and heat capacity changes,^{68,69} and correlations between ΔASA and heat capacity changes.⁷⁰⁻⁷⁴

(v) Global stability curve. Stability curves, $\Delta G(T)$, were calculated using $\Delta H(\Delta H_{\text{cal}})$, ΔC_p and T_m obtained from eqn (8) and the Gibbs-Helmholtz equation:^{128,129}

$$\Delta G(T) = \Delta H \left(1 - \frac{T}{T_m}\right) - \Delta C_p \left(T_m - T + T \ln\left(\frac{T}{T_m}\right)\right) \quad (10)$$

Maximal stability occurs at a temperature T_s where $d\Delta G/dT = 0 = \Delta S$ and $\Delta H_{\text{TS}} = \Delta G(T_s)$.

Abbreviations

BacTIM	Bacterial TIMs
CD	Circular dichroism
C_p TIM	TIM from <i>Clostridium perfringens</i>
D_r TIM	TIM from <i>Deinococcus radiodurans</i>
DSC	Differential scanning calorimetry
E_A	Activation energy
EukTIMs	Eukaryotic TIMs
Go TIM	TIM from <i>Gemmata obscuriglobus</i>
Gnd-HCl	Guanidinium hydrochloride
IrrevTIMs	TIMs with irreversible unfolding

<i>Np</i> TIM	TIM from <i>Nostoc punctiforme</i>
RevTIMs	TIMs with reversible unfolding
<i>Sc</i> TIM	TIM from <i>Streptomyces coelicolor</i>
TIM	Triosephosphate isomerase
T_m	Midpoint of thermal unfolding

Conflict of interest

The authors declare that they have no conflict of interest.

Note added in proof

In our description of monomeric TIM barrels with reversible thermal unfolding, we missed α -amylase, a system where an excellent discussion about protein adaptation to extreme temperatures has been published.¹³⁰

Acknowledgements

This work was supported by Consejo Nacional de Ciencia y Tecnología de México (Grant number 166472 to ARR, and 99857 to DAFV), Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica DGAPA-UNAM (PAPIIT grant number IN-112813 to MC, IN-207613 to ARR, and IN-219913 to DAFV), Facultad de Química and Facultad de Medicina, UNAM. SRR thanks Consejo Nacional de Ciencia y Tecnología de México (Grants numbers 313076 and 355134) and Posgrado en Ciencias Bioquímicas, UNAM, for financial support. We thank María Isabel Velázquez López and Georgina Espinosa Pérez for their competent technical support. X-Ray data were collected at the LANEM-IQ, UNAM. SRR MC ARR DAFV conceived and designed the experiments, SRR performed the experiments, SRR MC ARR DAFV analyzed the data, MC ARR DAFV contributed reagents/materials/analysis tools, SRR MC ARR DAFV wrote the paper, SRR MC DAFV performed the DSC experiments, SRR ARR performed the crystallization and structural determination, DAFV is the project director. All authors have given approval to the final version of the manuscript.

Notes and references

- K. A. Dill and J. L. MacCallum, *Science*, 2012, **338**, 1042–1046.
- H. Feng, Z. Zhou and Y. Bai, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 5026–5031.
- N. Koga, R. Tatsumi-Koga, G. Liu, R. Xiao, T. B. Acton, G. T. Montelione and D. Baker, *Nature*, 2012, **491**, 222–227.
- P. S. Huang, G. Oberdorfer, C. Xu, X. Y. Pei, B. L. Nannenga, J. M. Rogers, F. DiMaio, T. Gonen, B. Luisi and D. Baker, *Science*, 2014, **346**, 481–485.
- D. A. Benson, M. Cavanaugh, K. Clark, I. Karsch-Mizrachi, D. J. Lipman, J. Ostell and E. W. Sayers, *Nucleic Acids Res.*, 2013, **41**, D36–D42.
- H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov and P. E. Bourne, *Nucleic Acids Res.*, 2000, **28**, 235–245.
- M. D. Kumar, K. Abdulla-Bava, M. Michael-Gromiha, P. Prabakaran, K. Kitajima, H. Uedaira and A. Sarai, *Nucleic Acids Res.*, 2005, **34**, D204–D206.
- J. M. Scholtz, G. R. Grimsley and C. N. Pace, *Methods Enzymol.*, 2009, **466**, 549–565.
- P. L. Privalov, *Adv. Protein Chem.*, 1979, **33**, 167–241.
- J. M. Sanchez-Ruiz, *Annu. Rev. Phys. Chem.*, 2011, **62**, 231–255.
- C. M. Doyle, J. A. Rumfeldt, H. R. Broom, A. Broom, P. B. Stathopoulos, K. A. Vassall, J. J. Almey and E. M. Meiering, *Arch. Biochem. Biophys.*, 2013, **531**, 44–64.
- E. Braselman, J. L. Chaney and P. L. Clark, *Trends Biochem. Sci.*, 2013, **38**, 337–344.
- D. S. Goodsell and A. J. Olson, *Annu. Rev. Biophys. Biomol. Struct.*, 2000, **29**, 105–153.
- R. Sterner and B. Höcker, *Chem. Rev.*, 2005, **105**, 4038–4055.
- R. K. Wierenga, *FEBS Lett.*, 2001, **492**, 193–198.
- N. Nagano, C. A. Orengo and J. M. Thornton, *J. Mol. Biol.*, 2002, **321**, 741–765.
- W. R. Forsyth and C. R. Matthews, *J. Mol. Biol.*, 2002, **320**, 1119–1133.
- L. Carstensen, J. M. Sperl, M. Bocola, F. List, F. X. Schmid and R. Sterner, *J. Am. Chem. Soc.*, 2012, **134**, 12786–12791.
- B. N. Gangadhara, J. M. Laine, S. V. Kathuria, F. Massi and C. R. Matthews, *J. Mol. Biol.*, 2013, **425**, 1065–1081.
- R. Rudolph, R. Siebendritt and T. Kiefhaber, *Protein Sci.*, 1992, **1**, 654–666.
- A. A. Moosavi-Movahedi, B. Samiee and G. H. Hakimelahi, *J. Colloid Interface Sci.*, 1993, **161**, 53–56.
- S. Pyrpassopoulos, M. Vlassi, A. Tsortos, Y. Papanikolaou, K. Petratos, C. E. Vorgias and G. Nounesis, *Proteins*, 2006, **64**, 513–523.
- V. Mainfroid, S. C. Mande, W. G. J. Hol, J. Martial and K. Goraj, *Biochemistry*, 1996, **35**, 4110–4117.
- A. W. Rietveld and S. T. Ferreira, *Biochemistry*, 1996, **35**, 7743–7751.
- N. Beaucamp, A. Hofmann, B. Kellerer and R. Jeanicke, *Protein Sci.*, 1997, **6**, 2159–2165.
- A. W. Rietveld and S. T. Ferreira, *Biochemistry*, 1998, **37**, 933–937.
- R. S. Gokhale, S. S. Ray, H. Balaram and P. Balaram, *Biochemistry*, 1999, **38**, 423–431.
- A. M. Lambeir, J. Backmann, J. Ruiz-Sanz, V. Filimonov, J. E. Nielsen, I. Kursula, B. V. Norledge and R. K. Wierenga, *Eur. J. Biochem.*, 2000, **267**, 2516–2524.
- C. J. Morgan, D. K. Wilkins, L. J. Smith, Y. Kawata and C. M. Dobson, *J. Mol. Biol.*, 2000, **300**, 11–16.
- M. E. Cháñez-Cárdenas, D. A. Fernández-Velasco, E. Vázquez-Contreras, R. Coria-Ortega, G. Saab-Rincón and R. Pérez-Montfort, *Arch. Biochem. Biophys.*, 2002, **399**, 117–129.
- V. H. Moreau, A. W. Rietveld and S. T. Ferreira, *Biochemistry*, 2003, **42**, 14831–14837.
- H. Nájera, M. Costas and D. A. Fernández-Velasco, *Biochem. J.*, 2003, **370**, 785–792.

- 33 H. Pan, A. S. Raza and D. L. Smith, *J. Mol. Biol.*, 2004, **336**, 1251–1263.
- 34 M. E. Cháñez-Cárdenas, G. Pérez-Hernández, B. G. Sánchez-Rebollar, M. Costas and E. Vázquez-Contreras, *Biochemistry*, 2005, **44**, 10883–10892.
- 35 A. R. Vázquez-Pérez and D. A. Fernández-Velasco, *Biochemistry*, 2007, **46**, 8624–8633.
- 36 Y. Shi, J. H. Liu, H. J. Zhang and Y. Ding, *Protein Pept. Lett.*, 2008, **15**, 365–370.
- 37 V. Guzman-Luna and G. Garza-Ramos, *Proteins*, 2012, **80**, 1669–1682.
- 38 S. Lara-González, P. Estrella-Hernández, A. Ochoa-Leyva, M. Del Carmen Portillo-Téllez, L. A. Caro-Gómez, E. E. Figueroa-Angulo, H. Salgado-Lugo, J. F. T. Miranda-Ozuna, J. Ortega-López, R. Arroyo, L. G. Brieba and C. B. Benítez-Cardoza, *Proteins*, 2014, **82**, 22–33.
- 39 J. R. Knowles, *Nature*, 1991, **350**, 121–124.
- 40 J. P. Richard, *Biochemistry*, 2012, **51**, 2652–2661.
- 41 D. Maes, J. P. Zeelen, N. Thanki, N. Beaucamp, M. Alvarez, M. H. Thi, J. Backmann, J. A. Martial, L. Wyns, R. Jaenicke and R. K. Wierenga, *Proteins*, 1999, **37**, 441–453.
- 42 M. Peimbert, L. Domínguez-Ramírez, L. A. Téllez, L. M. Blancas and D. A. Fernández Velasco, in *Advances in Protein Physical Chemistry*, ed. E. García-Hernández and D. A. Fernández-Velasco, Transworld Research Network, Kerala, 1st edn, 2008, ch. 11, vol. 1, pp. 193–217.
- 43 A. Sun, Q. K. U. Yuksel and R. W. Gracy, *Arch. Biochem. Biophys.*, 1995, **322**, 361–368.
- 44 M. Alvarez, J. P. Zeelen, V. Mainfroid, F. Rentier-Delrue, J. A. Martial, L. Wyns, R. K. Wierenga and D. Maes, *J. Biol. Chem.*, 1998, **273**, 2199–2206.
- 45 M. Alvarez, J. Wouters, D. Maesi, V. Mainfroid, F. Rentier-Delrue, L. Wyns, E. Depiereux and J. A. Martial, *J. Biol. Chem.*, 1999, **274**, 19181–19187.
- 46 S. S. Ray, H. Balaram and P. Balaram, *Chem. Biol.*, 1999, **6**, 625–637.
- 47 C. G. Benítez-Cardoza, A. Rojo-Domínguez and A. Hernández-Arana, *Biochemistry*, 2001, **40**, 9049–9058.
- 48 L. Jiménez, D. A. Fernández-Velasco, K. Willms and A. Landa, *J. Parasitol.*, 2003, **89**, 209–214.
- 49 D. Mathur, G. Malik and L. C. Garg, *FEMS Microbiol. Lett.*, 2006, **263**, 229–235.
- 50 S. Enríquez-Flores, A. Rodríguez-Romero, G. Hernández-Alcántara, J. Oria-Hernández, P. Gutiérrez-Castrellón, G. Pérez-Hernández, I. de la Mora-de la Mora, A. Castillo-Villanueva, I. García Torres, S. T. Méndez, S. Gómez-Manzo, A. Torres-Arroyo, G. López-Velázquez and H. Reyes-Vivas, *Proteins*, 2007, **79**, 2711–2724.
- 51 L. A. Téllez, L. M. Blancas-Mejía, E. Carrillo-Nava, G. Mendoza-Hernández, D. A. Cisneros and D. A. Fernández-Velasco, *Biochemistry*, 2008, **47**, 11665–11673.
- 52 M. Costas, D. Rodríguez-Larrea, L. De Maria, T. V. Borchert, A. Gómez-Poyou and J. M. Sanchez-Ruiz, *J. Mol. Biol.*, 2009, **385**, 924–937.
- 53 Y. Aguirre, N. Cabrera, B. Aguirre, R. Pérez-Montfort, A. Hernandez-Santoyo, H. Reyes-Vivas, S. Enríquez-Flores, M. T. de Gómez-Puyou, A. Gómez-Puyou, J. M. Sanchez-Ruiz and M. Costas, *Proteins*, 2014, **82**, 323–335.
- 54 T. Cavalier-Smith, *Biol. Direct*, 2006, **1**, 1–19.
- 55 J. R. Lakowicz, *Principles of Fluorescence Spectroscopy*, Springer, New York, 3rd edn, 2006, ch. 16, vol. 1, pp. 529–575.
- 56 V. Mainfroid, P. Terpstra, M. Beaugard, J. M. Frère, S. C. Mande, W. G. J. Hol, J. A. Martial and K. Goraj, *J. Mol. Biol.*, 1996, **257**, 441–456.
- 57 A. Landa, A. Rojo-Domínguez, L. Jiménez and D. A. Fernández-Velasco, *Eur. J. Biochem.*, 1997, **247**, 348–355.
- 58 H. Reyes-Vivas, E. Martínez-Martínez, G. Mendoza-Hernández, G. López-Velázquez, R. Pérez-Montfort, M. T. de Gómez-Puyou and A. Gómez-Puyou, *Proteins*, 2002, **48**, 580–590.
- 59 G. López-Velázquez, D. Molina-Ortíz, N. Cabrera, G. Hernández-Alcántara, J. Peón-Peralta, L. Yépez-Mulia, R. Pérez-Montfort and H. Reyes-Vivas, *Proteins*, 2004, **55**, 824–834.
- 60 H. Reyes-Vivas, A. Díaz, J. Peón, G. Mendoza-Hernández, G. Hernández-Alcántara, I. de la Mora-de la Mora, S. Enríquez-Flores, L. Domínguez-Ramírez and G. López-Velázquez, *J. Mol. Biol.*, 2007, **365**, 752–763.
- 61 D. R. Trentham, C. H. McMurray and C. I. Pogson, *Biochem. J.*, 1969, **114**, 19–24.
- 62 G. Böhm, R. Muhr and R. Jaenicke, *Protein Eng.*, 1992, **5**, 191–195.
- 63 N. Sreerama and R. W. Woody, *Methods Enzymol.*, 2004, **383**, 318–351.
- 64 L. Whitmore and B. A. Wallace, *Biopolymers*, 2008, **89**, 392–400.
- 65 C. Vita, A. Fontana, J. R. Seeman and I. M. Chaiken, *Biochemistry*, 1979, **18**, 3023–3031.
- 66 L. López-Arenas, S. Solís-Mendiola and A. Hernández-Arana, *Biochemistry*, 1999, **38**, 15936–15943.
- 67 E. Freire, *Comments Mol. Cell. Biophys.*, 1989, **6**, 123–140.
- 68 D. Milardi, C. la Rosa, S. Fasone and D. Grasso, *Biophys. Chem.*, 1997, **69**, 43–51.
- 69 A. D. Robertson and K. P. Murphy, *Chem. Rev.*, 1997, **97**, 1251–1267.
- 70 K. P. Murphy and E. Freire, *Adv. Protein Chem.*, 1992, **43**, 313–361.
- 71 R. S. Spolar and M. T. Record, *Science*, 1994, **263**, 777–784.
- 72 G. I. Makhatadze and P. L. Privalov, *Adv. Protein Chem.*, 1995, **47**, 307–425.
- 73 J. K. Myers, C. N. Pace and J. M. Scholtz, *Protein Sci.*, 1995, **4**, 2138–2148.
- 74 B. Madan and K. A. Sharp, *Biophys. J.*, 2001, **81**, 1881–1887.
- 75 D. C. Rees and A. D. Robertson, *Protein Sci.*, 2001, **10**, 1187–1194.
- 76 K. M. Hart, M. J. Harms, B. H. Schmidt, C. Elya, J. W. Thornton and S. Marqusee, *PLoS Biol.*, 2014, **12**, 1–12.
- 77 R. G. E. Murray, in *The prokaryotes*, ed. A. Balows, H. G. Trüper, M. Dworkin, W. Harder and K. H. Schleifer, Springer-Verlag, New York, 2nd edn, 1992, ch. 204, vol. 4, pp. 3732–3744.
- 78 J. Bursy, A. U. Kuhlmann, M. Pittelkow, H. Hartmann, M. Jebbar, A. J. Pierik and E. Bremer, *Appl. Environ. Microbiol.*, 2008, **74**, 7286–7296.
- 79 J. Li and B. A. McClane, *Appl. Environ. Microbiol.*, 2006, **72**, 4561–4568.

- 80 E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M. R. Wilkins, R. D. Appel and A. Bairoch, in *The Proteomics Protocols Handbook*, ed. J. M. Walker, Humana Press, New York, 1st edn, 2005, ch. 52, vol. 1, pp. 571–607.
- 81 V. D. Trivedi, B. Raman, C. M. Rao and T. Ramakrishna, *FEBS Lett.*, 1997, **418**, 363–366.
- 82 A. Fersht, *Structure and mechanism in protein science: a guide to enzyme catalysis and protein folding*, Freeman, New York, 1st edn, 1999, ch. 17, vol. 1, pp. 508–539.
- 83 K. Shiraki, M. Kudou, S. Fujiwara, T. Imanaka and M. Takagi, *J. Biochem.*, 2002, **132**, 591–595.
- 84 J. L. Silva and G. Weber, *Annu. Rev. Phys. Chem.*, 1993, **44**, 89–113.
- 85 J. Roche, M. Dellarole, J. A. Caro, D. R. Norberto, A. E. Garcia, B. Garcia-Moreno, C. Roumestand and C. A. Royer, *J. Am. Chem. Soc.*, 2013, **135**, 14610–14618.
- 86 E. Mixcoha-Hernández, L. M. Moreno-Vargas, A. Rojo-Domínguez and C. G. Benítez-Cardoza, *Protein J.*, 2007, **26**, 491–498.
- 87 H.-J. Woo and B. Roux, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 6825–6830.
- 88 D. Li, B. Ji, K. Hwang and Y. Huang, *PLoS One*, 2011, **6**, e19268.
- 89 L. S. Cheung, D. J. Shea, N. Nicholes, A. Date, M. Ostermeier and K. Konstantopoulos, *Sci. Rep.*, 2015, **5**, 1–9.
- 90 L. A. Abriata and M. Dal Peraro, *Sci. Rep.*, 2015, **5**, 1–12.
- 91 J. M. Sanchez-Ruiz, *Biophys. Chem.*, 2010, **148**, 1–15.
- 92 J. Gomes and W. Steiner, *Food Technol. Biotechnol.*, 2004, **42**, 223–235.
- 93 A. Bhattacharya and B. I. Pletschke, *Enzyme Microb. Technol.*, 2014, **55**, 159–169.
- 94 S. Lindquist, *Annu. Rev. Biochem.*, 1986, **55**, 1151–1191.
- 95 A. Muga and F. Moro, *Curr. Protein Pept. Sci.*, 2008, **9**, 522–566.
- 96 M. J. Gething and J. Sambrook, *Nature*, 1992, **355**, 33–45.
- 97 N. A. Ryabova, V. V. Marchenkov, S. Y. Marchenkova, N. V. Kotova and G. V. Semisotnov, *Biochemistry*, 2013, **78**, 1405–1414.
- 98 P. Bozaykut, N. K. Ozer and B. Karademir, *Free Radical Biol. Med.*, 2014, **77**, 195–209.
- 99 M. J. Kerner, D. J. Naylor, Y. Ishihama, T. Maier, H. C. Chang, A. P. Stines, C. Georgopoulos, D. Frishman, M. Hayer-Hartl, M. Mann and F. Ulrich, *Cell*, 2005, **122**, 209–220.
- 100 F. D. Ciccarelli, T. Doerks, C. von Mering, C. J. Creevey, B. Snel and P. Bork, *Science*, 2006, **311**, 1283–1287.
- 101 K. Guruprasad, B. V. Reddy and M. W. Pandit, *Protein Eng.*, 1990, **4**, 155–161.
- 102 V. N. Uversky, *Biochemistry*, 1993, **32**, 13288–13298.
- 103 N. J. Greenfield, *Nat. Protoc.*, 2006, **1**, 2876–2890.
- 104 W. Kabsch, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2010, **66**, 125–132.
- 105 M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin and K. S. Wilson, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2011, **67**, 235–242.
- 106 P. A. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger and P. H. Zwart, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2010, **66**, 213–221.
- 107 P. Emsley and K. Cowtan, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2004, **60**, 2126–2132.
- 108 I. W. Davis, A. Leaver-Fay, V. B. Chen, J. N. Block, G. J. Kapral, X. Wang, L. W. Murray, W. B. Arendall III, J. Snoeyink, J. Richardson and D. C. Richardson, *Nucleic Acids Res.*, 2007, **35**, W375–W383.
- 109 *The PyMOL Molecular Graphics System, version 1.7.2*, Schrödinger, LLC, New York, U.S., 2014.
- 110 K. Katoh and D. M. Standley, *Mol. Biol. Evol.*, 2013, **30**, 772–780.
- 111 A. M. Waterhouse, J. B. Procter, D. M. A. Martin and G. J. Barton, *Bioinformatics*, 2009, **25**, 1189–1191.
- 112 P. Reche, *SIAS*, Universidad Complutense de Madrid, Madrid, Spain, 2013.
- 113 E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, *J. Comput. Chem.*, 2004, **25**, 1605–1612.
- 114 A. Nicholls, K. A. Sharp and B. Honig, *Proteins*, 1991, **11**, 281–296.
- 115 M. Heinig and D. Frishman, *Nucleic Acids Res.*, 2004, **32**, W500–W502.
- 116 N. R. Voss and M. Gerstein, *Nucleic Acids Res.*, 2010, **38**, W555–W562.
- 117 K. Berka, O. Hanák, D. Sehnal, P. Banás, V. Navrátilová, D. Jaiswal, C. M. Ionescu, R. Svobodová, J. Koca and M. Otyepka, *Nucleic Acids Res.*, 2012, **40**, W222–W227.
- 118 E. Krissinel and K. Henrick, *J. Mol. Biol.*, 2007, **372**, 774–797.
- 119 R. G. Coleman and K. A. Sharp, *J. Mol. Biol.*, 2006, **362**, 441–458.
- 120 S. J. Hubbard and J. M. Thornton, *NACCESS*, University College London, London, U.K., 1993.
- 121 L. Willard, A. Ranjan, H. Zhang, H. Monzavi, R. F. Boyko, B. D. Sykes and D. S. Wishart, *Nucleic Acids Res.*, 2003, **31**, 3316–3319.
- 122 I. McDonald and J. Thornton, *J. Mol. Biol.*, 1994, **238**, 777–793.
- 123 J. M. Sanchez-Ruiz, *Biophys. J.*, 1992, **61**, 921–935.
- 124 A. E. Lyubarev and B. I. Kurganov, *Biochemistry*, 1998, **63**, 434–440.
- 125 D. Rodríguez-Larrea, S. Minning, T. V. Borchert and J. M. Sanchez-Ruiz, *J. Mol. Biol.*, 2006, **360**, 715–724.
- 126 C. N. Pace, B. A. Shirley and J. A. Thomson, in *Protein structure: a practical approach*, ed. T. E. Creighton, IRL Press, Oxford, 2nd edn, 1997, ch. 12, vol. 1, pp. 311–330.
- 127 K. Takahashi and J. M. Sturtevant, *Biochemistry*, 1981, **20**, 6185–6190.
- 128 M. M. Santoro and D. W. Bolen, *Biochemistry*, 1992, **31**, 4901–4907.
- 129 S. Kumar, C. J. Tsai and R. Nussinov, *Biochemistry*, 2002, **41**, 5359–5374.
- 130 A. Cipolla, F. Delbrassine, J. L. Da Lage and G. Feller, *Biochimie*, 2012, **94**, 1943–1950.



Cite this: DOI: 10.1039/c6cp90084d

Correction: Reversibility and two state behaviour in the thermal unfolding of oligomeric TIM barrel proteins

 Sergio Romero-Romero,^a Miguel Costas,^b Adela Rodríguez-Romero^c and D. Alejandro Fernández-Velasco^{*a}

DOI: 10.1039/c6cp90084d

www.rsc.org/pccp

 Correction for 'Reversibility and two state behaviour in the thermal unfolding of oligomeric TIM barrel proteins' by Sergio Romero-Romero *et al.*, *Phys. Chem. Chem. Phys.*, 2015, **17**, 20699–20714.

The authors would like to correct an error on page 20706 of the article, left column, fourth line from bottom. The correct sentence should read as follows:

The size of the biggest cavity, located at the interface between β -strands 4–6 and α -helices 4–6 (comprising residues 90–120, 145–160 and 195–200; Fig. S6, ESI \ddagger), is the main responsible of the cavity volume difference between RevTIMs and IrrevTIMs (Table 4 and Fig. 7).

The Royal Society of Chemistry apologises for these errors and any consequent inconvenience to authors and readers.

^a Laboratorio de Físicoquímica e Ingeniería de Proteínas, Departamento de Bioquímica, Facultad de Medicina, Universidad Nacional Autónoma de México, 04510 Ciudad de México, Distrito Federal, Mexico. E-mail: fdaniel@unam.mx

^b Laboratorio de Biofísicoquímica, Departamento de Físicoquímica, Facultad de Química, Universidad Nacional Autónoma de México, 04510 Ciudad de México, Distrito Federal, Mexico

^c Laboratorio de Química de Biomacromoléculas 3, Departamento de Química de Biomacromoléculas, Instituto de Química, Universidad Nacional Autónoma de México, 04510 Ciudad de México, Distrito Federal, Mexico



Electronic Supplementary Information (ESI)

Reversibility and two state behavior in the thermal unfolding of oligomeric TIM barrel proteins

Sergio Romero-Romero¹, Miguel Costas², Adela Rodríguez-Romero³ and D. Alejandro Fernández-Velasco*¹

*1 - Laboratorio de Fisicoquímica e Ingeniería de Proteínas, Departamento de Bioquímica, Facultad de Medicina, Universidad Nacional Autónoma de México, 04510 Ciudad de México, Distrito Federal, México. **

E-mail: fdaniel@unam.mx

2 - Laboratorio de Biofisicoquímica, Departamento de Fisicoquímica, Facultad de Química, Universidad Nacional Autónoma de México, 04510 Ciudad de México, Distrito Federal, México.

3 - Laboratorio de Química de Biomacromoléculas 3, Departamento de Química de Biomacromoléculas, Instituto de Química, Universidad Nacional Autónoma de México, 04510 Ciudad de México, Distrito Federal, México.

Table of Contents:

Table of Contents	p. S2-S3
Fig. S1: Phylogenetic representations of BacTIMs.....	p. S4
Fig. S2: Far-UV CD spectra of BacTIMs at 25 °C	p. S5
Fig. S3: Intrinsic tryptophan fluorescence spectra of BacTIMs	p. S6
Fig. S4: Dependence of the T_m with respect to scan rate for BacTIMs	p. S7
Fig. S5: Far-UV CD spectra of BacTIMs at 80 °C	p. S8
Fig. S6: Amino acid sequence alignment of diverse TIMs used for comparison.....	p. S9-S10
Table S1: Data collection and refinement statistics of <i>Go</i> TIM, <i>Dr</i> TIM, <i>Sc</i> TIM and <i>Cp</i> TIM	p. S11
Table S2: Dimeric TIM PDB files used in the structural comparison of IrrevTIMs and RevTIMs.	p. S12
Table S3: Thermodynamic parameters of irreversible thermal unfolding of <i>Np</i> TIM and <i>Go</i> TIM.....	p. S13
Table S4: Matrix with identity sequence percentage and average RMSD values for IrrevTIMs and RevTIMs	p. S14

Table S5: Amino acidic composition used in the sequence comparison of IrrevTIMs and RevTIMs.....	p. S15
Table S6: Secondary structure elements and stabilizing interactions for IrrevTIMs and RevTIMs.....	p. S16
References for Electronic Supplementary Information.....	p. S17-S18

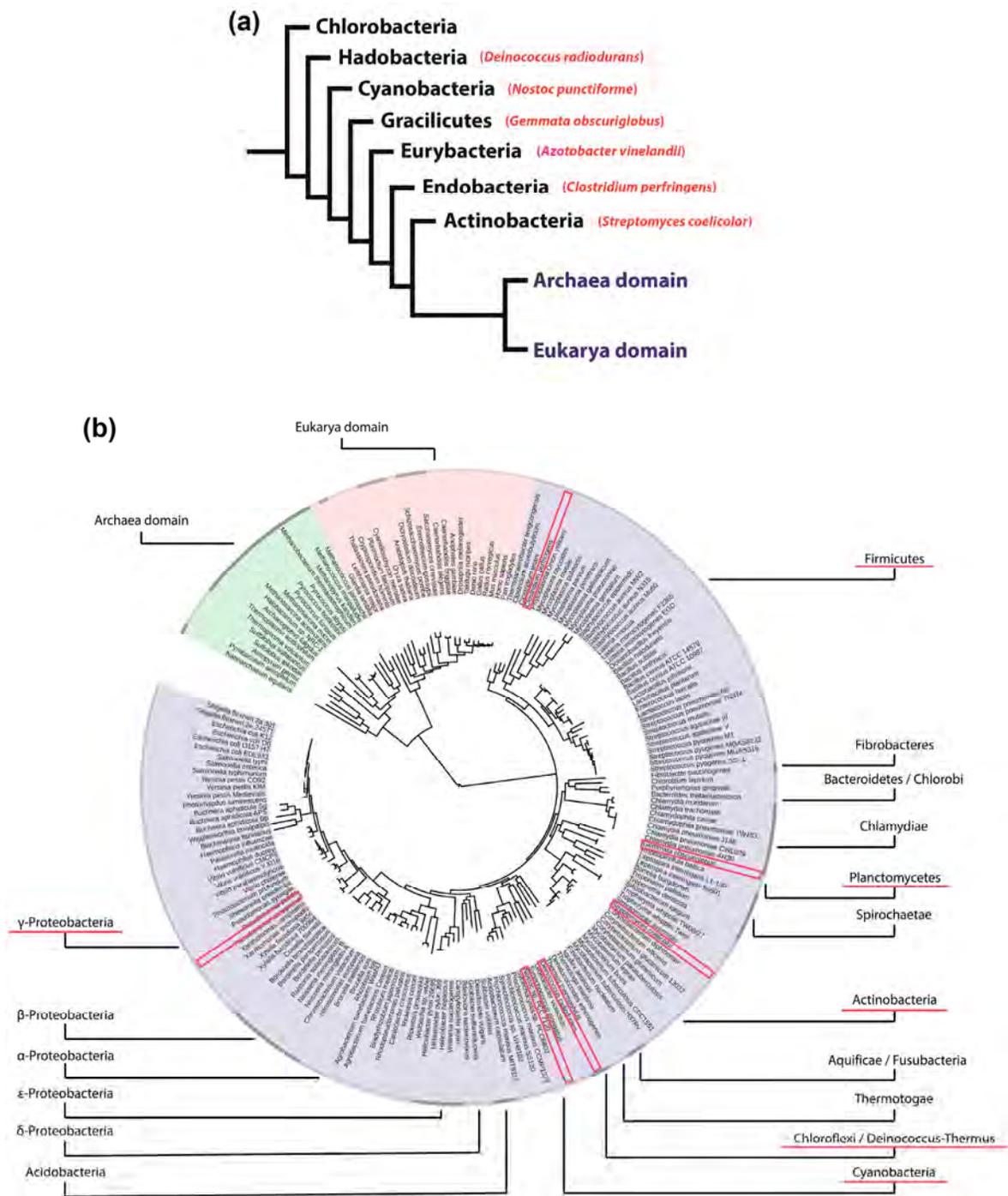


Fig. S1. Phylogenetic representations of BacTIMs. **(a)** Cladogram of bacterial supertaxa (according to ref. 1) showing the species from which the indicated TIM sequence was studied. **(b)** Phylogenetic tree showing the TIMs studied in this work (modified from ref. 2). For AvTIM phylum (γ -Proteobacteria) only an approximate location in the phylogenetic tree is shown.

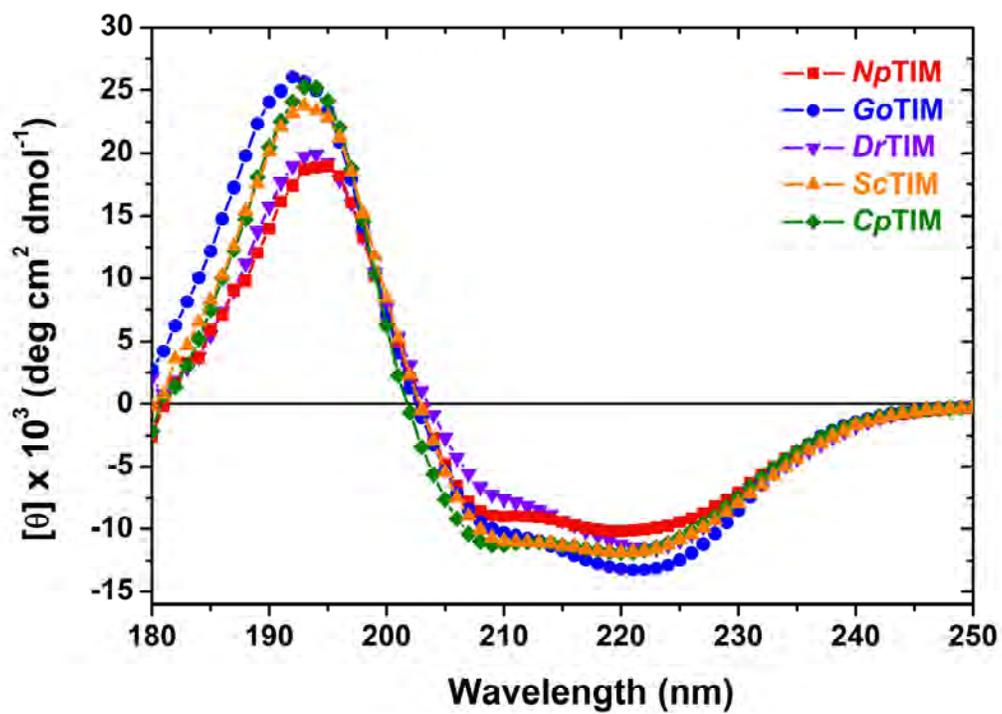


Fig. S2. Far-UV CD spectra of BacTIMs at 25 °C. Data shown are the average of five scans recorded using a protein concentration of 15 μM in buffer C (10 mM NaH_2PO_4 pH 8.0).

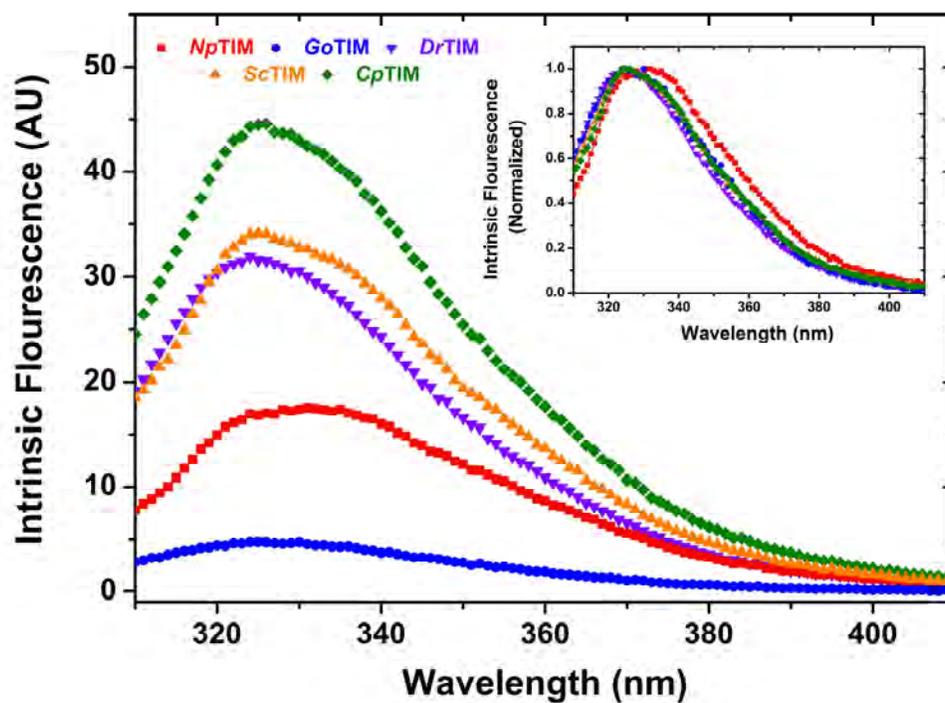


Fig. S3. Intrinsic tryptophan fluorescence spectra of BacTIMs. The spectra shown are the average of five scans recorded using a protein concentration of 15 μ M at 25 $^{\circ}$ C in buffer C (10 mM NaH_2PO_4 pH 8.0). The excitation wavelength was 295 nm. The inset shows normalized data.

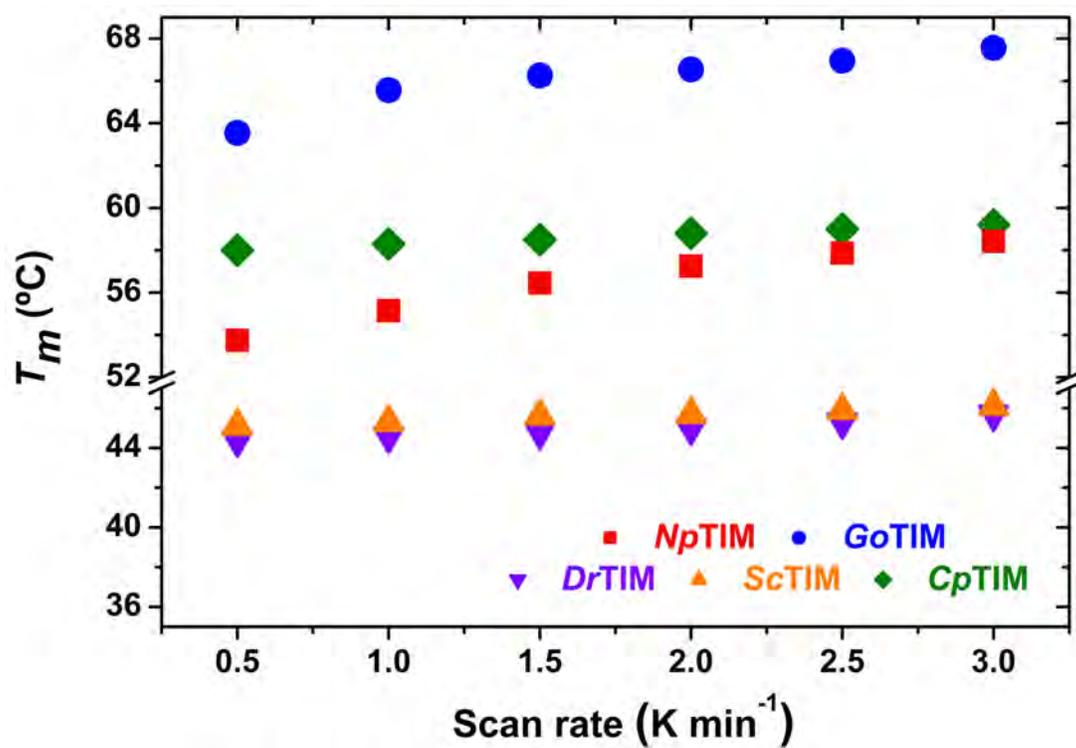


Fig. S4. Dependence of the T_m with respect to scan rate for BacTIMs. In all experiments, protein concentration was 15 μM .

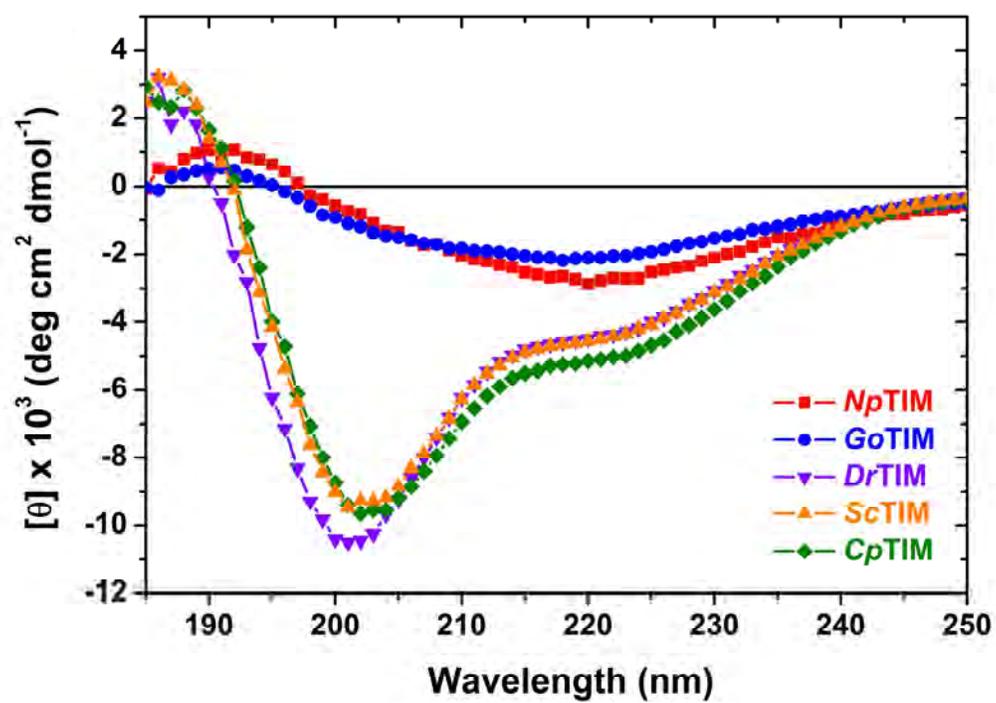


Fig. S5. Far-UV CD spectra of BacTIMs at 80 °C. Data shown are the average of five scans recorded using a protein concentration of 15 μM in buffer C (10 mM NaH_2PO_4 pH 8.0).

```

.....EcTIM..... 1 . . . MRHPLVMGNWKN . . GSRHMVHELVSNLRK . . . ELAGV . . AGCAVA IAPPEMY IDMAKREA . . . EGS HIM 60
.....EhTIM..... 1 . MSARKFVVGGNWKN . . GTLAS IETLTGKVAASVDAELA . . . KKVEV I VGVPI Y I PKVQQ I LAGEANGAN I L 68
.....GgTIM..... 1 MAP . RKFFVVGGNWKM . . GDKKSLGEL IHTLNG . . . AKLS . . . ADTEVVCGAPS IYLDFAQQKL . . . . . DAK I G 60
.....GITIM..... 1 . MPARRPF IGGNFKN . . GS LDF I KSHVAA IAA . . . HKIP . . . DSVDVY IAPS AVHLSTA I AAN . . . . . TSKQLR 61
.....GtTIM..... 1 . . . MRKPI IAGNWKMH . . KTLAEAVQFVEDVKG . . . HVPPA . . . DEVISVVCAPFLFLDRLVQAA . . . . . DGTDLK 60
.....HsTIM..... 1 MAPSRKFFVVGGNWKM . . GRKQSLGEL IGTLNA . . . AKVP . . . ADTEVVCAPPTAY IDFAQQKL . . . . . DPK I A 61
.....LmTIM..... 1 MSAKPQP IAAANWKN . . GTTAS IEKLVQVFNE . . . HTIS . . . HDVQC VVAPT FVH I PLVQAKL . . . . . RNPKYV 62
.....MmTIM..... 1 . . . MRHPVVMGNWKN . . GSKEMVVD LNLGNA . . . ELEGV . . . TGVDVAVAPPALFVDLAERTLT . . . EAGSA I I 62
.....MtTIM..... 1 . . . MSRKPL IAGNWKMN . . LNHYEA IALVQK IAF . . . SLPDKYYDRVDVAV IPPFTDLRSVQTLV . . . . . DGDKLR 63
.....OcTIM..... 1 MAPSRKFFVVGGNWKM . . GRKKNLGE L I TTLNA . . . AKVP . . . ADTEVVCAPPTAY IDFAQQKL . . . . . DPK I A 61
.....PFTIM..... 1 . . . MARKYFVAANWKN . . GTLES IKS L TNSFN . . . LDFDP . . . SKLDVVVFPVSVHYDHRKLL . . . . . QSKFS 60
.....SsTIM..... 1 MAPARKFFVVGGNWKM . . GRKNNLGE L I TTLNA . . . AKLP . . . ADTEVVCAPPTAY IDFAQQKL . . . . . DPK I A 61
.....TbTIM..... 1 . MSKPQP IAAANWKN . . GSQQSLSEL IDLFNS . . . TSIN . . . HDVQC VVAST FVH LAMTKERL . . . . . SHPKFV 61
.....TcTIM..... 1 MASKPQP IAAANWKN . . GSESLVPL IETLNA . . . ATFD . . . HDVQC VVAPT FLH I P MTKARL . . . . . TNPKFQ 62
.....TsTIM..... 1 . . . MTRKFLVGGNWKMN . . GSYSH INTFFDTLQK . . . ADTD . . . PNAD I V IGV PACYLKYAQDKA . . . . . PKG I K 59
.....TvTIM..... 1 . . . MRTFFVGGNWKANPKTVEEAEL I EMLNG . . . AKVE . . . GNVEVVVAAPF I FLPTLQQKL . . . . . RKDWK 59
.....YtTIM..... 1 . . . MARTFFVGGNFKN . . GSKQS I KE I VERLNT . . . ASIP . . . ENVEVY ICPPATYLDYSVSLV . . . . . KKPQVT 60
.....NpTIM..... 1 . . . MRKIV IAGNWKMF . . KTAETQEF LQGF LP . . . HLEETP . . . QGREV ILCPPFTDLSVLSKTL . . . . . HGSL I Q 61
.....GoTIM..... 1 . . . MPTRKKFVAGNWKMN . . TTLAEAKALGA AVAK . . . GVTD . . . DRVTVAVFPYPWLTA VGEVL . . . . . KGS PVA 61
.....DrTIM..... 1 . . . MQTLLALNWKMN . . KPT EARSWAEELT . . . KYAPA . . . EGVDLAVLAPALDLSALANL . . . . . PAG I A 58
.....ScTIM..... 1 . . . MTRTPTLMAGNWKMN . . LNHLEA IAHVQKLAF . . . ALADKDYDAVEAVLAPFTDLRSVQTLV . . . . . DGDKLR 64
.....CpTIM..... 1 . . . MRTPI IAGNWKMH . . YTIDEAVK LVEELKP . . . LVKD . . . AKCEVVVCPTFVCLDAVKKAV . . . . . EGN I K 59
Consensus R . . . . . GNWK - N G . . . . . L I . . . . . VV . . . . . PP . . . . . L . . . . .
M+PMRKPFVAGNWKMN -GTKAEA+ELVETLNA - - AKI++ -YADVEVVVAPPFVYLDFAQQKL - - - -DGPKI+
.....EcTIM..... 61 L . . GAQNVLDNLSGAFTGET SAAMLKDIGAQY I I IGHSE RRYHKESEDEL IAKKFAVLK EQGLTFVLC IGETE 131
.....EhTIM..... 69 V . . SAENAW .TKSGAYTGEVHVGM L VDCQVYV I I IGHSE RRYHKESEDEL IAKKFAVLK EQGLTFVLC IGETE 138
.....GgTIM..... 61 V . . AAQNCYKVPKGAFTGE I SPAM I KDIGA AWV I I IGHSE RRYHKESEDEL IAKKFAVLK EQGLTFVLC IGETE 131
.....GITIM..... 62 I . . AAQNVYLEGNGAWTGET SVEM LQDMGLKHV I V IGHSE RRYHKESEDEL IAKKFAVLK EQGLTFVLC IGETE 132
.....GtTIM..... 61 I . . GAQTMHFADQGAFTGEVSPVMLKDLGVTYV I I IGHSE RRYHKESEDEL IAKKFAVLK EQGLTFVLC IGETE 131
.....HsTIM..... 62 V . . AAQNCYKVTNGAFTGE I SPGMIKDCGATWVVLGHSE RRYHKESEDEL IAKKFAVLK EQGLTFVLC IGETE 132
.....LmTIM..... 63 I . . SAENAI . AKSGAFTGEVSPM I LKDIGVHWV I I IGHSE RRYHKESEDEL IAKKFAVLK EQGLTFVLC IGETE 132
.....MmTIM..... 63 L . . GAQNTDLNNSGAFTGDMSPAM LKEFGATH I I IGHSE RRYHKESEDEL IAKKFAVLK EQGLTFVLC IGETE 133
.....MtTIM..... 64 LTYGAQDLSPHDSGAYTGDVSGAF LAKLGC SYVVVGHSE RRYHKESEDEL IAKKFAVLK EQGLTFVLC IGETE 136
.....OcTIM..... 62 V . . AAQNCYKVTNGAFTGE I SPGMIKDCGATWVVLGHSE RRYHKESEDEL IAKKFAVLK EQGLTFVLC IGETE 132
.....PFTIM..... 61 T . . GIQNVSKFGNGSYTGEVSAE IAKDLNIEYV I I IGHSE RRYHKESEDEL IAKKFAVLK EQGLTFVLC IGETE 131
.....SsTIM..... 62 V . . AAQNCYKVANGAFTGE I GPGMIKDLGATWVVLGHSE RRYHKESEDEL IAKKFAVLK EQGLTFVLC IGETE 132
.....TbTIM..... 62 I . . AAQNAI . AKSGAFTGEVSLP I LKDFGVNWI V IGHSE RRYHKESEDEL IAKKFAVLK EQGLTFVLC IGETE 131
.....TcTIM..... 63 I . . AAQNAI . TRSGAFTGEVSLQ I LKDYGIKWWV I I IGHSE RRYHKESEDEL IAKKFAVLK EQGLTFVLC IGETE 132
.....TsTIM..... 60 I . . AAENCYKVGSGAFTGE ISTEM I KDCGCEWV I I IGHSE RRYHKESEDEL IAKKFAVLK EQGLTFVLC IGETE 130
.....TvTIM..... 60 V . . SAENVFTKPNGAFTGEVTPM I KSFGEIWT I I IGHSE RRYHKESEDEL IAKKFAVLK EQGLTFVLC IGETE 130
.....YtTIM..... 61 V . . GAQNAYLKASGAFTGENSV DQIKDVGAKWV I I IGHSE RRYHKESEDEL IAKKFAVLK EQGLTFVLC IGETE 131
.....NpTIM..... 62 L . . GAQNIHWEEFGAFTGE I SGPM LTESGVRFV I V IGHSE RRYHKESEDEL IAKKFAVLK EQGLTFVLC IGETE 132
.....GoTIM..... 62 L . . GAQDVSEKKG AFTGEVSPAM LLETCKYAL IGHSE RRYHKESEDEL IAKKFAVLK EQGLTFVLC IGETE 132
.....DrTIM..... 59 F . . GGQDVSAHESGAYTGE I SAAMLK DAGASC VVGHSE RRYHKESEDEL IAKKFAVLK EQGLTFVLC IGETE 129
.....ScTIM..... 65 I KYGAQDISAHDGGAFTGE I SGPM LAKLKCTYVAVGHSE RRYHKESEDEL IAKKFAVLK EQGLTFVLC IGETE 137
.....CpTIM..... 60 V . . GAQNMHFEKGAFTGE I APRMLEAMNIDYV I I IGHSE RRYHKESEDEL IAKKFAVLK EQGLTFVLC IGETE 130
Consensus AQN . . . . . GA=TGE S MLKD G V+ . . . GHSERR . . . . . E DE . . . . . K . . . . . A . . . . . GL . . . . . C GE L
V-YGAQNVYKVKSGAFTGE+SPAM LKDLGATWV I IGHSE RRYHKESEDELVA+KVA+ALAEGLTV I I C IGETL

```

Continue in the next page...

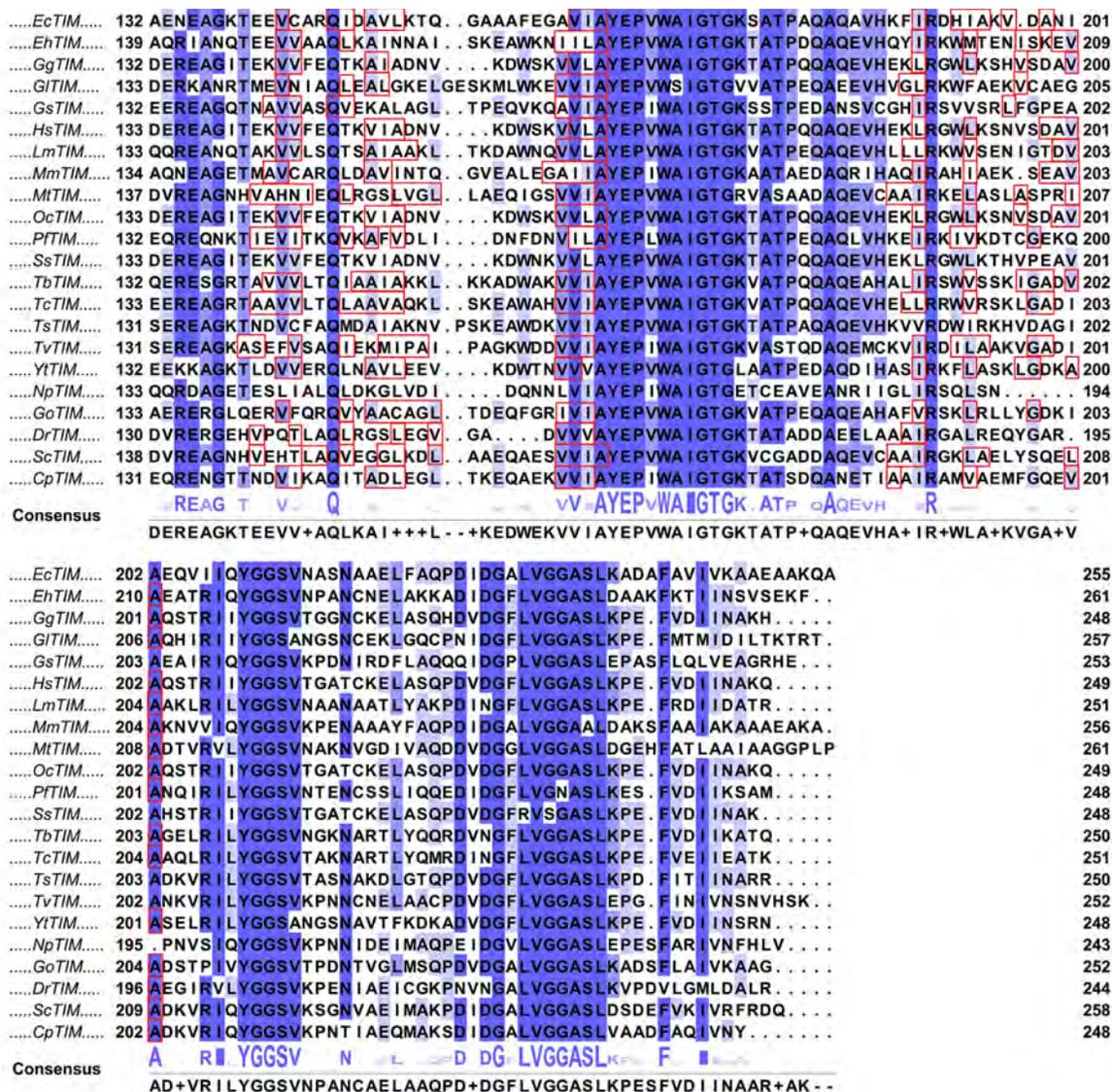


Fig. S6. Sequence alignment of diverse TIMs used for comparison. The conserved residues in these TIMs are shown on the *consensus line* and represented from blue light to dark in accordance to increase conservation. In red squares are shown the residues (both internal and boundary) involved in the formation of the biggest cavity in these structures, except for *NpTIM*, *SsTIM* and *TsTIM* for which there is not structural data reported.

Table S1. Thermodynamic parameters for the irreversible thermal unfolding of *Np*TIM and *Go*TIM.

Thermodynamic parameter	Bacterial TIMs			Eukaryotic TIMs			
	<i>Np</i> TIM	<i>Go</i> TIM	<i>Gl</i> TIM ^a	<i>Hs</i> TIM ^a	<i>Lm</i> TIM ^b	<i>Tb</i> TIM ^b	<i>Tc</i> TIM ^b
Activation free energy (kJ mol ⁻¹) from average of individual fits to DSC thermograms (Eqs. (3) and (4))	368 ± 46	716 ± 63	NR	NR	312 ± 5	397 ± 5	809 ± 7
Activation free energy (kJ mol ⁻¹) from Arrhenius plot (Eq. (2))	368 ± 17	716 ± 25	367 ± 7	361 ± 4	315 ± 3	398 ± 5	793 ± 18
Activation free energy (kJ mol ⁻¹) from consistency test using several scanning rates (Eq. (5))	343 ± 21	715 ± 88	NR	NR	347 ± 39	420 ± 55	715 ± 137
Activation free energy average (kJ mol ⁻¹)	360 ± 25	715 ± 46	NR	NR	325 ± 21	406 ± 29	774 ± 75
m^\ddagger (kJ mol ⁻¹ M ⁻¹)	1.67	8.12	3.94	2.77	1.47	1.69	8.32
m_{eq} (kJ mol ⁻¹ M ⁻¹)	27.95 ^c	27.77	30.02	30.09	30.80	30.35	30.67
m^\ddagger / m_{eq}	0.06	0.29	0.13	0.09	0.05	0.06	0.27
Number of residues unfolded in the transition state ^d	30	148	67	45	24	28	136

^a From Ref. 18. ^b From Ref. 19. ^c Because the structure of *Np*TIM is not known, the Δ ASA value used for the calculation of m_{eq} is the average of Δ ASA from all three-dimensional structures employed for structural comparison (57388 ± 1365 Å²; Table 4). ^d Calculated from: $N_r(m^\ddagger/m_{eq})$, where N_r is the total number of residues in the dimeric protein, and m^\ddagger/m_{eq} is the degree of unfolding as estimated from the urea m values. NR: value not reported.

Table S2. Data collection and refinement statistics of *Go*TIM, *Dr*TIM, *Sc*TIM and *Cp*TIM.

Data collection ^a				
	<i>Go</i> TIM	<i>Dr</i> TIM	<i>Sc</i> TIM	<i>Cp</i> TIM
PDB ID	4Y96	4Y90	4Y9A	4Y8F
Resolution range (Å)	36.36-1.58 (1.64-1.58)	36.13-2.09 (2.17-2.09)	43.03-2.29 (2.37-2.29)	37.59-1.54 (1.59-1.54)
Space group	P 6 ₅ 2 2	R 3 2	P 4 ₃	C 1 2 1
Unit cell dimensions				
a, b, c, (Å)	124.8, 127.8, 134.2	169.6, 169.6, 202.3	86.1, 86.1, 134.0	75.3, 49.6, 71.5
α, β, γ, (°)	90.0, 90.0, 120.0	90.0, 90.0, 120.0	90.0, 90.0, 90.0	90.0, 120.0, 90.0
Total reflections	779530 (107888)	261090 (36321)	128420 (19245)	142402 (17348)
Unique reflections	81335 (8192)	64962 (6342)	43475 (4241)	33930 (3130)
Multiplicity	9.6 (9.0)	4.0 (3.9)	3.0 (3.1)	4.2 (3.7)
Completeness (%)	99.3 (97.1)	99.7 (98.6)	99.7 (98.1)	99.2 (94.4)
Mean I / sigma (I)	29.0 (4.8)	11.0 (2.9)	10.0 (3.0)	14.6 (3.1)
R-merge	0.040 (0.487)	0.097 (0.464)	0.071 (0.371)	0.062 (0.448)
Refinement statistics				
	<i>Go</i> TIM	<i>Dr</i> TIM	<i>Sc</i> TIM	<i>Cp</i> TIM
R _{work} / R _{free} (%)	17.8 / 20.1	13.8 / 18.6	23.1 / 28.2	14.8 / 18.1
Average B-value (Å ²)	24.9	25.9	35.8	16.6
Protein	23.1	24.8	35.9	14.1
Ligand	32.8	49.3	---	31.5
Solvent	35.2	33.2	32.7	29.5
Number of atoms	4446	8069	7943	2362
Protein	3777	7142	7690	1977
Ligand	14	87	---	11
Water	639	832	250	372
Protein residues	500	976	1018	251
RMS (bonds) (Å)	0.007	0.011	0.009	0.010
RMS (angles) (°)	1.11	1.30	1.33	1.24
Ramachandran favored (%)	98.43	97.74	96.15	98.47
Ramachandran allowed (%)	1.18	1.85	3.56	1.15
Ramachandran outliers (%)	0.39	0.41	0.30	0.38
Clashcore	1.83	1.25	5.41	1.76

^a Statistics for the highest-resolution shell are shown in parentheses.

Table S3. Matrix with identity sequence percentage (below) and average RMSD values (top) for IrrevTIMs and RevTIMs.^a

		RMSD values																							
		IrrevTIMs															RevTIMs								
		<i>Ec</i> TIM	<i>Eh</i> TIM	<i>Gg</i> TIM	<i>Gl</i> TIM	<i>Gs</i> TIM	<i>Hs</i> TIM	<i>Lm</i> TIM	<i>Mm</i> TIM	<i>Mt</i> TIM	<i>Oc</i> TIM	<i>Pf</i> TIM	<i>Ss</i> TIM	<i>Tb</i> TIM	<i>Tc</i> TIM	<i>Ts</i> TIM	<i>Tv</i> TIM	<i>Yt</i> TIM	<i>Np</i> TIM	<i>Go</i> TIM	<i>Dr</i> TIM	<i>Sc</i> TIM	<i>Cp</i> TIM		
IrrevTIMs	<i>Ec</i> TIM	-----	1.51	1.30	1.53	1.48	1.39	1.13	1.24	1.10	1.50	1.27	1.41	1.48		1.24	1.08		1.49	1.37	1.36	1.46			
	<i>Eh</i> TIM	41	-----	0.83	1.50	1.16	0.97	1.09	1.65	1.15	0.93	1.04	0.82	1.07		0.90	1.21		1.51	1.20	0.87	1.07			
	<i>Gg</i> TIM	45	50	-----	1.01	0.92	0.40	0.65	1.20	1.05	0.34	0.85	0.61	0.73		0.62	0.68		1.36	1.00	0.87	1.22			
	<i>Gl</i> TIM	40	43	44	-----	1.21	1.01	1.05	1.09	1.37	0.98	1.30	1.10	0.88		1.04	1.12		0.86	1.49	1.45	1.52			
	<i>Gs</i> TIM	39	41	39	38	-----	0.99	1.49	1.51	1.20	0.94	1.30	1.13	1.30		1.11	1.24		1.33	1.07	0.88	0.78			
	<i>Hs</i> TIM	44	49	90	45	38	-----	0.93	1.26	1.06	0.22	0.96	ND	0.69	0.79	ND	0.64	0.87	ND	1.13	1.29	0.91	1.17		
	<i>Lm</i> TIM	42	51	50	46	43	50	-----	1.16	1.13	0.91	0.84	0.46	0.51		1.03	0.92		1.46	1.05	1.23	1.39			
	<i>Mm</i> TIM	65	41	40	35	41	40	38	-----	1.42	1.33	1.36	1.40	1.17		1.17	0.99		1.31	1.39	1.44	1.26			
	<i>Mt</i> TIM	41	36	37	38	43	37	38	41	-----	1.11	0.93	1.11	1.25		1.15	1.10		1.49	0.61	0.70	1.40			
	<i>Oc</i> TIM	44	49	89	45	38	98	50	40	37	-----	1.10	0.76	0.87		0.63	0.86		1.14	1.13	0.92	1.11			
RevTIMs	<i>Pf</i> TIM	39	44	43	40	37	42	45	38	35	42	-----	0.87	0.93		1.07	1.01		1.52	1.10	1.01	1.42			
	<i>Ss</i> TIM	43	48	87	44	37	93	49	39	35	93	42	-----					ND							
	<i>Tb</i> TIM	42	43	51	44	40	52	69	38	38	51	43	39	-----	0.47				0.89	0.98		1.48	1.07	0.96	1.35
	<i>Tc</i> TIM	42	45	50	46	41	52	68	39	39	51	43	50	74	-----		ND		0.91	0.88	ND	1.38	1.49	1.17	1.48
	<i>Ts</i> TIM	49	48	61	47	39	59	46	39	38	59	41	58	47	50	-----					ND				
	<i>Tv</i> TIM	40	44	48	41	43	47	43	40	39	47	38	46	42	42	48	-----	1.02			1.15	1.01	0.89	0.95	
	<i>Yt</i> TIM	44	43	53	48	38	52	46	42	39	52	42	51	48	46	49	46	-----	ND		1.49	1.27	1.15	1.47	
	<i>Np</i> TIM	39	38	37	33	47	36	36	43	40	36	33	35	34	32	34	38	35	-----			ND			
	<i>Go</i> TIM	46	41	46	39	45	45	39	43	44	45	40	43	41	40	45	39	40	38	-----	1.25	1.43	1.44		
	<i>Dr</i> TIM	38	34	38	36	41	38	37	39	48	37	33	37	37	39	38	39	37	40	40	-----	0.76	1.35		
<i>Sc</i> TIM	42	40	38	38	48	37	40	40	65	37	34	36	37	38	40	43	37	46	45	49	-----	0.94			
<i>Cp</i> TIM	45	42	41	38	55	40	42	44	41	40	40	40	40	40	41	43	40	48	44	42	52	-----			

^a ND: average RMSD value not determined because there is no structural data for *Np*TIM, *Ss*TIM and *Ts*TIM.

Table S4. Amino acidic composition used in the sequence comparison of IrrevTIMs and RevTIMs.

Structural property	<i>Dr</i> TIM	<i>Sc</i> TIM	<i>Cp</i> TIM	Average RevTIMs	Average IrrevTIMs
Amino acid composition (%)					
Alanine (A)	17.0	13.6	12.4	14.3 ± 2.0	12.0 ± 3.1
Cysteine (C)	1.2	1.6	2.4	1.7 ± 0.5	1.5 ± 0.4
Aspartic acid (D)	4.9	7.4	4.8	5.7 ± 1.2	4.4 ± 1.4
Glutamic acid (E)	8.1	7.4	9.6	8.3 ± 0.9	7.2 ± 1.3
Phenylalanine (F)	0.4	1.6	2.8	1.6 ± 1.0	3.2 ± 0.9
Glycine (G)	10.9	8.9	7.2	9.0 ± 1.5	8.8 ± 1.4
Histidine (H)	2.0	3.1	2.0	2.4 ± 0.5	2.3 ± 0.6
Isoleucine (I)	3.2	5.0	8.0	5.4 ± 1.9	6.9 ± 1.2
Lysine (K)	3.6	7.0	7.6	6.1 ± 1.7	6.9 ± 1.8
Leucine (L)	10.1	8.5	4.8	7.8 ± 2.2	7.1 ± 1.3
Methionine (M)	1.6	1.9	3.2	2.2 ± 0.7	1.6 ± 0.7
Asparagine (N)	3.2	2.3	4.8	3.4 ± 1.0	4.0 ± 1.2
Proline (P)	4.0	2.3	2.8	3.1 ± 0.7	3.5 ± 0.9
Glutamine (Q)	2.8	3.9	3.6	3.4 ± 0.4	4.0 ± 1.1
Arginine (R)	4.9	3.5	2.8	3.7 ± 0.9	3.6 ± 0.8
Serine (S)	4.9	3.9	2.4	3.7 ± 1.0	5.1 ± 1.1
Threonine (T)	4.5	4.3	7.2	5.3 ± 1.3	5.1 ± 0.9
Valine (V)	8.9	9.7	8.8	9.1 ± 0.4	9.2 ± 1.1
Tryptophan (W)	1.2	0.8	0.8	0.9 ± 0.2	1.5 ± 0.5
Tyrosine (Y)	2.4	3.5	2.4	2.8 ± 0.5	2.0 ± 0.6
Physicochemical amino acid properties (%)					
Charged (DEHKR)	23.5	28.3	26.7	26.2 ± 2.0	24.4 ± 1.9
Positively charged (HKR)	10.5	13.6	12.4	12.1 ± 1.2	12.7 ± 1.3
Negatively charged (DE)	13.0	14.7	14.3	14.0 ± 0.8	11.6 ± 1.4
Aliphatic (AGILPV)	54.3	48.1	43.8	48.7 ± 4.3	47.5 ± 3.3
Aromatic (FHXY)	6.1	8.9	8.0	7.7 ± 1.2	9.0 ± 0.8
Polar (DEKNQR)	27.5	31.4	33.1	30.7 ± 2.3	30.1 ± 2.4
Neutral, polar (CNQSTY)	19.0	19.4	22.7	20.4 ± 1.7	21.8 ± 2.7
Neutral, non-polar (AFGILMPVW)	57.5	52.3	50.6	53.5 ± 2.9	53.8 ± 2.9
Hydrophobic (CFILMVW)	26.7	29.1	30.7	28.8 ± 1.6	31.0 ± 1.4
Small size (ACGSTV)	47.4	41.9	40.2	43.2 ± 3.1	41.8 ± 3.0
Medium size (DEHILMN PQ)	40.1	41.9	43.4	41.8 ± 1.4	41.0 ± 2.9
Large size (FKRWY)	12.6	16.3	16.3	15.1 ± 1.8	17.2 ± 2.1

Table S5. Secondary structure elements and stabilizing interactions for IrrevTIMs and RevTIMs.

Structural property	<i>Dr</i> TIM	<i>Sc</i> TIM	<i>Cp</i> TIM	Average RevTIMs	Average IrrevTIMs
Secondary structure elements (%)					
α -helix	44	48	43	45.0 \pm 2.2	44.9 \pm 2.1
β -strand	21	20	21	20.7 \pm 0.5	21.8 \pm 1.0
Random coil	35	32	36	34.3 \pm 1.7	33.3 \pm 1.8
Stabilizing interactions					
Total H-bonds of the oligomer	490	503	510	501 \pm 8	521 \pm 24
Interface H-bonds	27	19	24	23 \pm 3	28 \pm 5
Total salt bridges of the oligomer	37	50	44	44 \pm 5	34 \pm 7
Interface salt bridges	2	2	12	5 \pm 5	6 \pm 3

Table S6. Dimeric TIM PDB files used in the structural comparison of IrrevTIMs and RevTIMs.

Organism	Code	PDB ID	Resolution (Å)	Spatial group	Asymmetric unit ^a	Conformation of active site loops		Reference
						Subunit A	Subunit B	
<i>Clostridium perfringens</i>	<i>Cp</i> TIM	4Y8F	1.54	C 1 2 1	Monomer	Open	Open	This work
<i>Deinococcus radiodurans</i>	<i>Dr</i> TIM	4Y90	2.10	R 3 2	Tetramer	Open	Open	This work
<i>Entamoeba histolytica</i>	<i>Et</i> TIM	1M6J	1.50	P 2 ₁ 2 ₁ 2	Dimer	Open	Open	3
<i>Escherichia coli</i>	<i>Ec</i> TIM	4K6A	1.80	P 2 ₁ 2 ₁ 2 ₁	Dimer	Open	Open	4
<i>Gallus gallus</i>	<i>Gg</i> TIM	1TPH	1.80	P 2 ₁ 2 ₁ 2 ₁	Dimer	Closed	Closed	5
<i>Gemmata obscuriglobus</i>	<i>Go</i> TIM	4Y96	1.58	P 6 ₅ 2 2	Dimer	Closed	Open	This work
<i>Geobacillus stearothermophilus</i>	<i>Gs</i> TIM	1BTM	2.80	P 2 ₁ 2 ₁ 2	Dimer	Closed	Closed	6
<i>Giardia lamblia</i>	<i>Gl</i> TIM	2DP3	2.10	I 2 2 2	Monomer	Closed	Open	7
<i>Homo sapiens</i>	<i>Hs</i> TIM	2JK2	1.70	P 2 ₁ 2 ₁ 2 ₁	Dimer	Open	Open	8
<i>Leishmania mexicana</i>	<i>Lm</i> TIM	1AMK	1.83	C 1 2 1	Monomer	Closed	Closed	9
<i>Moritella marina</i>	<i>Mm</i> TIM	1AW2	2.65	P 1 2 ₁ 1	Dimer	Open	Open	10
<i>Mycobacterium tuberculosis</i>	<i>Mt</i> TIM	3TA6	1.41	C 1 2 1	Dimer	Open	Open	11
<i>Oryctolagus cuniculus</i>	<i>Oc</i> TIM	1R2R	1.50	P 2 ₁ 2 ₁ 2 ₁	Dimer	Open	Open	12
<i>Plasmodium falciparum</i>	<i>Pf</i> TIM	1YDV	2.20	C 1 2 1	Dimer	Open	Open	13
<i>Saccharomyces cerevisiae</i>	<i>Yt</i> TIM	1NF0	1.60	P 2 ₁ 2 ₁ 2 ₁	Dimer	Open	Closed	14
<i>Streptomyces coelicolor</i>	<i>Sc</i> TIM	4Y9A	2.30	P 4 ₃	Dimer	Open	Open	This work
<i>Trichomonas vaginalis</i> ^b	<i>Tv</i> TIM	3QSR	2.05	P 2 2 ₁ 2 ₁	Monomer	Open	Open	15
<i>Trypanosoma brucei</i>	<i>Tb</i> TIM	5TIM	1.83	P 2 ₁ 2 ₁ 2 ₁	Dimer	Open	Closed	16
<i>Trypanosoma cruzi</i>	<i>Tc</i> TIM	1TCD	1.83	P 2 ₁ 2 ₁ 2 ₁	Dimer	Closed	Open	17

^a The number of molecules in asymmetric unit was determined by the Matthews coefficient analysis by solvent content percentage in the crystal.

^b Structural comparison of *Tv*TIM was realized with the three-dimensional structure of its dimeric Ile45 variant.

References for Electronic Supplementary Information

1. T. Cavalier-Smith, *Biol. Direct.*, 2006, **1**, 1-19.
2. I. Letunic and P. Bork, *Nucleic Acids Res.*, 2011, **39**, W475-W478.
3. A. Rodríguez-Romero, A. Hernandez-Santoyo, L. del Pozo-Yauner, A. Kornhauser and D. A. Fernández-Velasco, *J. Mol. Biol.*, 2002, **322**: 669-75.
4. G. Minasov, M. Kuhn, A. Halavaty, L. Shuvalova, I. Dubrovskaya, J. Winsor, S. Grimshaw, W. F. Anderson and Center for Structural Genomics of Infectious Diseases, 2013, Forthcoming.
5. Z. Zhang, S. Sugio, E. A. Komives, K. D. Liu, J. R. Knowles, G. A. Petsko and D. Ringe, *Biochemistry*, 1994, **33**, 2830-2837.
6. L. F. Delboni, S. C. Mande, F. Reintier-Delrue, V. Mainfroid, S. Turtley, F. M. D. Vellieux, J. A. Martial and W. G. J. Hol, *Protein Sci.*, 1995, **4**, 2594-2604.
7. H. Reyes-Vivas, A. Díaz, J. Peón, G. Mendoza-Hernández, G. Hernández-Alcántara, I. de la Mora-de la Mora, S. Enríquez-Flores, L. Domínguez-Ramírez and G. López-Velázquez, *J. Mol. Biol.*, 2007, **365**, 752-763.
8. C. Rodríguez-Almazán, R. Arreola, D. Rodríguez-Larrea, B. Aguirre-López, M. T. de Gómez-Puyou, R. Pérez-Montfort, M. Costas, A. Gómez-Puyou and A. Torres-Larios, *J. Biol. Chem.*, 2008, **283**, 23254-23263.
9. J. C. Williams, J. P. Zeelen, G. Neubauer, G. Vriend, J. Backmann, P. A. Michels, A. M. Lambeir and R. K. Wierenga, *Protein Eng.*, 1999, **12**, 243-250.
10. M. Alvarez, J. P. Zeelen, V. Mainfroid, F. Reintier-Delrue, J. A. Martial, L. Wyns, R. K. Wierenga and D. Maes, *J. Biol. Chem.*, 1998, **273**, 2199-2206.
11. S. E. Connor, G. C. Capodagli, M. K. Deaton and S. D. Pegan, *Acta Crystallogr. D Biol. Crystallogr.*, 2011, **67**, 1017-1022.
12. R. Aparicio, S. T. Ferreira and I. Polikarpov, *J. Mol. Biol.*, 2003, **334**, 1023-1041.
13. S. S. Velanker, S. S. Ray, R. S. Gokhale, S. Suma, H. Balaram, P. Balaram and M. R. N. Murthy, *Structure*, 1997, **5**, 751-761.
14. G. Jogl, S. Rozovsky, A. E. McDermott and L. Tong, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 50-55.

15. S. Lara-González, P. Estrella-Hernández, A. Ochoa-Leyva, M. Del Carmen Portillo-Téllez, L. A. Caro-Gómez, E. E. Figueroa-Angulo, H. Salgado-Lugo, J. F. T. Miranda-Ozuna, J. Ortega-López, R. Arroyo, L. G. Briebe and C. B. Benítez-Cardoza, *Proteins*, 2014, **82**, 22-33.
16. R. K. Wierenga, M. E. M. Noble, G. Vriend and S. Nauche, *J. Mol. Biol.*, 1991, **220**, 995-1015.
17. E. Maldonado, M. Soriano-García, A. Moreno, N. Cabrera, G. Garza-Ramos, M. T. de Gómez-Puyou, A. Gómez-Puyou and R. Pérez-Montfort, *J. Mol. Biol.*, 1998, **283**: 193-203.
18. Y. Aguirre, N. Cabrera, B. Aguirre, R. Pérez-Montfort, A. Hernandez-Santoyo, H. Reyes-Vivas, S. Enríquez-Flores, M. T. de Gómez-Puyou, A. Gómez-Puyou, J. M. Sanchez-Ruiz and M. Costas, *Proteins*, 2014, **82**, 323-335.
19. M. Costas, D. Rodríguez-Larrea, L. De Maria, T. V. Borchert, A. Gómez-Poyou and J. M. Sanchez-Ruiz, *J. Mol. Biol.*, 2009, **385**, 924-937.

Side chain flexibility and coupling between the S4-S5 linker and the TRP domain in thermo-sensitive TRP channels: Insights from protein modeling

Sergio Romero-Romero,¹ Froylan Gomez Lagunas,² and Daniel Balleza^{2*}

¹ Facultad de Medicina, Departamento de Bioquímica, Universidad Nacional Autónoma de México, Mexico city, MEXICO

² Facultad de Medicina, Departamento de Fisiología, Universidad Nacional Autónoma de México, Mexico city, MEXICO

ABSTRACT

The transient receptor potential (TRP) superfamily is subdivided into several subfamilies on the basis of sequence similarity, which is highly heterogeneous but shows a molecular architecture that resembles the one present in members of the Kv channel superfamily. Because of this diversity, they produce a large variety of channels with different gating and permeability properties. Elucidation of these particular features necessarily requires comparative studies based on structural and functional data. The present study aims to compile, analyze, and determine, in a coherent way, the relationship between intrinsic side-chain flexibility and the allosteric coupling in members of the TRPV, TRPM, and TRPC families. Based on the recently determined structures of TRPV1 and TRPV2, we have generated protein models for single subunits of TRPV5, TRPM8, and TRPC5 channels. With these models, we focused our attention on the apparently crucial role of the GP dipeptide at the center of the S4-S5 linker and discussed its role in the interaction with the TRP domain, specifically with the highly-conserved Trp during this coupling. Our analysis suggests an important role of the S4-S5L flexibility in the thermo-sensitivity, where heat-activated channels possess rigid S4-S5 linkers, whereas cold-activated channels have flexible ones. Finally, we also present evidence of the key interaction between the conserved Trp residue of the TRP box and of several residues in the S4-S5L, importantly the central Pro.

Proteins 2017; 85:630–646.
© 2017 Wiley Periodicals, Inc.

Key words: S4-S5 linker; TRP box; local flexibility; allosteric coupling; tryptophan interactions.

INTRODUCTION

The superfamily of transient receptor potential (TRP) channels conforms one of the largest groups of ion channels. Despite its diversity, all TRP channels share the common property of functioning as polymodal cellular sensors.^{1,2} TRPs are commonly grouped into six subfamilies, namely: TRPC for “canonical” or “classical” (TRPC1–7), TRPV for “vanilloid” (TRPV1–6), TRPM for “melastatin” (TRPM1–8), TRPP for “polycystin” (TRPP2–5), TRPML for “mucolipin” (TRPML1–3), and TRPA for “ankyrin” (TRPA1). All known TRPs are activated by many, polymodal stimuli and often form heterotetrameric, cation selective channels with a large diversity of functions. In mammal cells, these channels are implicated in sensory pathways responding to temperature, mechanical forces, pH, odors, tastants, and auditory stimuli.^{2,3} In particular, thermo-sensitive TRPs

are classified as channels whose activation directly responds to thermal stimuli, including heating (as in TRPV1, -V2, -V3, -V4, TRPM2, -M3, -M4, -M5) or cooling (TRPM8, TRPA1, and TRPC5).² Moreover, although channel activation is poorly voltage-dependent,

Additional Supporting Information may be found in the online version of this article.

Daniel Balleza's current address is Departamento de Química ICET, Universidad Autónoma de Guadalajara, Patria 1201, Lomas del Valle, 45129 Zapopan Jalisco, México

Grant sponsor: DGAPA-UNAM; Grant number: IN220916; Grant sponsor: ConaCyT (to E.G.L.); Grant number: 153504; Grant sponsor: ConaCyT (to S.R.R.); Grant numbers: 313076 and 355134; Grant sponsor: Posgrado en Ciencias Bioquímicas, UNAM.

*Correspondence to: daniel.balleza@mail.uag.mx Dr. Daniel Balleza, Departamento de Fisiología, Facultad de Medicina, Torre de Investigación, 1er. Piso. Laboratorio de Canales Iónicos, Col. Copilco Universidad, 04510, Cd. Universitaria, Ciudad de México. E-mail: daniel.balleza@mail.uag.mx

Received 16 August 2016; Revised 1 December 2016; Accepted 3 January 2017
Published online 9 January 2017 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/prot.25243

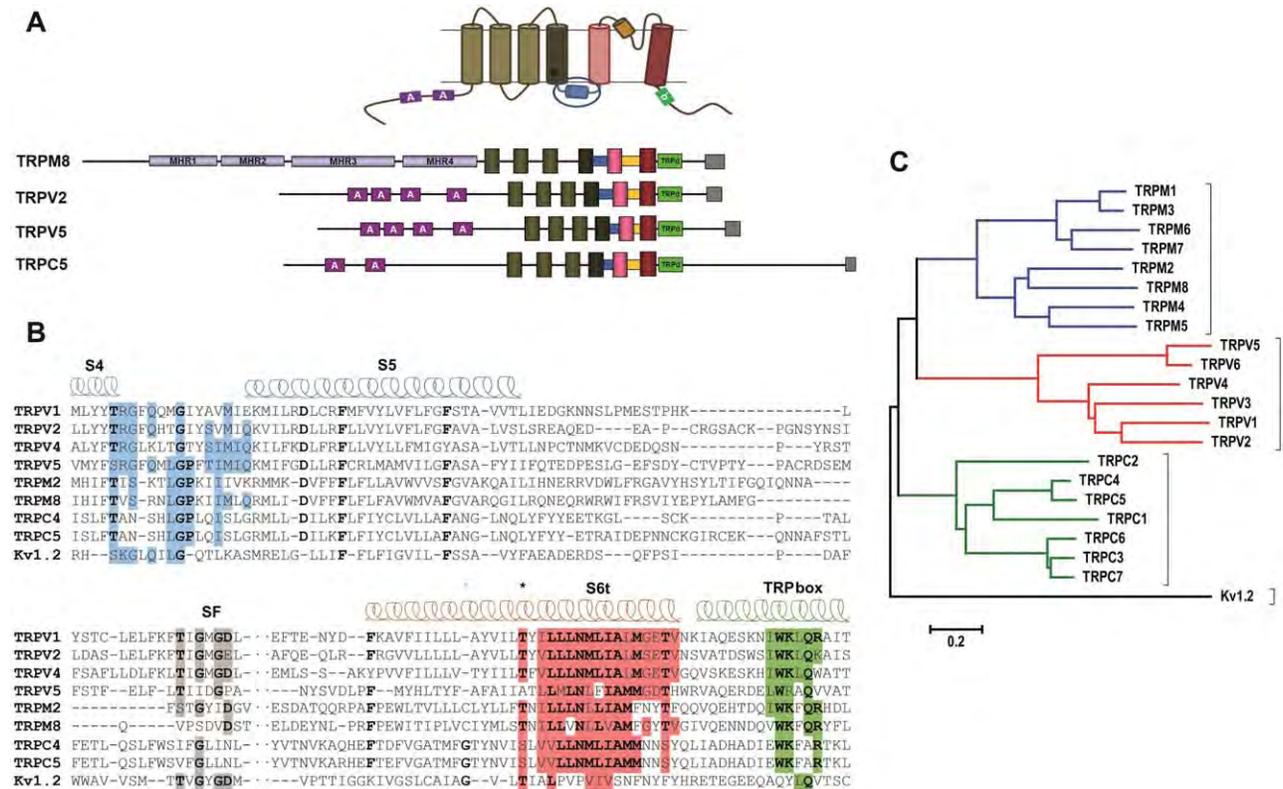


Figure 1

A. Transmembrane topology of TRP channels and location of various structural domains and protein interaction sites found in TRP channels. A ankyrin-like repeats (purple), *MHR* melastatin homology regions (silver), *TRPd* TRP domain (green). Positions of these domains are proportionally represented. B. Alignment of the S4-S5L, S6, and the TRP domain in several TRP and Kv channels. The more conserved sites at the S4-S5L (cyan), selectivity filter (gray), the S6t segment (red), and the TRP box (green) are indicated. Gray asterisk shows the position where a glycine (bold letter) forms a hinge point in Kv channels. In TRP channels, an equivalent point of high flexibility is located one helical turn below (black asterisk). C. Phenogram of TRPs channels. The tree was inferred using the Neighbor-Joining method and was built in MEGA6 (Ref. 12). Scale bar represents evolutionary distance expressed as the number of substitutions per amino acid. [Color figure can be viewed at wileyonlinelibrary.com]

it has been demonstrated that some members of TRPV, TRPM and TRPC subfamilies exhibit voltage dependency and, in consequence, voltage can also be considered as a partial activator of these channels.⁴⁻⁶

As in Kv channels, TRPs are intrinsic membrane proteins formed by six transmembrane (TM) segments (S1-S6), with the pore domain formed by segments S5, S6, and the loop containing the selectivity filter that connects them.⁷ The intracellular amino (-N) and carboxy (-C) termini are involved in the modulation of channel function, membrane trafficking, and their lengths vary significantly between subfamilies.^{1,8} In Kv channels, the voltage-sensor modulus is encompassed by segments S1 to S4, which responds directly to the membrane electric field, through the movements of the four S4 segments, considered the core of the V-sensor.^{9,10} These movements are then coupled to the motion of the pore's activation gate, which implies a physical interaction between both modules, carried out through the intervention of the S4-S5 linker (S4-S5L).^{10,11} In a similar manner, at least in members of TRPV, TRPM, and TRPC channels,

an equivalent interaction is present between S4-S5L residues and the distal part of S6 (S6t), including an intracellular domain known as the TRP domain, not present in Kv channels (Fig. 1).¹³⁻¹⁶

The recently determined cryo-EM structures of TRPV1 in different conformational states^{17,18} and TRPV2 from diverse organisms^{19,20} offer valuable information to understand the nature of this interaction. Thus, these structures show that: (i) the lower, C-terminal, part of S6 and the conserved TRP domain interacts with the S4-S5L; (ii) the α -helical TRP domain is located parallel to the inner leaflet of the membrane; (iii) an invariable Trp residue near the middle of this TRP domain forms an H-bond with the carbonyl oxygen of a Phe residue (F559) located at the beginning of the S4-S5L in TRPV1; and (iv) in TRPV2, the middle inner part of the S6 segment switches between two different arrangements during gating, namely a low-energy α -helix and a high-energy π -helix configurations, which contribute to gate the pore.¹⁹ This subtle feature has been proposed to be an alternative to the Gly-hinge and the Pro-kink

motifs which in Kv channels facilitate the bending and opening of the internal gate.^{21,22}

In a previous work, we described the regional flexibility of the S4-S5L and the terminal part of S6 implicated in the electromechanical coupling in several families of Kv and HCN channels.²³ Herein, taking advantage of the recently reported structures of two different conformational states of TRPV1, and based on molecular modeling using TRPV2 structure as template, we generated new protein models for TRPV5, TRPM8 and TRPC5 channels. Then, analyzing the residue conservation of the S4-S5L, the C-terminal part of the S6 segment and the TRP domain, we correlated the encoded intrinsic flexibility with the detailed structural information provided by the models to improve our understanding of the allosteric coupling between the sensor domain and the pore domain of these channels. In particular, we have found a correlation between S4-S5L flexibility and thermosensitivity in those channels.

METHODS

Database screening of homologous thermo-TRP channels

For our analysis, we selected 1814 sequences searching for homology using the Basic Local Alignment Search Tool for each subfamily. We used a non-redundant protein sequences database and organized all the sequences in three different groups. For each studied group, we chose, as a query, sequences that had been previously studied with respect to the interaction between the S4-S5L and S6t: (1) TRPV1-6, 609 sequences, query: TRPV1 channel from rat (*gi*: 14010883, version: NP_114188.1, Uniprot: O35433.1); (2) TRPM1-8, 741 sequences, query TRPM8 from rat (*gi*: 52851375, version: NP_599198.2, UniProt: Q7Z2W7); (3) TRPC1-7, 464 sequences, query: TRPC5 channel from mouse (*gi*: 3046909, version: AAC13550.1, UniProt: Q9QX29).

Multiple sequences alignment, sequence logos, and phenogram

Alignments were assembled using the CLUSTAL-W algorithm.²⁴ Location of TM segments studied was compared using the aligned profiles with the Uniprot server and contrasted them with previous reports on literature. Sequence alignments of interest were graphically represented with a sequence logo using the WebLogo server²⁵ for the graphical representation of amino acids, displaying the grade of conservation through the height of the letter. To estimate intrinsic flexibility for each subfamily, we generated consensus by this mean, taking the most conserved residue for such estimation. The solved structural data were: TRPV1^{17,26} (PDB codes: 3J5P, 3J5Q, 5IRZ, 5IRX), TRPV2^{19,20} (5AN8, and 5HI9), and TRPV6²⁷

(5IWK). The phenogram was inferred using the Neighbor-Joining method. Evolutionary distances were computed using the Poisson correction method.²⁸ Evolutionary analyses were conducted in MEGA6.¹²

Estimation of flexibility indices for protein sequences

To assess the flexibility of the S4-S5L and the S6t-TRP box, flexibility plots were generated analyzing windows corresponding to these domains in the selected protein sequences. Therefore, we estimated the intrinsic flexibility of a residue in accordance with preliminary estimations,²⁹ and additionally, the amino acid flexibility index was estimated calculating the mean B-factor for each of the 20 amino acids, depending on whether the two adjacent residues were rigid or flexible, comparing them with a Gumbel distribution using a dataset previously reported.²⁹ Finally, as the flexibility of a residue is strongly influenced by the solvent accessibility of the adjacent neighbors and as mobile portions of proteins often had high solvent accessibility,³⁰ we incorporated this information to our analysis.

Protein modeling of the thermoTRP TM region and model validation

Protein modeling was done with the TM region sequence and ~30 residues in the C-terminal containing the TRP domain. All TRP channels were modeled using I-TASSER.³¹ Initial modeling was performed from all structure templates identified by LOMETS³² from the PDB library. In all cases, truncated TRPV2 structure¹⁹ was the template of the highest significance in the threading alignments (the significance of which was measured by the Z-score: >0.88 for the three models). This template was used for modeling TRPV5 (*gi*: 31377510, version: NP_446239.2, UniProt: Q9JIP0), and the two query sequences for: (i) TRPM8 and (i) TRPC5 channels (see above). From this target, I-TASSER generated a large ensemble of structural conformations and the top five 3D models were examined. Among them, the best model based in a highest confidence value was selected for each protein. The confidence of each model was quantitatively measured by the C-score and the TM-score, respectively: 0.35/0.79 for TRPV5, 0.62/0.83 for TRPM8, and 0.38/0.76 for TRPC5. The models were then subjected to conjugate gradient energy minimizations using Amber 16 package with Amber ff99SB-ILDN protein force field.³³ After generating the 3D model, structure validation and stereo-chemical analyses were performed using three different evaluation and validation tools: PROCHECK,³⁴ WHAT_CHECK,³⁵ and the ERRAT server (<http://services.mbi.ucla.edu/ERRAT/>). From these analyses, the stereo-chemical quality of the protein structures was assessed by analyzing residue-by-residue geometries. Overall, structure geometry results in that >95% of the

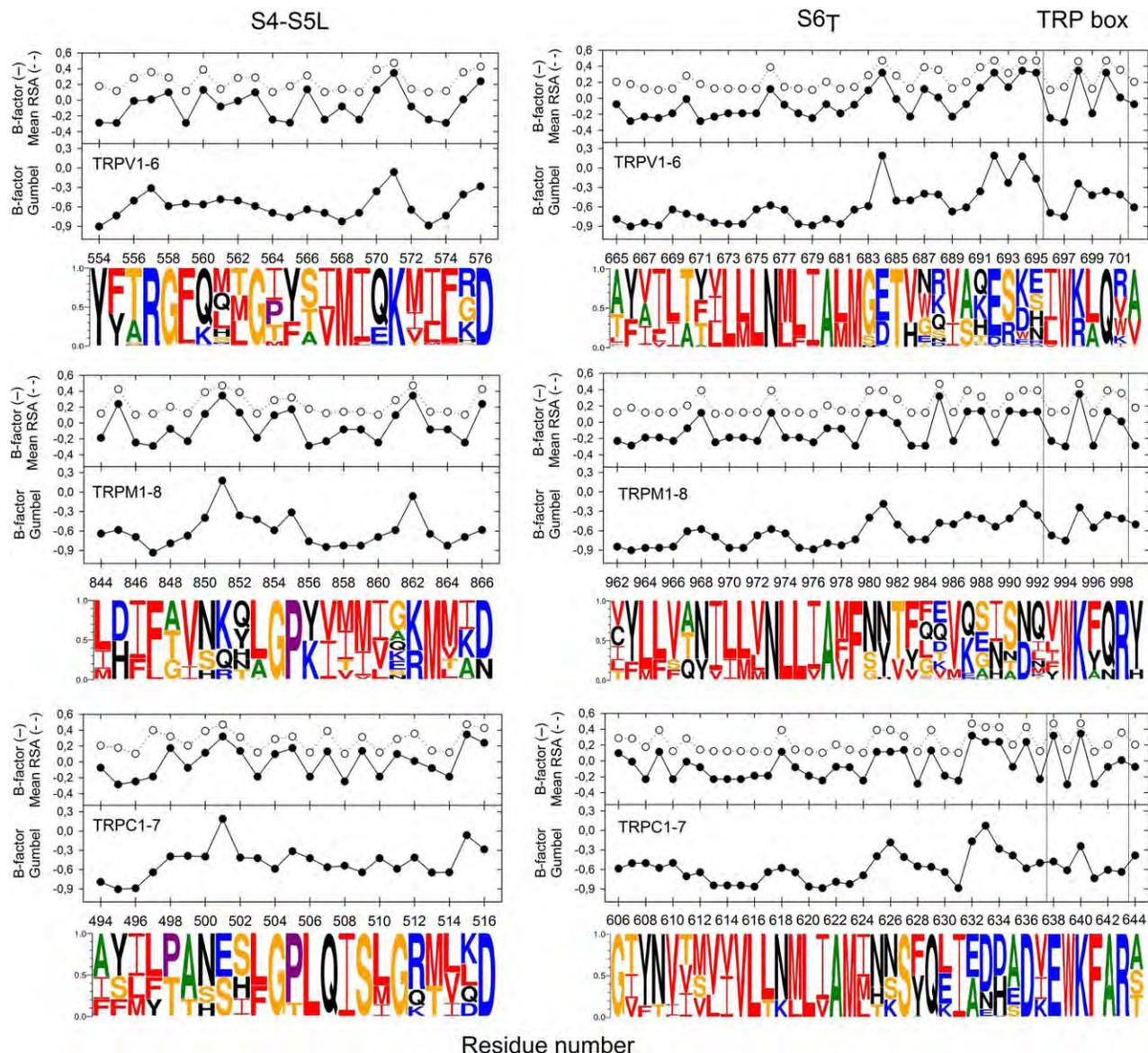
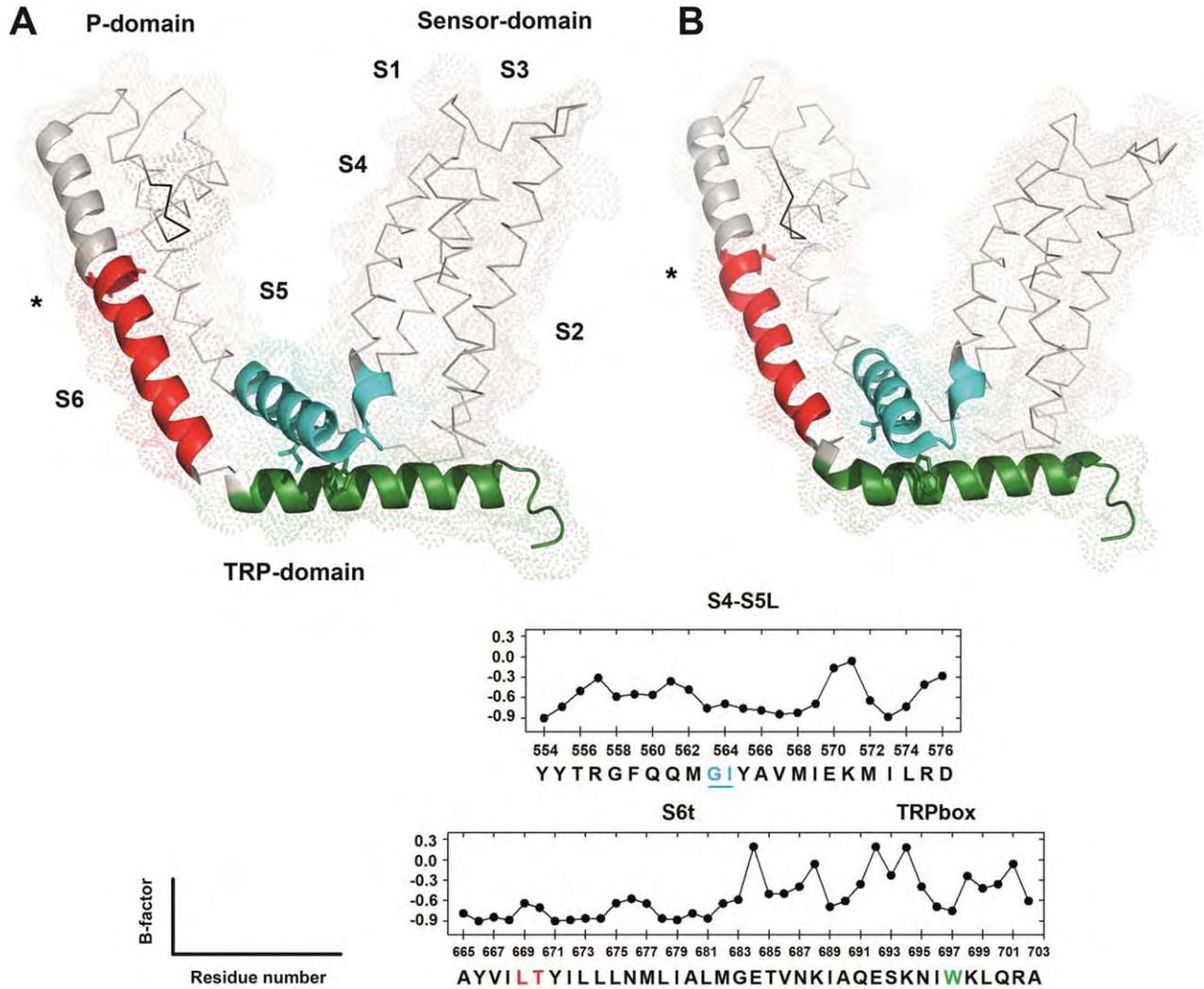


Figure 2

Sequence conservation and local flexibility index for the S4–S5L and S6T segments of TRP subfamilies -V, -C, and -M. *Upper* plots depict mean B-factors and mean relative solvent accessibility for each amino acid according Ref. 30. *Middle* plots show B-factors normalized to a Gumbel distribution according Ref. 29. *Lower* panel depicts sequence logos of the multiple sequence alignments analyzed by the WebLogo server to calculate the relative frequency of amino acids (Ref. 25). The height of a letter is proportional to the relative frequency of that residue at a particular position. Residue numbering is according query sequences (see Methods). Helix-promoting Ala is shown in green; helix-disrupting Pro in purple; highly flexible residues are colored in orange (G,S,T) and blue (K,R,D,E); Cys, Tyr, Asn, Gln, and His, encompassing flexible and rigid residues are shown in black; nonpolar, rigid residues (M,L,V,I,F,W) are shown in red.

residues were in correct regions in preferred/allowed regions in Ramachandran plots for the three models. All results exposed here exhibited high quality of homology modeling and reasonable accuracy in terms of dihedral distribution and stereo-chemical features. Finally, the generated models were submitted and deposited in the Protein Model Database (PMDb) with PMDb identifiers PM0080542, PM0080543, and PM0080544 for TRPV5, TRPM8, and

TRPC5, respectively. Structural comparisons and analyses were done using the following structures: TRPV1 in its fully closed (3J5P, 5IRZ), intermediate-open (3J5R), and in open (3J5Q, 5IRX) states,^{18,26} as well as with a truncated TRPV2 in its putative desensitized state (5AN8)¹⁹ and with a full-length ortholog (5HI9).²⁰ Structures were visualized using the PyMOL Molecular Graphics System, v1.7.2. Graphs were prepared using Sigmaplot v.11.1.

**Figure 3**

Conformational changes of a TRPV1 monomer involving a flexibility hinge at S6 based on the structure of TRPV1 determined by cryo-EM in closed (A), and open state in complex with DkTx and RTx (B). Sensor- (S1–S4) and pore domains (S5–P-S6) are depicted. The S4–S5L is colored in cyan, terminal part of S6 (S6t) in red, and TRP domain in green. The GI dipeptide at the S4–S5L (underlined, cyan), the dipeptide at the S6t starting at position L669 (underlined red and asterisk), and the conserved Trp at the TRP box are shown as sticks. Lower panel shows the flexibility profile for the S4–S5 and the S6–TRP domain linkers with the structural key features in corresponding colors according to Ref. 29. [Color figure can be viewed at wileyonlinelibrary.com]

RESULTS

Identification of the homology regions

The whole channel dataset analyzed consisting of 1814 sequences, represents three different families of thermo-sensitive TRP channels. Figure 1 depicts the multiple ClustalW alignments, going from the last part of S4 to the TRP box of members of TRPV, TRPM, and TRPC families, as well as taxonomic relationships among all members of each family, as indicated in the phenogram. From this diversity, we focused our study on: (i) the cannabidiol and heat-activated TRPV2 channel for which the structure of two different orthologues is now known;^{19,20} (ii) TRPM8, a channel activated by low temperatures as well

as by plant-derived compounds, like menthol and eucalyptol^{36,37}; and, (iii) the canonical stretch-sensitive and cold-activated channel TRPC5, which is also activated on stimulation of receptors coupled to G α -dependent signaling pathways, and receptor tyrosine kinases.^{3,38} Finally, for reasons stated below, we included TRPV5, another member of the vanilloid family in our analysis. We focused on these proteins in an attempt to establish common structural elements relevant to TRPs activation gating in four phylogenetically distant channels.

Estimation of the intrinsic flexibility

B-factors reflect protein internal dynamics, providing a flexibility map of the ground-state protein conformation.³⁹

Table I
Correlation Between Mean B-Factor and Sequence Identity for TRP Channel Subdomains

Channel	Sequence	Mean Bf	Identity (%) ^a			
			TRPV1	TRPV2	TRPV5	TRPM8
S4-S5L						
TRPV1	RGFQQMGIYAVMIE	-0.6003	—			
TRPV2	RGFQHTGIYSVMIQ	-0.5889	71.4 (54)	—		
TRPV5	RGFQMLGPFTIMIQ	-0.5744	50.0 (44)	57.1 (43)	—	
TRPM8	TVSRNLGPKIIMLQ	-0.4956	18.2 (15)	50.0 (15)	66.7 (29)	—
TRPC5	TANSHLGPLQISLG	-0.5184	25.0 (8)	66.7 (13)	50.0 (21)	46.2 (25)
S6t-TRP box						
TRPV1	LTYILLNMLIALMGETVNKIAQESKNIWKLQR	-0.5191	—			
TRPV2	LTYVLLNMLIALMSETVNHVADNSWSIWKLQK	-0.5671	72.7 (122)	—		
TRPV5	ATLLMLNLFIAMMGDTHWRVAQERDELWRAQV	-0.5494	44.8 (77)	37.9 (56)	—	
TRPM8	STNILLVLLVAMFGYTVGIVQENNDQVWKFQR	-0.5396	43.8 (80)	40.6 (74)	37.9 (61)	—
TRPC5	GTYNVISLVLLNMLIAMMNNSYQLIADHADIEWKFAR	-0.6141	39.4 (64)	39.4 (66)	33.3 (53)	33.3 (61)

^aNon-intersecting alignments between two sequences using the SIM alignment tool for protein sequences (<http://web.expasy.org/sim/>). Scores are shown in parentheses.

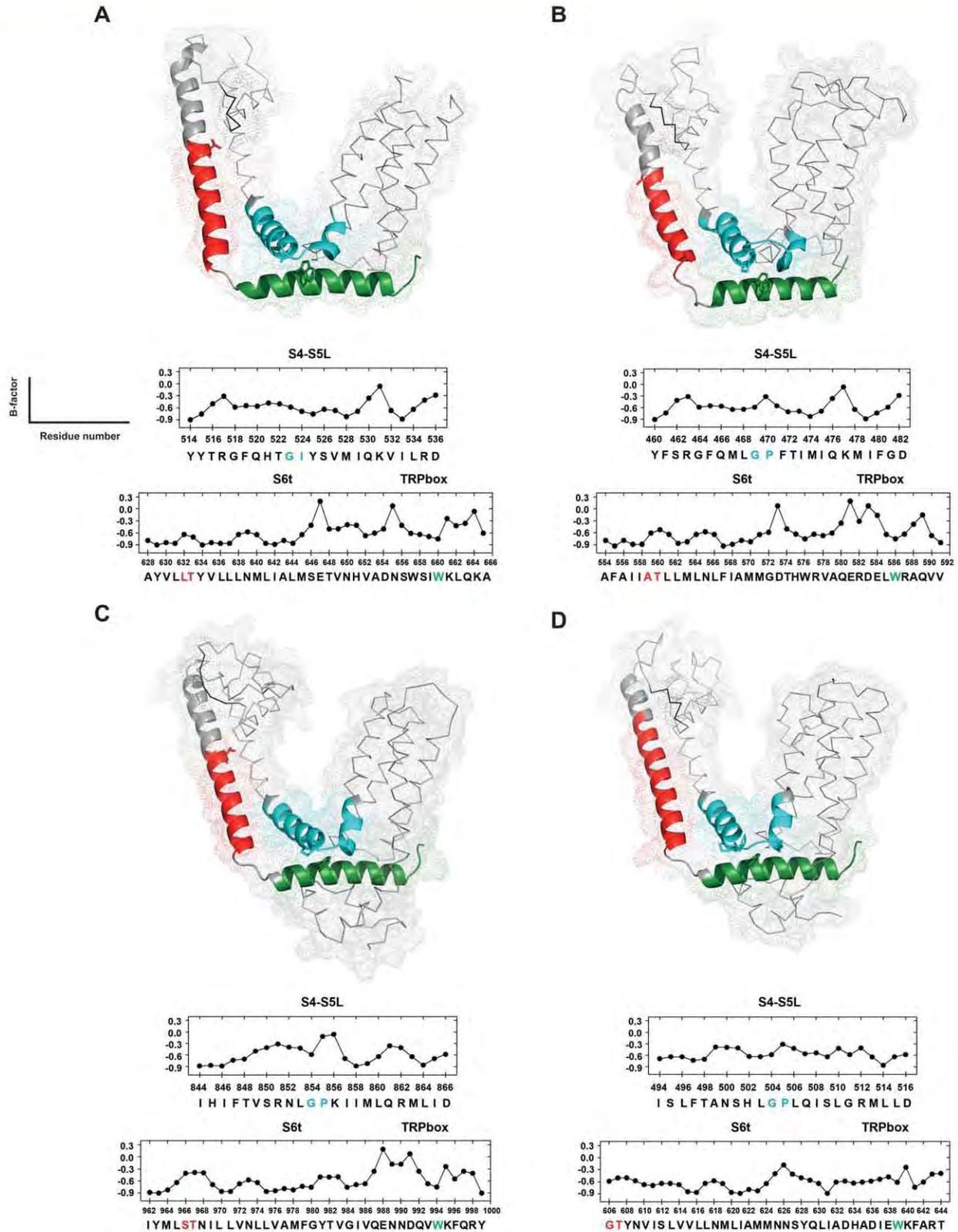
To examine the intrinsic flexibility concomitant to the S4-S5L and S6t subdomains, Figure 2 associates a group of six sequence logos with the B-factor of the C_α atoms derived from X-ray crystallographic studies (black symbols). We also considered the relationship between residue flexibility and mean relative solvent accessibility (RSA, open symbols) which are two strongly correlated parameters.³⁰ Additionally, we included the Gumbel B-factor, which comprises the contribution of the two neighboring residues to have an accurate estimate of the flexibility for a specific amino acid.²⁹ According to this, ten residues are considered as intrinsically flexible (G, S, T, D, N, R, Q, E, and K), whereas the rest are rigid. Proline is an atypical amino acid as it has been considered either flexible²⁹ or exceptionally rigid, in regards to other amino acids.⁴⁰ B-factors and mean RSA shown on top of the sequence logos are those corresponding to the top (consensus) aminoacids (see Methods). Note that, overall, both measurements yield comparable results. Thus, Figure 2 provides a richer, statistically grounded description of the compared primary-sequence segments, in contrast to simple consensus-sequences, allowing the comparison of the segmental flexibility of each sequence against the relative conservation of each position. Thereafter, from the TRP diversity illustrated in Figures 1 and 2, and with the aim of determining common structural elements relevant to the activation gating among these TRP subfamilies, we focused our attention in the relative maxima of specific B-factors as these positions are intrinsically flexible.

The S4-S5L of thermo-sensitive TRP channels

The S4-S5L forms an α -helical structure, which physically links the V-sensor domain to the pore modulus in Kv channels.^{9–11} The recent solved structures of TRPV1 and TRPV2 channels determined by cryo-EM also confirm the helical nature of S4-S5L in these proteins, as well as its proper location to constrict or expand the S6

inner helical bundle of the pore, and its close proximity to the TRP domain (Fig. 3).^{17,19,20} A remarkable characteristic of the selected protein families studied here is the presence of a conserved Gly located at about the middle of the S4-S5L, an amino acid that imprints significant flexibility to polypeptide chains (Fig. 2 and see below). Therefore, to analyze this region better, and considering that TRPV1 and V2 have Ile linked to this conserved Gly, while in contrast TRPM8 and TRPC5, have Pro instead of Ile, we included TRPV5 in our analysis because it also has a Gly-Pro dipeptide. As Pro residues create prominent bends in proteins and thus are important α -helix disruptors, such as Gly,^{40,41} we analyzed TRP sequences showing the GP dipeptide in more detail. Hence, in order to correlate the high intrinsic GP flexibility with the allosteric coupling through its interaction with the TRP domain,^{15,42} we compared it with Gly-Ile and Gly-X dipeptides which conform more rigid linkers (Fig. 3 and Table I). Therefore, taking advantage of protein modeling, we examined the flexibility contribution of the persistent presence of such conserved GP dipeptides at the middle of the S4-S5L in TRPV5 (G469), TRPM8 (G854), and TRPC5 (G504).

In the TRPV family, the Pro residue is only present in members of TRPV5 and TRPV6 subfamilies (31% of our TRPV database) whereas the same position can be replaced by Ile (41%), Thr (22%), Met (6%), and Val (~1%). In the TRPC family the central Gly precedes a Thr residue in 16% of the occurrences, and only in members of the subfamily TRPC2. Finally, in the TRPM family the GP dipeptide is absolutely conserved (Fig. 2). These findings are consistent with the relations shown in Figure 1. Thus, to gain more insight about the putative role of the GP dipeptide at the middle of S4-S5L, as well as on the possible roles of the alternative residues taking the position of that Pro residue, we wanted to explore the contribution to flexibility of alternative residues at those positions. Figures 4 and 5 highlight the

**Figure 4**

Comparison between the structure of a subunit of TRPV2 (A) which was used as a template to generate protein models for TRPV5 (B), TRPM8 (C), and TRPC5 (D). Equivalent dipeptides at the S4-S5L and S6t as well as the Trp at the TRP box are depicted as sticks. Plots show B-factors according Ref. 29. Colors are shown as in Figure 3. [Color figure can be viewed at wileyonlinelibrary.com]

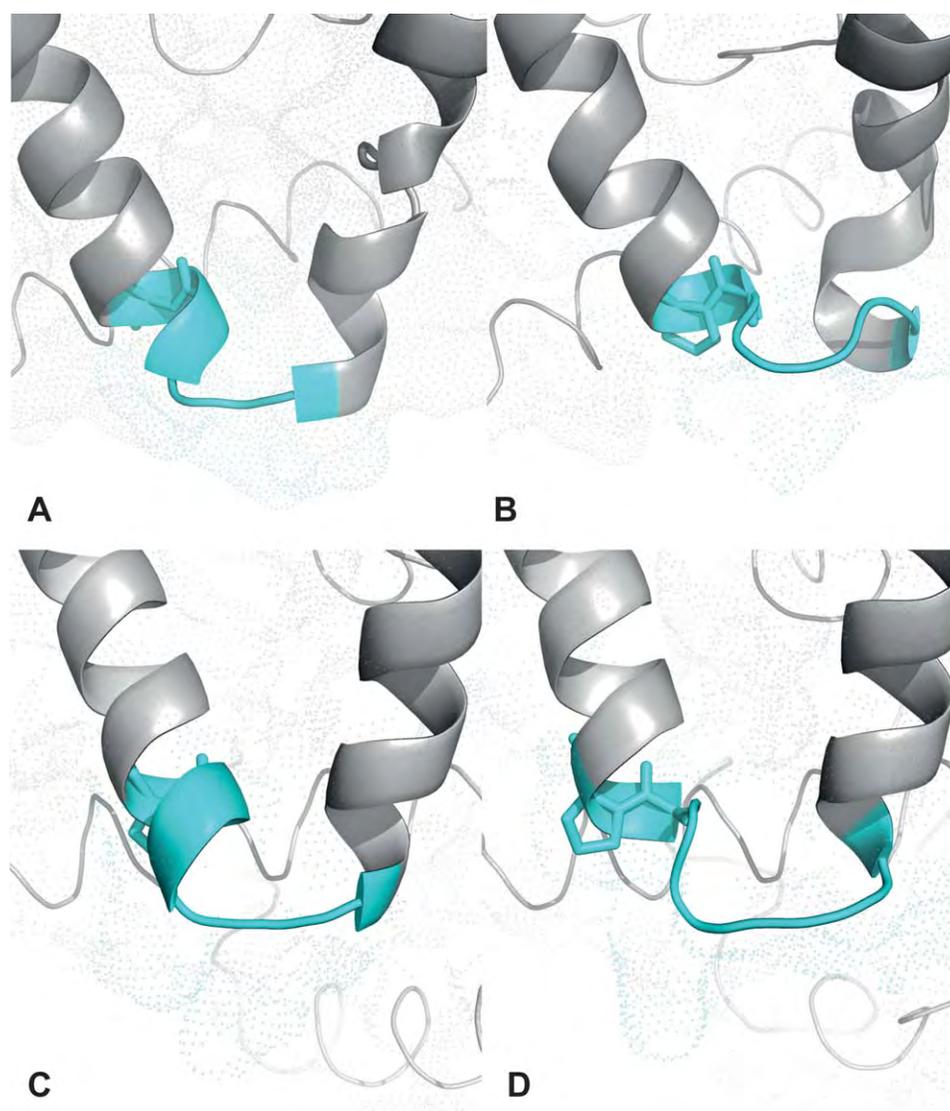


Figure 5

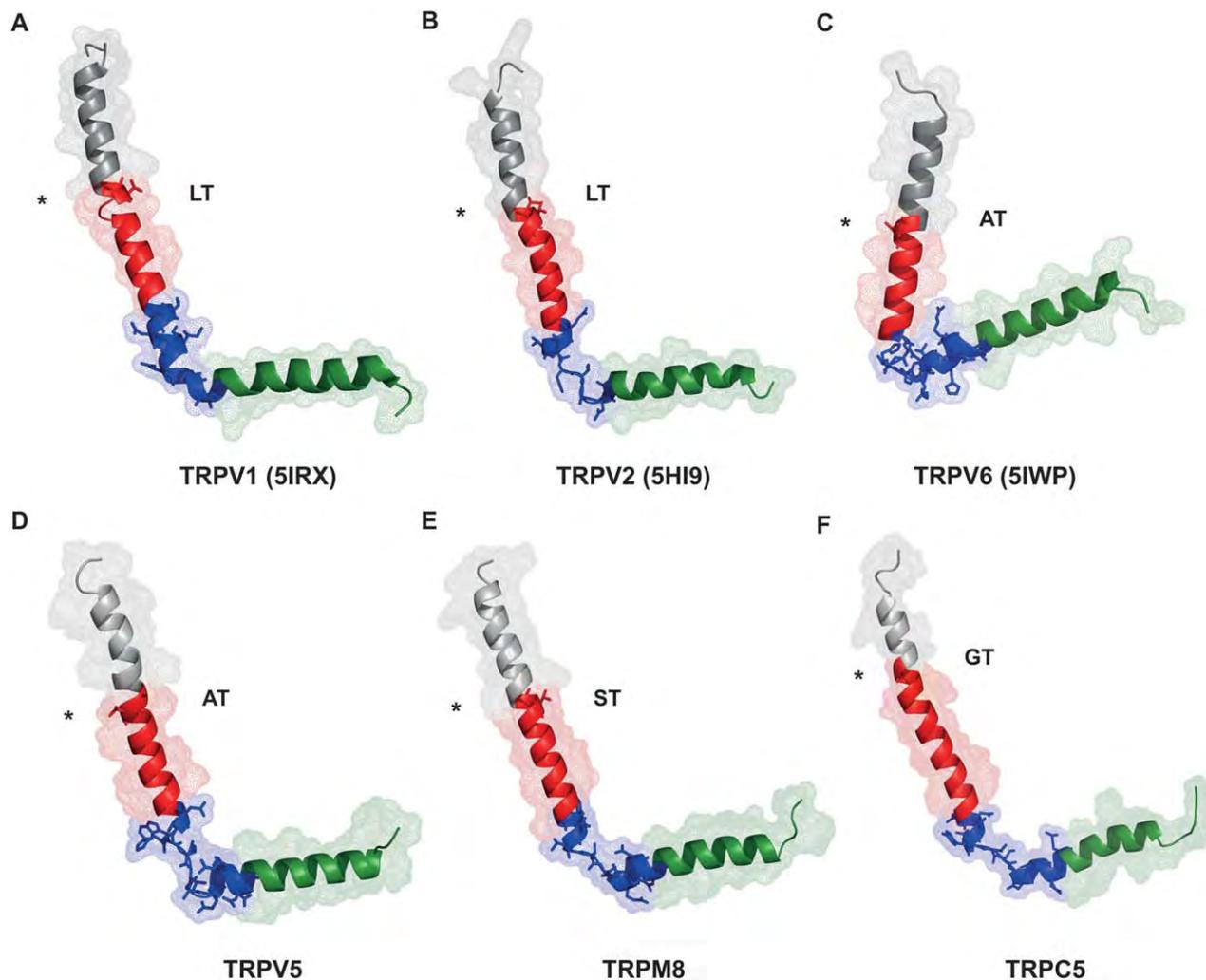
Comparison of the S4-S5 linkers. GI dipeptide present in TRPV2 induces a 3-residue loop (A). GP dipeptide increases the length of the loop: 4-residues in TRPV5 (B) and TRPM8 (C), and 6-residues in TRPC5 (D). Dipeptides are depicted as sticks, respectively. [Color figure can be viewed at wileyonlinelibrary.com]

contribution of the GP dipeptide to the linker flexibility. Note that the presence of proline favors the formation of more flexible connecting loops, at the bending of the S4-S5L, with variable and larger lengths: four (TRPV5, TRPM8), or six (TRPC5) residues, as opposed to the 3-residue loop present in TRPV2 (GI dipeptide). These results and the recent cryo-EM structures showing that these connecting loops change depending on the protein conformation of the pore,^{18,26} highlighting the dynamic role of S4-S5L during the activation gating of TRP channels (Supporting Information Fig. S1). Besides, we decided to build and validate three additional models for TRPC2 and TRPV4 which show different dipeptides. We found that TRPC2 (GT, 8 residues) connecting-loop is

even longer, and more flexible than in TRPC5, whereas TRPV4 (GT, 3 residues) and TRPV4 from *Xenopus* (GM, 3 residues) have shorter, more rigid loops, similar to TRPV2 (*data not shown*). In sum, depending on the sub-family, the central Pro in S4-S5L changes to alternative residues (Ile, Thr, or Met) endowing them with differential intrinsic local flexibilities.

The segment S6 and the S6-TRP domain linker in thermo-sensitive TRP channels

Regarding S6, we noted that the highly conserved Gly-hinge in S6 of Kv channels^{9,21,22} is apparently present only in the TRPC family (Gly 606 of TRPC5), as part of

**Figure 6**

Comparison of S6-TRP domain linkers according to the solved structures of TRPV1, -V2, and -V6 (A to C), and according to the protein models of TRPV5, -M8 and -C5 (D to F). Putative hinge-like Thr-dipeptides (*) and flexible motifs linking S6 to TRP domain (in deep blue) are depicted as sticks. [Color figure can be viewed at wileyonlinelibrary.com]

the G-[TIV]-[YF] motif. In contrast, the corresponding, and more rigid motif starting at position 665 of TRPVs, is [ATL]-[YF]-[VAI] whereas in TRPM channels, the equivalent sequence changes to [VCI]-[YF]-[LM] at position 962 of TRPM8, also intrinsically rigid (Fig. 2). Thus, in the case of TRP channels, at first sight only sequences corresponding to members of the TRPC family seems to have a recognizable, putative, Gly-hinge near the middle of S6, whereas ~1 helical turn downstream—near the place where the Gly-hinge should appear in Kv channels—⁹ a less conserved Thr/Ser is present at position 611/612 in TRPC channels. In Kv channels, the movable Gly-hinge has been described as the pivot point where pore-opening is facilitated by a helical rotation from the symmetry axis of the pore.^{21,43} Interestingly, the corresponding position in members of TRPV and

TRPM families, is occupied by two analogous sequence segments, which are coincident with the presence of a small peak in the flexibility profile of that region (T670 in TRPV1 and T967 in TRPM8) (Figs. 3 and 4).

To further explore this local point of flexibility, we took advantage of the recent cryo-EM structures obtained for TRPV1 in different conformations,^{18,26} as well as the TRPV2,^{19,20} TRPA1,⁴⁴ and the X-ray crystal structure of the Ca²⁺-permeable TRPV6 channel.²⁷ All of them have provided essential information to understand how different protein subdomains are coupled in TRP channel activation. As expected, the TRPV1 structure reveals slight differences between the ligand-free (apo) and the ligand-bound structures.^{18,26} For this reason, in Figure 3(A) we chose to compare the ligand-free (closed conformation) against the open-pore state

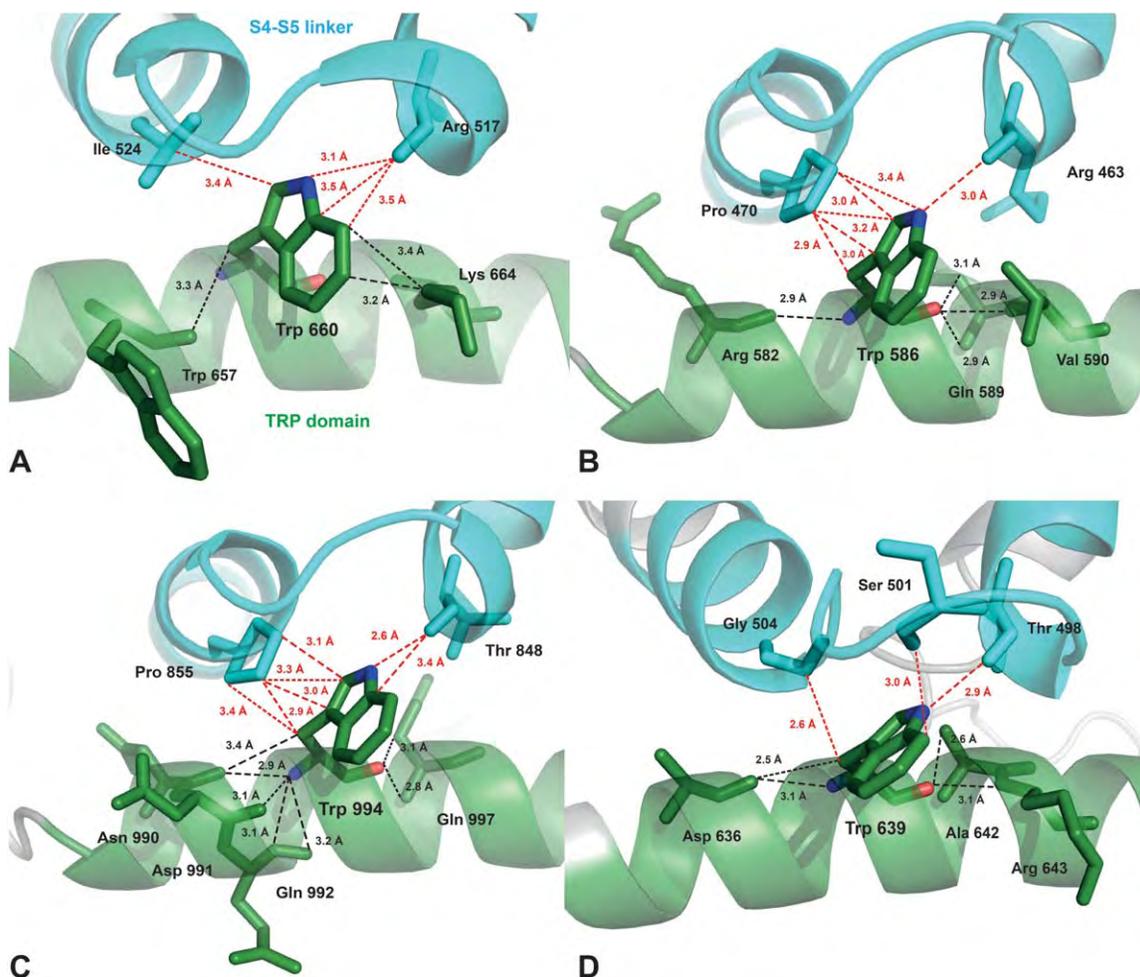


Figure 7

Tryptophan potential interactions in the complex between TRP box and S4-SL according to the structure of TRPV2 (A) in comparison with the corresponding interactions predicted in the protein models for TRPV5 (B), TRPM8 (C), and TRPC5 (D). Black lines depict intradomain interactions whereas interdomain interactions are shown in red lines. [Color figure can be viewed at wileyonlinelibrary.com]

achieved on binding of the vanilloid agonist resiniferatoxin (RTx) from *Euphorbia resinifera* and the bizarre double-knot toxin DkTx from the tarantula. As it can be appreciated, the S6t flexibility plot in TRPV1 predicts a peak around the middle of the S6 helix [see asterisks in Figs. 1(B) and 3], two residues downstream of the position where the Gly-hinge is present in Kv-channels.²¹ The relevant residues in TRPV1 are Leu669/Thr670, and interestingly, in the open conformation a slight kink is evident at that position, consistent with its putative role as a hinge in this channel (Fig. 3). To corroborate this, it was useful to compare the amphipol-stabilized structures shown before against recent structures of the same protein determined in lipidic nanodiscs, where structures represent a more realistic conformation due to the presence of some annular lipids.²⁶ As shown in Supporting Information Fig. S1 and S2, the presence of a S6-inflexion at that position is more evident in presence of

RTx/DkTx. Indeed, it is seen that the activated TRPV1 shows a 3-residue flexible loop composed by Tyr671, Ile672, and Leu673. However, unlike what is seen in the case of TRPV1, we did not find the corresponding S6-hinge in the TRPV5, TRPM8, and TRPC5 models, even when a peak of minimal flexibility is evident in each profile around the middle of the S6 (Fig. 4). The putative hinge is not evident either in the truncated TRPV2, although in the full-length structure the kink corresponding to the flexibility peak predicted in our plots is most conspicuous (Fig. 6). Similarly, in the crystal structure of TRPV6 that irregularity is not evident; however, as it shows a close packing of aromatic side chains that rigidifies the structure, and the S4-S5L shows not interpretable density data, we should interpret this with caution.²⁷

Protein models also show the corresponding regions of high flexibility within the S6-TRP domain linker. In the

solved structures, this region has a significant torsion to form the intracellular helix which runs parallel to the inner layer of the membrane.^{17,18} In TRPV1, the flexible sequence (~12 residues) starting at E684 is rich in flexible residues and encompasses the portion of each monomer where S6 connects the TRP domain. Similarly, in TRPV2 the homolog flexible sequence initiates at E647 but it is smaller (~9 residues), whereas in TRPV5 it is again larger (~12 residues), starting at D573. Finally, in TRPM8 the corresponding flexible sequence is longer (~13 residues) starting at G980, but in TRPC5, it begins at N625 (~14 residues) but it shows a more rigid profile for the S6-TRP domain linker (compare flexibility plots in Fig. 4 with Fig. 6).

The conserved tryptophan at the TRP domain

Finally, taking into account that different allosteric models describe the gating of thermo-sensitive TRP channels,^{4,45,46} and considering that in all of the cases the corresponding region to the TRP domain at the C-end of the inner gate plays a central role along with the S4-S5L during the activation of the channels,^{14,15,47–49} our attention was turned to the putative interactions of S4-S5L with the highly conserved Trp residue present in the TRP box. The results in Figure 7 show that several potential interactions can be established between the Trp residue and the S4-S5L. Thus, in TRPV2, Trp660 is oriented toward the linker interacting with Arg 517 and Ile524 (from the GI dipeptide). Trp could also interact with Lys664 of the same TRP domain; in TRPV5, Trp586 is also oriented to the linker, and the only candidate to establish interactions is Pro470 precisely, of the GP dipeptide, and probably Arg463. Interestingly, in this case Pro470 and Trp586 seems to have the proper geometry to potentially establish several interactions between the pyrrolidine and the indole rings (see Discussion). A similar pattern was observed in the case of TRPM8, where the highly conserved Trp994 is also exposed to the linker and it also has the proper orientation to establish several interactions with Thr848 and mainly with Pro855 (from the GP dipeptide). The molecular environment changes for TRPC5, where Trp639 runs more parallel to the TRP domain and, in consequence, relevant interactions could be established with Thr498, Ser501, and Gly504 from the GP dipeptide, but not with Pro505 which is farther.

DISCUSSION

In a previous work, we found the presence of a highly conserved Gly residue at the center of the S4-S5 linker in Kv channels, and corroborated the presence of the conserved Gly-hinge at the middle of S6 in all members of the Kv superfamily, and pointed out its absence in the HCN

family.²³ The starting point of the present study was the generation of flexibility plots of several TRP families, where points of relatively maximal flexibility were identified in the two modules that couple the pore, S5-S6, and the sensor, S1-S4, domains. The observations reported here agree well with previous reports showing: (i) a close relation between the TRPV and TRPM subfamilies⁸; (ii) conserved motif sequences at the S4-S5L and near to the channels activation gate with a 6-mer segment (the TRP box) showing an invariant Trp residue which interacts with the S4-S5L^{14,15}; (iii) we have also shown that, as in Kv channels, there is a highly-conserved Gly at the center of S4-S5L, which in the TRPM and TRPC families (except TRPC4 and C5) is always linked to a Pro residue, thus forming a GP dipeptide. TRPC4 and C5 have instead a GT dipeptide, whereas in TRPV channels the central Gly forms less conserved dipeptides: GI, GP, GT, and GM. Additionally, two other points of interest were identified, that we propose to be relevant for the activation gating of these channels, and therefore were studied in more detail: (i) a putative Thr-hinge analogous to the Gly-hinge described in potassium channels, and (ii) a region of high flexibility that connects the S6 segment with the TRP domain, which in turn, allows it to interact with specific residues within S4-S5L. To conclude, as the generated models are consistent with a large amount of functional data, conclusions discussed in next sections have structural relevance.

Proline makes more flexible thermo-TRP S4-S5 linkers

Because of the rigid nature of peptide bonds, the conformational flexibility of polypeptide chains is limited to rotations at the bonds of C α atoms and to the contribution of the side chains mobility.⁵⁰ In α -helical backbones, Gly and Pro are known as “helix breakers” because they disrupt the helix regularity.^{21,41} In this regard, a pivotal role of Gly is to prevent constraining interactions with critical residues in its vicinity,⁴³ whereas Pro induces kinks that also facilitate the helical packing assembly and due to its intrinsic rigidity it may provide significant structural distortions.⁵¹ Furthermore, it has been shown that when Gly precedes Pro, this dipeptide distorts helices by changing the direction of the axis about 20°.⁵² Hence, herein we explored the contribution of this GP dipeptide to the S4-S5L flexibility of several TRP channels. The B-factors summarized in Table I indicated that when Pro is present in the S4-S5L (as in TRPV5, TRPM8, and TRPC5) there is a bigger segmental mobility than when Ile takes its place (as in TRPV1 and V2). Notably, this estimate was reflected in the specific flexibility plots in Figures 3 and 4, which was consistent with the protein models showing longer loops immediately adjacent to the Pro residue in Figure 5.

Interestingly, the S4-S5L from TRPM8 shows the more flexible profile among the studied proteins (Fig. 4 and

Table II

Thermodynamic Parameters of the TRP Channels Studied in This Work

Channel	T (°C)	Mean Bf(S4-S5L/ S6t-TRPbox)	Q_{10} (°C)	ΔS (kcal/molK)	ΔH (kcal/mol)	Activators	References
TRPV1	>43	-0.5597	27	0.26	90	Capsaicin, RTx, Piperine, Endocannabinoid lipids	37,58-62
TRPV2	>52	-0.5780	>100	0.59	200	2-APB ^a , LPC ^b , Probenecid, Mechanical stretch	63-66
TRPV5	n.d.	-0.5619	n.d.	n.d.	n.d.	Constitutive active ^c , PIP ₂	62,67
TRPM8	<20-28	-0.5176	24	-0.38	-112	Menthol, Eukalyptol, Icilin, PIP ₂	37,59,68,69
TRPC5	25-37	-0.5663	~10	n.d.	-40	Gq-PLC ^d , PIP ₂ , Nitric oxide, H ₂ O ₂ , Mechanical stretch	38,70-72
Shaker	—	—	>4	-1.1 ^e	-16	Voltage, PIP ₂	73-75

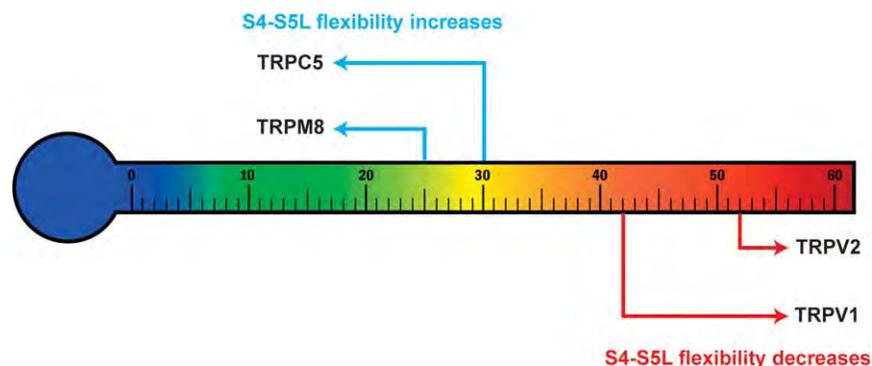
^a2-aminoethoxydiphenyl borate.^bLysophosphatidylcholine.^cRegulated via PKC.^dHeterotrimeric Gq protein subunit that activates phospholipase C.^eEntropic change for the C to O transition at 20°C.

n.d. not determined.

Table I); in contrast, the corresponding profiles for TRPV1 and TRPV2 are the most rigid. TRPM8 channels are activated by cold ($\sim 28^\circ\text{C}$), whereas TRPV1 and TRPV2 are activated by heat ($\sim 43^\circ\text{C}$ and $\sim 52^\circ\text{C}$ respectively).^{37,46} Thus, it is tempting to hypothesize a direct relationship between a high flexibility S4-S5L profile with activation by cooling as well as a restricted flexibility (i.e., an enhanced rigidity) in heat activated TRP channels (see Scheme). In support of this idea, TRPC5, which has been described also highly cold sensitive ($37-25^\circ\text{C}$)³⁸ exhibits a highly flexible S4-S5L profile (Table I). In terms of protein thermal stability, this relationship has recently received noticeable attention.⁵³ In the case of TRPV5, its S4-S5L flexibility profile was intermediate between those of cold-sensitive (TRPM8 and TRPC5) and hot-sensitive (TRPV1 and TRPV2) channels, but we did not find reports about the temperature sensitivity of this channel. Although we cannot verify our hypothesis in this case, we believe that this channel could be activated by warm stimuli as in the case of TRPV4.^{54,55} The mobile nature of protein loops has also been decoded

in terms of composition, being Lys, Met, and Asp, part of “fast” loops (as in TRPM8 S4-S5L), while Cys, Tyr, Ile Phe, and Leu are often part of static loops (as in TRPV1, and V2),⁵⁶ which is consistent with our results. Nevertheless, although internal motion (flexibility) importantly depends on aminoacid composition, several other factors have to be considered to make a universal postulate about this. Consider for example the heat capacity, conformational entropy, salt bridge networks, electrostatic and dipolar interactions, as well as the hydrophobic effect,⁵³ which cannot be determined here.

We have also noted that S4-S5L from TRPV1 has a stiffer profile than TRPV2, which is not consistent with the T-sensitivity showed by these two channels (see Table I). This can be explained if one also considers the contribution of the interacting S6t segment, since these two subdomains are involved in the intersubunit coupling during the allosteric activation.^{12-15,57} Taking into account this, we found that the S4-S5L/S6T-TRP box profile of TRPV2 (activated by noxious heat) is slightly

**Scheme 1**

Schematic representation of four thermo-TRPs indicating their reported temperature sensitivity associated with their relative S4-S5L flexibility. [Color figure can be viewed at wileyonlinelibrary.com]

more flexible than that of its homologous, TRPV1. The latter is also true at least for TRPM8 because it shows the more flexible combined profile for such interaction, consistent with the aforementioned hypothesis (Table II). Thus, as protein mobility increases with temperature,⁷⁶ and considering that specific and regional protein motions are linked to the efficiency of enzymatic catalysis in some enzymes,⁷⁷ it was interesting to assess the contribution of the intrinsic flexibility of these two sub-domains during allosteric gating of these thermo-sensitive TRP channels.

A S6 Thr-hinge in thermo-sensitive TRP channels?

Previous reports suggest that the lower part of S6 (S6t) in TRPV1 moves to open the activation gate in presence of specific ligands.^{78,79} Supporting this, the new cryo-EM data of TRPV1 show such motion of S6t and also suggest that it may be caused by an outward movement of the S4-S5L.^{17,18} In TRPV2, whose structure has been reported as a putative desensitized state, the intracellular activation gate has been positioned closer to the cytosol by one helical turn in comparison with TRPV1.¹⁹ This peculiarity was interpreted to arise from a change in the helical organization of S6 in both channels as result of two different non-conductive states. Thus, whereas in TRPV1 S6 adopts a π -helix in the middle of this segment, S6 of TRPV2 adopts a canonical α -helix conformation. Such subtle differences have profound implications for this work because residues that define the point at which S6 bend in the TRPV1 structure is Tyr671, just one residue after the LT dipeptide defining the lower gate flexibility peak for this subfamily (Fig. 2). Interestingly, Tyr761 constitute the most constricted point in the lower gate of the TRPV1 channel structure¹⁷ and it might contribute to the allosteric coupling between thermal- and capsaicin dependent activation mechanisms.⁸⁰ However, in TRPV2, the equivalent position (Tyr632) shows no bending point probably because in this case, the protein conformation represents a different closed state. We cannot rule out the possibility of a change in the helical conformation depending on the substate of the channel, but we also cannot discard a minimum contribution around the intrinsic flexibility point we found in our analysis for the three TRP subfamilies reported here.

In Kv channels, experimental evidence indicates that the highly conserved Gly residue at the middle of S6 acts as a local flexibility point.^{11,21,22} This “molecular hinge” has a subtle fluctuation of $\pm 2.5^\circ$ during gating transitions in several K^+ -channels.⁸¹ Moreover, the high-conserved PVP motif in these channels also contributes to the S6 flexibility through a kink angle fluctuation from 18° to 30° during the closed to open transition.⁸¹ Notably, this “Pro-kink” motif aligns near (one helical turn) the flexibility peak in TRPV and TRPM channels (Fig. 1) but in the case of the TRPC subfamily, the corresponding point of flexibility is less defined, ranging from residue 606 to 610. Indeed, we

found an invariant Gly at position 606 in TRPC channels as part of the G-[TIV]-[YF] motif, which is similar to the 3-residue motif in Kv10-12 channels with a high-conserved Gly followed by at least one flexible residue and then a well-conserved rigid one.²³ This motif is around one and a half helical turn above the equivalent position to the putative Thr-hinge (T670) in TRPV channels,⁸⁰ and it shows high local-flexibility in TRPC channels (Fig. 2). However, if Gly606 has a role as a hinge in TRPC channels, it remains to be experimentally determined. In contrast, the corresponding position in TRPV channels (T665 in TRPV1) is low-conserved because it may also be occupied by an Ala or a Leu residue. However, due to its poor local flexibility, it probably excludes that motif as an additional bending point in S6 and reinforces the idea of a Thr-hinge at position 670. Hence, TRP channels do not possess a typical point of local flexibility like Gly or Pro at the middle of S6, except in the structure of TRPA1 where the equivalent position has already a Pro residue as a point of high flexibility.⁴⁴ Indeed, in dynamic simulations performed by Susankova et al. (Ref. 80) bending becomes apparent in the region around T670 in TRPV1 channels, in good agreement with our predictions and very recent structural data (Supporting Information Fig. S1 and S2). Unfortunately, this putative Thr-hinge has been severely questioned since no functional evidence was provided by the authors.^{78,82}

Likewise, a highly conserved Ala at position 680 in TRPV channels, present in TRPM (Ala977) and TRPC (Ala622) channels also have been considered as an alternative “bending-hinge” in those families.⁸⁰ In this particular case, the high helix-stabilizing propensity of Ala relative to Gly, Ser, and Thr, as well as the fact that these Ala residues are immersed in rigid motifs, leads us to not correlate its low-flexibility profile with such hypothetical function (Fig. 2). In sum, we found three different dipeptides showing slight intrinsic flexibility, all of them involving a Thr residue near to the middle part of S6 in the three different subfamilies analyzed here. When such positions were structurally evaluated in our models, a bending point was difficult to be evidenced in S6, and only in the case of TRPC5 a subtle bending can be anticipated at that position (Fig. 6). Nevertheless, as our models were generated using the structure of TRPV2, which represent a desensitized state of that channel, and considering that density maps determined with nanodiscs including annular lipids are in general more close to a physiological condition, the fine details revealed for TRPV1 structures reinforces our prediction for these particular bending points (Supporting Information Fig. S2),²⁶ although we present this aspect of the lower gate cautiously.

The flexible nature of the S6-TRP domain linker

Finally, we evaluated the three different TRP domains in our models and we found that they assume an

α -helical structure that presumably also runs parallel to the inner face of the membrane, just as it was anticipated by the structures of TRPV1, -V2, and -V6.^{17,19,20,27} This domain starts with a sharp bend at the end of S6, having an important role as a tetramerization domain and as a molecular determinant of channel gating.^{47,48} Notably, we found that the higher flexibility profile found for this region corresponds precisely to the S6-TRP domain linker, a subdomain originally described as a sharp bend after S6.¹⁷ In the case of TRPV1 this flexible motif starts at Glu684 (part of segment S6) and it ends at Asn695 (in the TRP domain), which are two highly flexible residues²⁹; then, the flexibility profile goes down where the 6-mer TRP box initiates with a more rigid profile (Fig. 3). This is quite consistent with the cryo-EM data, where electronic density for the TRP domain, especially the part immediately following S6 (the helical “elbow”) does not have a good side-chain resolution, which probably reflects the dynamic nature of this region.¹⁸ Indeed, this channel in closed state shows a S6 α -helix until Ile689 and then a kink appears at Ala690, which has been solved as a loop in the structure (Supporting Information Fig. S1). Thus, TRP domain starts at Gln691 but in the open state the flexible loop disappears which is consistent with a stabilization of the structure in presence of specific ligands.^{18,83} The linker between S6 and the TRP domain was also assessed in the rest of the models, finding that flexibility profiles were equivalent, providing a pronounced kink in the structure at the end of S6 in TRPV5, TRPM8, and TRPC5 (Fig. 6). TRPV6 structure, in contrast, differs markedly because the S4-P-S5 pore domain and the S1-S4 sensor domain of the same protomer are packed against each other.²⁷ As TRPV5 and TRPV6 are unique members of the TRP superfamily and closely phylogenetically related,^{1,3,8} we must mention that the TRPV5 model (based on the relatively distant TRPV2 structure) should be interpreted with this information in mind. Even so, taken all these observations together, our results agree with the general notion of a correlation between a high intrinsic/segmental flexibility in the S6-TRP box linker as a key structural and a functional determinant that is essential for the allosteric coupling of diverse stimuli in these proteins.^{15,46}

Putative interactions for the conserved Trp at the TRP box

Tryptophan has the largest rigid side-chain among the 20 amino acids, so because of its size and mixed hydrophobic/hydrophilic character, it has an important structural role in proteins. Trp also is unique due to: (i) its amphipathic nature enabling it to participate in both nonpolar interactions as well as in H-bonding; (ii) its potential to perform $\text{NH}\cdots\pi$ and $\text{CH}\cdots\pi$ interactions; and (iii) its electrostatic interactions via the inherent

quadrupoles of the aromatic ring.⁸⁴ Furthermore, Trp is an important contributor to the stabilization of the folded structures.⁸⁵ Importantly for our study, a recent study with TRPV4 concluded that specific interactions of the TRP box, involving Trp733 (Trp697 in TRPV1) with the S4-S5L stabilize the closed state of the channel.⁵⁷ Thus, on ligand binding, the TRP box undergoes conformational changes unlocking the TRP domain as a negative control of the gate opening.⁴⁷ In this model, hydrogen interactions between the S4-S5L and the S6t region are disrupted, presumably promoting opening of the gate.¹⁸

These findings suggest that both intrasubunit as well as intersubunit interactions involving this Trp residue are pivotal for an efficient channel gating and also anticipate that such motif may be involved in the functional coupling of different stimuli leading to opening the pore. Regarding this, we have delimited our analysis to the intrasubunit interactions without considering the intersubunit ones; nevertheless, this does not limit our main findings. In Figure 7, the most interesting potential interactions involving such Trp are depicted. The criterion to establish relevant local interactions between the Trp side-chain and the backbone and lateral groups of residues at the S4-S5L was the close proximity which was considered to be below 3.5 Å. This argument includes hydrophobic interactions, van der Waals forces, and hydrogen bonding.⁸⁶ Remarkably, in all the models we found that Trp was always oriented to the S4-S5L, exposing the indole ring and establishing putative interactions with Arg517/Ile524 (TRPV2), Arg463/Pro470 (TRPV5), Thr848/Pro855 (TRPM8), and Thr498/Ser501 (TRPC5) with interatomic distances $\text{N-H}\cdots\text{O}=\text{C}$ ranging from 2.6 to 3.4 Å, which suggest H-bonding. Indeed amino-aromatic interactions have been defined ranging 3.4 to 6 Å.⁸⁷ In addition, some putative $\text{C-H}\cdots\text{O}=\text{C}$ interactions were also revealed in our models involving other different residues in the S4-S5L as well as in the TRP box itself. However, if the major contribution to possibly lock the lower gate in these channels should be assigned to a particular interaction, it should be attributed mainly to the $\text{N-H}\cdots\text{O}=\text{C}$ H-bonding because its energy is greater than the $\text{C-H}\cdots\text{O}=\text{C}$ interaction.⁸⁸ In the case of TRPV5, and TRPM8 Trp also seems to interact with the central Pro at the S4-S5L (Fig. 7). Trp-Pro interactions are energetically strong (~ 7 kcal/mol)⁴⁰ so contact between a Pro and a Trp can result in very stable interactions because Pro has high propensity to interact with Trp through stereo-specific interactions involving the C-H groups of the pyrrolidine ring as a H-bond donor in the Pro and the π electron cloud of the aromatic ring in Trp.⁸⁹ This is also consistent with the hypothetical role of the Trp-Pro interaction as a locking point in TRPV5 and TRPM8 channels. Obviously, the nature of H-bonding and other kind of interactions varies from system to system and determine if the Pro-Trp interactions between the S4-S5L and the TRP box are real necessarily requires to be experimentally explored.

To finish, it is well known that Trp prefers to exist in the interfacial region of lipid bilayers, between the hydrocarbon core and the aqueous milieu, functioning as a molecular anchor in membrane proteins.⁹⁰ The chemical nature of Trp (in this case the benzene ring) allow it to interact directly with the interfacial water at the membrane surface rather than being buried into the hydrophobic core of the lipid bilayer.⁹¹ Similarly, evidence for cation- π interactions between the lipid choline groups and the indole has also been revealed.⁹² Although these preferences could have important consequences on protein modeling, it is well established from the structures discussed here that the position of the conserved Trp residue in the TRP box is almost invariable, running parallel to the membrane and interacting with the intracellular soluble domains (i.e., the S4-S5L) in each case. In consequence, the TRP domain is considered a point of structural integration which facilitates the allosteric coupling between different domains in these channels.^{17,19,20,26,27}

CONCLUSIONS

As polymodal gating implies the action of different sensors that may work independently or in concert to activate a channel, understanding of allosteric coupling controlling TRP channel activation should seriously consider the concept of intrinsic/local flexibility. The most salient contributions of this study can be summarized as follows: (i) A central Pro in the S4-S5L always increases its flexibility. (ii) Thermo-dependence at least in TRPV1, TRPV2, TRPM8, and probably also in TRPC5 is directly related to the flexibility of the S4-S5L. (iii) Segments S6t in TRPV, TRPM, and TRPC channels show different flexibility profiles, with at least one putative hinge point differentially positioned along such helical structure. (iv) The non-covalent interactions of Pro, combined with its intrinsic rigidity due to its backbone torsion, might be the explanation for its putative stabilizing role of the closed/desensitized state in some TRP channels because it could form a lock through its interaction with the highly conserved Trp at the TRP box; this Trp residue is probably interacting with the inner leaflet of the membrane, stabilizing the closed state. The ideas and predictions stated in this work are mapped on the two available structures of TRPV1 in closed and open states (Supporting Information Fig. S2). Integrate these observations and conciliate them with several previous experimental results and structural data to create a consistent model which includes the role of local flexibility, conformational changes associated to ligand-bonding, as well as the response to temperature, voltage, and mechanical aspects of the lipid bilayer, requires further experiments including the electrophysiological approach. Recent independent works by Chugunov et al. (2016)⁹³ and Wen et al. (2016)⁹⁴ using MD have shed new light on early steps

of TRPV1 activation during an increase in temperature. According to these authors, heat induces subtle conformational motions of the structure formed by S6 helix, which exhibit a heat-activated flexibility during the closed state, and bending of the TRP domains, leading to opening of the lower gate. However, the most unexpected phenomenon observed in these studies is the temperature-dependent opening at the level of the upper gate, which corresponds to the selectivity filter of the pore. Detailed analysis show a partial channel opening in the lower gate, being independent of the upper gate mechanics. Hence, during the closed-to-open conformational simulations at high temperatures, both the lower and the upper gates are dynamic, but the upper gate undergoes flickery opening/closing events, and therefore it may open more readily than the lower one during the gating transition.

ACKNOWLEDGMENTS

The authors thank to Diego Granados for advice using Amber-16, to Gustavo Martinez for helpful discussions during the first part of this work, and to Mrs. Josefina Bolado for editing the English version of this manuscript. Authors claim no conflict interest associated with this report.

REFERENCES

- Owsianik G, D'Hoedt D, Voets T, Nilius B. Structure-function relationship of the TRP channel superfamily. *Rev Physiol Biochem Pharmacol* 2006; 156:61–90.
- Voets T. TRP channels and thermosensation. In: Nilius B, Flockerzi V, editors. *Mammalian transient receptor potential (TRP) cation channels*. Handbook of Experimental Pharmacology; 2014. pp 729–741.
- Ramsey IS, Delling M, Clapham DE. An introduction to TRP channels. *Annu Rev Physiol* 2006; 68:619–647.
- Matta JA, Ahern GP. Voltage is a partial activator of rat thermosensitive TRP channels. *J Physiol* 2007; 585:469–482.
- Voets T, Owsianik G, Janssens A, Talavera K, Nilius B. TRPM8 voltage sensor mutants reveal a mechanism for integrating thermal and chemical stimuli. *Nat Chem Biol* 2007; 3:174–182.
- Obukhov AG, Nowycky MC. TRPC5 channels undergo changes in gating properties during the activation-deactivation cycle. *J Cell Physiol* 2008; 216:162–171.
- Kalia J, Swartz KJ. Exploring structure-function relationships between TRP and Kv channels. *Sci Rep* 2013; 3:1523.
- Venkatachalam K, Montell C. TRP channels. *Annu Rev Biochem* 2007; 76:387–417.
- Long SB, Campbell EB, Mackinnon R. Voltage sensor of Kv1.2: structural basis of EM coupling. *Science* 2005; 309:903–908.
- Blunck R, Batulan Z. Mechanism of electromechanical coupling in voltage-gated potassium channels. *Front Pharmacol* 2012; 3:166.
- Labro AJ, Snyders DJ. Being flexible: the voltage-controllable activation gate of Kv channels. *Front Pharmacol* 2012; 3:168.
- Tamura K, Stecher G, Paterson D, Filipinski A, Kumar S. *MEGA6: molecular evolutionary genetics analysis version 6.0*. *Mol Biol Evol* 2013;30:2725–2729.
- Boukalova S, Marsakova L, Teisinger J, Vlachova V. Conserved residues within the putative S4–S5 region serve distinct functions among thermosensitive vanilloid transient receptor potential (TRPV) channels. *J Biol Chem* 2010; 285:41455–41462.

14. Beck A, Speicher T, Stoerger C, Sell T, Dettmer V, Jusoh SA, Abdulmughni A, Cavalié A, Philipp SE, Zhu MX, Helms V, Wissenbach U, Flockerzi V. Conserved gating elements in TRPC4 and TRPC5 channels. *J Biol Chem* 2013; 288:19471–19483.
15. Taberner FJ, López-Córdoba A, Fernández-Ballester G, Korchev Y, Ferrer-Montiel A. The region adjacent to the C-end of the inner gate in transient receptor potential melastatin 8 (TRPM8) channels plays a central role in allosteric channel activation. *J Biol Chem* 2014; 289:28579–28594.
16. Poblete H, Oyarzún I, Olivero P, Comer J, Zuñiga M, Sepulveda RV, Báez-Nieto D, González Leon C, González-Nilo F, Latorre R. Molecular determinants of phosphatidylinositol 4,5-bisphosphate (PIP₂) binding to transient receptor potential V1 (TRPV1) channels. *J Biol Chem* 2015;290:2086–2098.
17. Liao M, Cao E, Julius D, Cheng Y. Structure of the TRPV1 ion channel determined by electron cryo-microscopy. *Nature* 2013; 504:107–112.
18. Cao E, Liao M, Cheng Y, Julius D. TRPV1 structures in distinct conformations reveal activation mechanisms. *Nature* 2013; 504:113–118.
19. Zubcevic L, Herzik MA, Jr, Chung BC, Liu Z, Lander GC, Lee SY. Cryo-electron microscopy structure of the TRPV2 ion channel. *Nat Struct Mol Biol* 2016;23:180–186.
20. Huynh KW, Cohen MR, Jiang J, Samanta A, Lodowski DT, Zhou ZH, Moiseenkova-Bell VY. Structure of the full-length TRPV2 channel by cryo-EM. *Nat Commun* 2016; 7:11130.
21. Magidovich E, Yifrach O. Conserved gating hinge in ligand- and voltage-dependent K⁺ channels. *Biochemistry* 2004;43:13242–13247.
22. Ding S, Ingleby L, Ahern CA, Horn R. Investigating the putative glycine hinge in Shaker potassium channel. *J Gen Physiol* 2005; 126:213–226.
23. Balleza D, Carrillo E, Gómez-Lagunas F. Conservation analysis of residues in the S4-S5 linker and the terminal part of the S5-P-S6 pore modulus in Kv and HCN channels: flexible determinants for the electromechanical coupling. *Pflugers Arch* 2015; 467:2069–2079.
24. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 1994; 22:4673–4680.
25. Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: a sequence logo generator. *Genome Res* 2004; 14:1188–1190.
26. Gao Y, Cao E, Julius D, Cheng Y. TRPV1 structures in nanodiscs reveal mechanisms of ligand and lipid action. *Nature* 2016; 534:347–351.
27. Saotome K, Singh AK, Yelshanskaya MV, Sobolevsky AI. Crystal structure of the epithelial calcium channel TRPV6. *Nature* 2016; 534:506–511.
28. Zuckerkandl E, Pauling L. Evolutionary divergence and convergence in proteins. In: Bryson V, Vogel HJ, editors. *Evolving genes and proteins*. New York: Academic Press; 1965. pp 97–166.
29. Smith DK, Radivojac P, Obradovic Z, Zhu G. Improved amino acid flexibility parameters. *Protein Sci* 2003; 12:1060–1072.
30. Zhang H, Zhang T, Chen K, Shen S, Ruan J, Kurgan L. On the relation between residue flexibility and local solvent accessibility in proteins. *Proteins* 2009; 76:617–636.
31. Yang J, Yan R, Xu D, Poisson J, Zhang Y. The I-TASSER Suite: protein structure and function prediction. *Nat Methods* 2015; 12:7–18.
32. Wu S, Zhang Y. LOMETS: a local meta-threading-server for protein structure prediction. *Nucleic Acids Res* 2007; 35:3375–3382.
33. Case DA, Betz RM, Botello-Smith W, Cerutti DS, Cheatham TE, Darden TA, Duke RE, Giese TJ, Gohlke H, Goetz AW, Homeyer N, Izadi S, Janowski P, Kaus J, Kovalenko A, Lee TS, LeGrand S, Li P, Lin C, Luchko T, Luo R, Madej B, Mermelstein D, Merz KM, Monard G, Nguyen H, Nguyen HT, Omelyan I, Onufriev A, Roe DR, Roitberg A, Sagui C, Simmerling CL, Swails J, Walker RC, Wang J, Wolf RM, Wu X, Xiao L, York DM, Kollman PA. Amber, Version 16. San Francisco: University of California, 2016.
34. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK, a program to check the stereochemical quality of protein structures. *J App Cryst* 1993; 26:283–291.
35. Hooft RWW, Vriend G, Sander C, Abola EE. Errors in protein structures. *Nature* 1996; 381:272.
36. Latorre R, Brauchi S, Madrid R, Orío P. A cool channel in cold transduction. *Physiol* 2011; 26:273–285.
37. Voets T, Droogmans G, Wissenbach U, Janssens A, Flockerzi V, Nilius B. The principle of temperature-dependent gating in cold- and heat-sensitive TRP channels. *Nature* 2004; 430:748–754.
38. Zimmermann K, Lennerz JK, Hein A, Link AS, Kaczmarek JS, Delling M, Uysal S, Pfeifer JD, Riccio A, Clapham DE. Transient receptor potential cation channel, subfamily C, member 5 (TRPC5) is a cold-transducer in the peripheral nervous system. *Proc Natl Acad Sci USA* 2011; 108:18114–18119.
39. Yuan Z, Bailey TL, Teasdale RD. Prediction of protein B-factor profiles. *Proteins* 2005; 58:905–912.
40. Biedermannova L, E Riley K, Berka K, Hobza P, Vondrasek J. Another role of proline: stabilization interactions in proteins and protein complexes concerning proline and tryptophane. *Phys Chem Chem Phys* 2008; 10:6350–6359.
41. Tieleman DP, Shrivastava IH, Ulmschneider MR, Sansom MS. Proline-induced hinges in transmembrane helices: possible roles in ion channel gating. *Proteins* 2001; 44:63–72.
42. Gregorio-Teruel L, Valente P, González-Ros JM, Fernández-Ballester G, Ferrer-Montiel A. Mutation of I696 and W697 in the TRP box of vanilloid receptor subtype I modulates allosteric channel activation. *J Gen Physiol* 2014; 143:361–375.
43. Rosenhouse-Dantsker A, Logothetis DE. New roles for a key glycine and its neighboring residue in potassium channel gating. *Biophys J* 2006; 91:2860–2873.
44. Paulsen CE, Armache JP, Gao Y, Cheng Y, Julius D. Structure of the TRPA1 ion channel suggests regulatory mechanisms. *Nature* 2015; 520:511–517.
45. Baez D, Raddatz N, Ferreira G, Gonzalez C, Latorre R. Gating of thermally activated channels. *Curr Top Membr* 2014; 74:51–87.
46. Diaz-Franulic I, Poblete H, Miño-Galaz G, González C, Latorre R. Allosterism and structure in thermally activated transient receptor potential channels. *Annu Rev Biophys* 2016; 45:371–398.
47. García-Sanz N, Valente P, Gomis A, Fernández-Carvajal A, Fernández-Ballester G, Viana F, Belmonte C, Ferrer-Montiel A. A role of the transient receptor potential domain of vanilloid receptor I in channel gating. *J Neurosci* 2007; 27:11641–11650.
48. Valente P, García-Sanz N, Gomis A, Fernández-Carvajal A, Fernández-Ballester G, Viana F, Belmonte C, Ferrer-Montiel A. Identification of molecular determinants of channel gating in the transient receptor potential box of vanilloid receptor I. *FASEB J* 2008; 22:3298–3309.
49. Gregorio-Teruel L, Valente P, Liu B, Fernández-Ballester G, Qin F, Ferrer-Montiel A. The integrity of the TRP domain is pivotal for correct TRPV1 channel gating. *Biophys J* 2015; 109:529–541.
50. Carugo O, Argos P. Correlation between side chain mobility and conformation in protein structures. *Protein Eng* 1997; 10:777–787.
51. Jacob J, Duclohier H, Cafiso DS. The role of proline and glycine in determining the backbone flexibility of a channel-forming peptide. *Biophys J* 1999; 76:1367–1376.
52. Varela MF, Sansom CE, Griffith JK. Mutational analysis and molecular modeling of an amino acid sequence motif conserved in antiporters but not symporters in a transporter superfamily. *Mol Membr Biol* 1995; 12:313–319.
53. Karshikoff A, Nilsson L, Ladenstein R. Rigidity versus flexibility: the dilemma of understanding protein thermal stability. *FEBS J* 2015; 282:3899–3917.
54. Tominaga M. The role of TRP channels in thermosensation. In: Liedtke WB, Heller S, editors. *TRP ion channel function in sensory transduction and cellular signaling cascades*. Boca Raton (FL): CRC Press/Taylor & Francis; 2007. Chapter 20.

55. Nilius B, Watanabe H, Vriens J. The TRPV4 channel: structure-function relationship and promiscuous gating behaviour. *Pflugers Arch* 2003; 446:298–303.
56. Gu Y, Li DW, Bruschweiler R. Decoding the mobility and time scales of protein loops. *J Chem Theory Comput* 2015; 11:1308–1314.
57. Teng J, Loukin SH, Anishkin A, Kung C. L596-W733 bond between the start of the S4-S5 linker and the TRP box stabilizes the closed state of TRPV4 channel. *Proc Natl Acad Sci USA* 2015; 112:3386–3391.
58. Caterina MJ, Schumacher MA, Tominaga M, Rosen TA, Levine JD, Julius D. The capsaicin receptor: a heat-activated ion channel in the pain pathway. *Nature* 1997; 389:816–824.
59. Voets T, Talavera K, Owsianik G, Nilius B. Sensing with TRP channels. *Nat Chem Biol* 2005; 1:85–92.
60. Liu B, Hui K, Qin F. Thermodynamics of heat activation of single capsaicin ion channels VR1. *Biophys J* 2003; 85:2988–3006.
61. Yao J, Liu B, Qin F. Modular thermal sensors in temperature-gated transient receptor potential (TRP) channels. *Proc Natl Acad Sci USA* 2011; 108:11109–11114.
62. Vennekens R, Owsianik G, Nilius B. Vanilloid transient receptor potential cation channels: an overview. *Curr Pharm Des* 2008; 14: 18–31.
63. Beech DJ, Muraki K, Flemming R. Non-selective cationic channels of smooth muscle and the mammalian homologues of *Drosophila* TRP. *J Physiol* 2004; 559:685–706.
64. Muraki K, Iwata Y, Katanosaka Y, Ito T, Ohya S, Shigekawa M, Imaizumi Y. TRPV2 is a component of osmotically sensitive cation channels in murine aortic myocytes. *Circ Res* 2003; 93:829–838.
65. Bang S, Kim KY, Yoo S, Lee SH, Hwang SW. Transient receptor potential V2 expressed in sensory neurons is activated by probenecid. *Neurosci Lett* 2007; 425:120–125.
66. Shibasaki K. Physiological significance of TRPV2 as a mechanosensor, thermosensor and lipid sensor. *J Physiol Sci* 2016; 66:359–365.
67. de Groot T, Bindels RJ, Hoenderop JG. TRPV5: an ingeniously controlled calcium channel. *Kidney Int* 2008; 74:1241–1246.
68. Brauchi S, Orio P, Latorre R. Clues to understanding cold sensation: thermodynamics and electrophysiological analysis of the cold receptor TRPM8. *Proc Natl Acad Sci USA* 2004; 101:15494–15499.
69. McKemy DD, Neuhauser WM, Julius D. Identification of a cold receptor reveals a general role for TRP channels in thermosensation. *Nature* 2002; 416:52–58.
70. Yoshida T, Inoue R, Morii T, Takahashi N, Yamamoto S, Hara Y, Tominaga M, Shimizu S, Sato Y, Mori Y. Nitric oxide activates TRP channels by cysteine S-nitrosylation. *Nat Chem Biol* 2006; 2:596–607.
71. Zholos AV. TRPC5. *Handb Exp Pharmacol* 2014; 222:129–156.
72. Gomis A, Soriano S, Belmonte C, Viana F. Hypoosmotic- and pressure-induced membrane stretch activate TRPC5 channels. *J Physiol* 2008; 586:5633–5649.
73. Nobile M, Olcese R, Toro L, Stefani E. Fast inactivation of Shaker K⁺ channels is highly temperature dependent. *Exp Brain Res* 1997; 114:138–142.
74. Rodríguez BM, Sigg D, Bezanilla F. Voltage gating of Shaker K⁺ channels. The effect of temperature on ionic and gating currents. *J Gen Physiol* 1998; 112:223–242.
75. Abderemane-Ali F, Es-Salah-Lamoureux Z, Delemotte L, Kasimova MA, Labro AJ, Snyders DJ, Fedida D, Tarek M, Baró I, Loussouarn G. Dual effect of phosphatidylinositol (4,5)-bisphosphate PIP₂ on Shaker K⁺ channels. *J Biol Chem* 2012; 287:36158–36167.
76. Thomas DD. Saturation transfer EPR studies of microsecond rotational motions in biological membranes. In: Martonosi AN, editor. *The enzymes of biological membranes*. Vol. 1. Membrane structure and dynamics. Springer, Plenum Press, New York, 1985. pp 287–312.
77. Liang ZX, Lee T, Resing KA, Ahn NG, Klinman JP. Thermal-activated protein mobility and its correlation with catalysis in thermophilic alcohol dehydrogenase. *Proc Natl Acad Sci USA* 2004; 101:9556–9561.
78. Salazar H, Jara-Oseguera A, Hernandez-Garcia E, Llorente I, Arias-Olguin I, Soriano-Garcia M, Islas L, Rosenbaum T. Structural determinants of gating in the TRPV1 channel. *Nat Struct Mol Biol* 2009; 16:704–711.
79. Zheng J, Ma L. Structure and function of the thermoTRP channel pore. *Curr Top Membr* 2014; 74:233–257.
80. Susankova K, Ettrich R, Vyklicky L, Teisinger J, Vlachova V. Contribution of the putative inner-pore region to the gating of the transient receptor potential vanilloid subtype 1 channel (TRPV1). *J Neurosci* 2007; 27:7578–7585.
81. Denning EJ, Woolf TB. Cooperative nature of gating transitions in K⁽⁺⁾ channels as seen from dynamic importance sampling calculations. *Proteins* 2010; 78:1105–1119.
82. Islas LD, Salazar H, Jara-Oseguera A, Nieto-Posadas A, Llorente I, Rangel-Yescas G, Rosenbaum T. The helical character of the S6 segment of TRPV1 channels. *Channels* 2009; 3:311–313.
83. Bohlen CJ, Priel A, Zhou S, King D, Siemens J, Julius D. A bivalent tarantula toxin activates the capsaicin receptor, TRPV1, by targeting the outer pore domain. *Cell* 2010; 141:834–845.
84. Mahalakshmi R, Sengupta A, Raghothama S, Shamala N, Balam P. Tryptophan-containing peptide helices: interactions involving the indole side chain. *J Pept Res* 2005; 66:277–296.
85. Petrella RJ, Karplus M. The role of carbon-donor hydrogen bonds in stabilizing tryptophan conformations. *Proteins* 2004; 54:716–726.
86. Rose GD, Wolfenden R. Hydrogen bonding, hydrophobicity, packing, and protein folding. *Annu Rev Biophys Biomol Struct* 1993; 22: 381–415.
87. Burley SK, Petsko GA. Amino-aromatic interactions in proteins. *FEBS Lett* 1986; 203:139–143.
88. Scheiner S. Relative strengths of NH··O and CH··O hydrogen bonds between polypeptide chain segments. *J Phys Chem B* 2005; 109: 16132–16141.
89. Bhattacharyya R, Chakrabarti P. Stereospecific interactions of proline residues in protein structures and complexes. *J Mol Biol* 2003; 331:925–940.
90. Ridder AN, Morein S, Stam JG, Kuhn A, de Kruijff B, Killian JA. Analysis of the role of interfacial tryptophan residues in controlling the topology of membrane proteins. *Biochemistry* 2000; 39:6521–6528.
91. Johnston AJ, Zhang YR, Busch S, Pardo LC, Imberti S, McLain SE. Amphipathic solvation of indole: implications for the role of tryptophan in membrane proteins. *J Phys Chem B* 2015; 119:5979–5987.
92. Norman KE, Nymeyer H. Indole localization in lipid membranes revealed by molecular simulation. *Biophys J* 2006; 91:2046–2054.
93. Chugunov AO, Volynsky PE, Krylov NA, Nolde DE, Efremov RG. Temperature-sensitive gating of TRPV1 channel as probed by atomistic simulations of its trans- and juxtamembrane domains. *Sci Rep* 2016; 6:33112. doi:10.1038/srep33112.
94. Wen H, Qin F, Zheng W. Toward elucidating the heat activation mechanism of the TRPV1 channel gating by molecular dynamics simulation. *Proteins* 2016; 84:1938–1949. doi:10.1002/prot.25177.

Calorimetría Diferencial de Barrido

Sergio Romero-Romero¹, Miguel Costas²

¹Laboratorio de Físicoquímica e Ingeniería de Proteínas, Departamento de Bioquímica, Facultad de Medicina, Universidad Nacional Autónoma de México.

²Laboratorio de Biofísicoquímica, Departamento de Físicoquímica, Facultad de Química, Universidad Nacional Autónoma de México.

1. Introducción

El proceso por el cual la cadena polipeptídica de una proteína alcanza su estructura tridimensional biológicamente funcional -conocida como la estructura nativa- se denomina plegamiento, y es posiblemente el ejemplo más fundamental de autoensamblaje molecular en un contexto biológico. Un conocimiento a profundidad de este proceso es de gran importancia para (i) acercarse a alcanzar el objetivo de poder predecir la estructura proteica a partir de la secuencia de aminoácidos, (ii) poder realizar el diseño *de novo* de proteínas con propiedades nuevas y útiles, (iii) profundizar en muchos aspectos de las teorías evolutivas y (iv) aumentar el conocimiento de los orígenes moleculares de muchas enfermedades que son atribuidas al plegamiento anómalo de proteínas.¹ Con frecuencia, en los laboratorios de investigación se estudia el proceso inverso al plegamiento, conocido como desplegamiento o desnaturalización. Los procesos de plegamiento o de desplegamiento de una proteína conllevan la formación o ruptura, respectivamente, de numerosas interacciones débiles (no covalentes), y están caracterizados por un cambio en la energía libre de Gibbs de la proteína al pasar del estado inicial al final. Esta diferencia de energía entre el estado nativo y otras conformaciones "no nativas" o desnaturalizadas permite analizar y conocer la estabilidad de la proteína.²

Una estrategia experimental para conocer los cambios energéticos que ocurren durante el desplegamiento de una proteína consiste en emplear la técnica de Calorimetría Diferencial de Barrido (DSC en inglés).³ La Calorimetría Diferencial de Barrido (CDB) tiene una larga historia de interesantes y útiles aplicaciones en el campo del estudio de las proteínas. De hecho, puede afirmarse que los estudios con CDB en la desnaturalización

térmica de proteínas han jugado un papel central en el conocimiento de las bases moleculares que determinan la estabilidad proteica. La técnica fue desarrollada por E. S. Watson y M. J. O'Neill en 1962, y se introdujo en el mercado en 1963 en la Conferencia de Química y Espectroscopía Analítica y Aplicada de Pittsburgh.⁴ Inicialmente, el instrumento fue diseñado para medir los cambios conformacionales y la temperatura de transición vítrea inducidos por el aumento de la temperatura en diversos materiales y biopolímeros en estado sólido. Sin embargo, el primer calorímetro diferencial de barrido adiabático que pudo ser utilizado en bioquímica fue desarrollado por P. L. Privalov y D. R. Monaselidze en 1964.⁵

La CDB es actualmente aceptada como la técnica de elección para determinar la energética del plegamiento/desplegamiento de proteínas y los mecanismos termodinámicos que subyacen a estos procesos. En ella, se determina la capacidad calorífica de una proteína u otra macromolécula como función de la temperatura. Cuando dentro del intervalo de temperatura empleado en el experimento calorimétrico la proteína en estudio sufre la transición del estado nativo al desnaturalizado, la CDB permite determinar la cantidad de energía absorbida (o calor, si se mantiene la presión constante) para que ocurra dicha transición.⁶ Así, empleando los datos obtenidos en un experimento de CDB se puede realizar una caracterización termodinámica de la transición. De forma general, la información obtenida de un experimento de CDB permite conocer: ⁷ (i) la capacidad calorífica de la proteína a diferentes temperaturas, (ii) los cambios en la entalpía (ΔH), en la entropía (ΔS) y en la capacidad calorífica (ΔC_p) asociados con el desplegamiento inducido por temperatura, (iii) la temperatura media de desnaturalización (T_m), y (iv) la población de estados intermedios durante el desplegamiento y sus parámetros termodinámicos. En resumen, la CDB es una técnica poderosa que caracteriza los cambios conformacionales inducidos por temperatura en proteínas y otras macromoléculas biológicas. En este capítulo revisaremos los fundamentos de la técnica, la instrumentación y las aplicaciones de la CDB en el campo de proteínas, haciendo algunas observaciones al final sobre su uso en otros campos de interés biotecnológico y biomédico.

2. Instrumentación

Un calorímetro diferencial de barrido consiste en dos celdas idénticas: la celda de referencia que contiene el amortiguador utilizado para solubilizar la macromolécula, y la celda de trabajo o de muestra que contiene la macromolécula de interés en ese mismo amortiguador (para una detallada descripción del instrumento, ver referencias 8 y 9). Ambas celdas están colocadas dentro de una chaqueta adiabática que las aísla del exterior, evitando así la pérdida o ganancia de energía desde el entorno (figura 1). El calorímetro establece una diferencia de temperatura (ΔT) muy pequeña entre las dos celdas y de ahí el uso de la palabra diferencial en su nombre. En un experimento de CDB ambas celdas son simultáneamente calentadas a la misma velocidad (conocida como velocidad de barrido) y, como son idénticas, el ΔT es el mismo a toda temperatura hasta que ocurre algún evento que absorbe o libera energía. Si en la celda de trabajo hay una proteína, este evento es el desplegamiento de la molécula que, al ocurrir, altera el valor de ΔT . Los calentadores y circuitos de retroalimentación del calorímetro proporcionan a las celdas de trabajo o de referencia (según se haya absorbido o liberado energía) la cantidad de energía necesaria para restablecer el ΔT a su valor original. En el caso de una proteína, esa cantidad de energía es igual a la energía absorbida por ella al desplegarse, que es, como se describirá más adelante, una de los valores más importantes que proporciona la CDB.

La energía que suministra el instrumento a las celdas para restablecer el ΔT original se realiza aplicando una corriente eléctrica (I) a un voltaje constante (V). Es decir, el equipo aplica una potencia eléctrica (P) que es igual al producto de V e I ($P = VI$) y cuyas unidades son watts. Dado que $1 \text{ watt} = 1 \text{ Joule/s}$ (energía/tiempo), al dividir P entre la tasa de calentamiento empleada en el experimento (dT/dt , donde $T =$ temperatura y $t =$ tiempo) se obtiene que dicho cociente ($P/(dT/dt)$) tiene unidades de energía/grado.¹⁰ Estas son las unidades (J/K o cal/K) de la propiedad extensiva llamada capacidad calorífica. Al dividir entre el número de moles de la proteína presente en la celda, se obtiene la capacidad calorífica molar (con unidades de $J/K \text{ mol}$ o $\text{cal}/K \text{ mol}$) que es una propiedad intensiva.¹¹ Dado que el experimento se realiza a presión constante, el experimento de CDB proporciona el C_P del sistema (proteína en amortiguador, por ejemplo) que esté colocado en

la celda de trabajo. En resumen, un equipo de CDB mide el C_p como función de la temperatura.

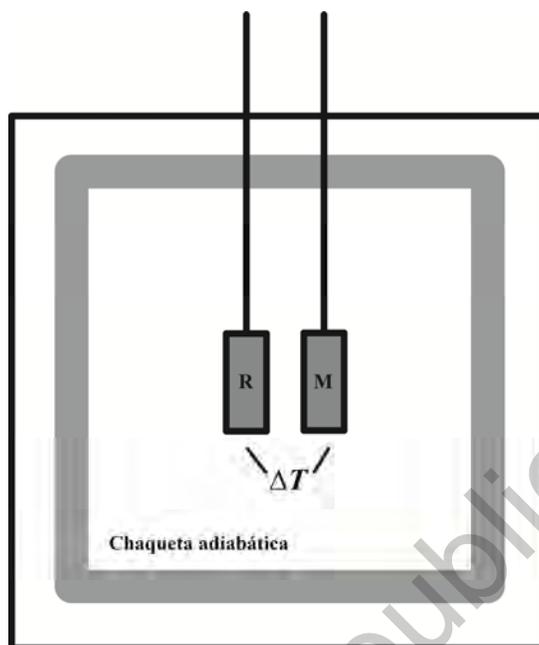


Figura 1. Ilustración esquemática de un Calorímetro Diferencial de Barrido (CDB). Consta de dos celdas idénticas, la que contiene la muestra M de interés (por ejemplo, una solución de proteína), y la de referencia R que se llena con el mismo disolvente (por ejemplo, una solución amortiguadora) en el que está la proteína. ΔT es la diferencia de temperatura entre las dos celdas.

Desde el primer calorímetro descrito en 1964, los CDB han evolucionado muy rápidamente. Conservando el principio de medición descrito anteriormente, en los más de 50 años que han transcurrido desde su introducción, muchas mejoras han sido implementadas, producto del desarrollo de la tecnología. Una de las primeras fue lograr construir chaquetas adiabáticas mucho más eficientes. El siguiente avance clave fue la miniaturización de las celdas, lo cual mejoró la sensibilidad, drásticamente reduciendo los gradientes de temperatura que ocurrían en muestras grandes. Además, este cambio en el diseño también permitió la eliminación de mecanismos de agitación. El reemplazo de celdas cilíndricas por tubos capilares, que tienen un muy alto cociente superficie/volumen, minimizó dramáticamente los efectos de viscosidad y los gradientes de calor. También, el uso de celdas presurizadas permitió el aumentar el límite de temperatura superior de los experimentos.⁹ Los equipos modernos son capaces de calentar desde velocidades muy bajas

(0.2 grados cada minuto) hasta muy rápidas (4.2 grados cada minuto), y operan aplicando una presión ligera (entre 50 y 60 psi) a ambas celdas, aumentando con ello el punto de ebullición de las disoluciones que se encuentran contenidas en éstas (el contenido de las celdas es básicamente agua) y permitiendo llegar incluso hasta los 125 °C de calentamiento. Esta presurización también tiene el efecto de inhibir la formación de burbujas durante el calentamiento de los gases disueltos en la muestra.

El éxito de un experimento de CDB depende no sólo de la calidad del calorímetro sino también del empleo de un diseño experimental adecuado. Al respecto, a continuación se proporcionan un conjunto de recomendaciones y cuidados específicos útiles, para el caso en el que el sistema de interés sea una proteína:

- Un experimento de CDB consiste de: (i) un barrido de amortiguador-amortiguador (presente en ambas celdas) para obtener la línea base física, (ii) un barrido de proteína-amortiguador, colocando la proteína disuelta en el amortiguador en la celda de trabajo y amortiguador en la de referencia, para obtener una primera endoterma de desnaturalización y (iii) un segundo barrido a esa misma muestra después de haber sido enfriada, con el fin de saber si la transición nativa a desplegada es o no reversible. Típicamente, un proceso es considerado calorimétricamente reversible si el 85-90% del área bajo la curva de la endoterma es recuperada en el experimento de recalentamiento.⁷
- Para garantizar la calidad de los datos obtenidos se requiere: (i) que la pureza e integridad de la proteína bajo estudio sea de la más alta calidad posible; (ii) que las muestras de proteína sean dializadas contra el buffer a emplear, centrifugadas para remover agregados, y desgasificadas; (iii) que el amortiguador desgasificado, usado en el último paso de diálisis, sea el introducido en la celda de referencia con el fin de que los disolventes en ambas celdas sean idénticos; esto es especialmente crítico si se desea obtener los valores absolutos de las capacidades caloríficas; (iv) que la concentración de la proteína sea determinada con alta exactitud dado que cualquier error en la concentración se verá reflejado directamente en un error en los parámetros termodinámicos a obtenerse; esta cuantificación debe hacerse justo antes de que la muestra sea introducida en la celda del calorímetro; (v) que al cargar las

muestras en ambas celdas se cuide que no se generen burbujas; (vi) que después de los experimentos, se laven las celdas a profundidad para eliminar rastros de las corridas previas; además, periódicamente hay que realizar controles para asegurarse de que las celdas calorimétricas están limpias.

- El uso de un amortiguador adecuado es de importancia. En la elección del amortiguador se recomienda utilizar sales con bajas entalpías de ionización, ya que sales con alta entalpía de ionización provocarán grandes cambios en el pH a altas temperaturas y, por lo tanto, efectos en la estabilidad proteica. Sales como fosfatos, formatos y acetatos regularmente son buenos sistemas de amortiguamiento para los experimentos realizados (ΔH_{ion} cercana a 0). El empleo de ditioneitol (DTT) en la disolución amortiguadora no es recomendado debido a su inestabilidad y facilidad para oxidarse en presencia de aire (aun con los amortiguadores previamente desgasificados); como alternativa se puede utilizar tris(2-carboxietil)fosfina (TCEP) o β -mercaptoetanol (BME).¹² Algunos aditivos como detergentes o glicerol pueden causar problemas en el momento de dializar; se puede prolongar el tiempo de diálisis o utilizar otro método de diálisis como una columna de desalado.

3. La traza calorimétrica

A la representación de la capacidad calorífica contra la temperatura se le conoce como traza calorimétrica. Una traza típica de CDB para una proteína globular en solución es mostrada en la figura 2. Dado que la estructura nativa de cada proteína es distinta, la traza calorimétrica de cada una de ellas es diferente. Por ello, la traza calorimétrica de una proteína puede considerarse como su huella digital. Incluso, para una misma proteína, la traza calorimétrica es diferente de acuerdo a las condiciones experimentales en las que se encuentra la macromolécula (es decir, pH, presencia de solutos, amortiguador, etc.).

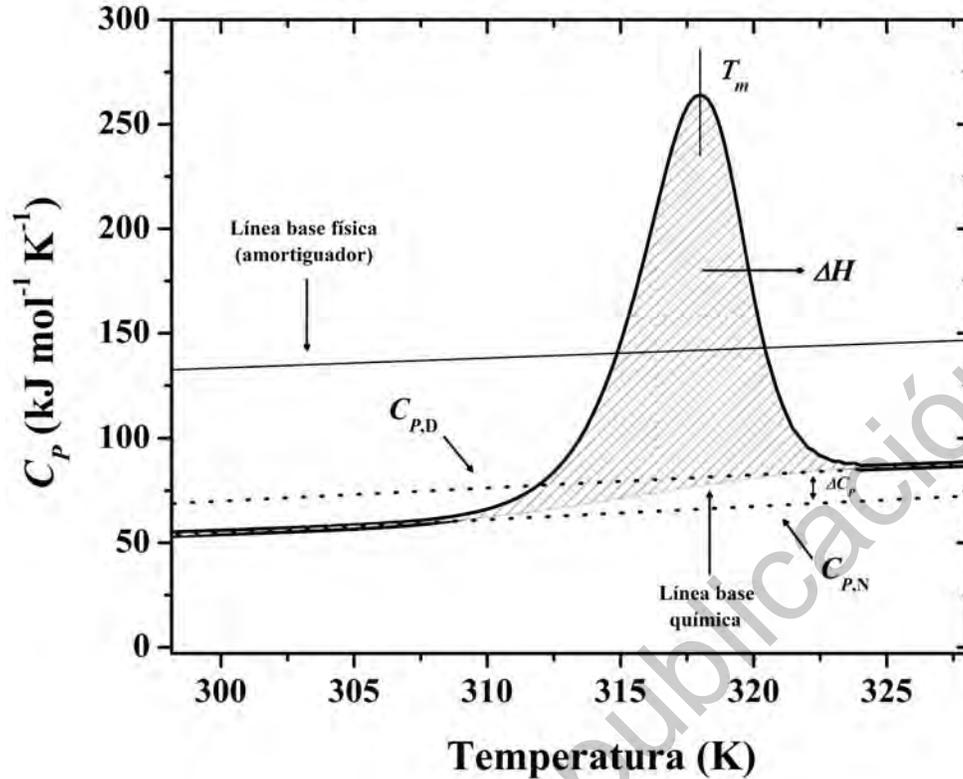


Figura 2. Traza calorimétrica típica para una proteína globular en solución. La capacidad calorífica molar del estado nativo ($C_{P,N}$) y desnaturalizado ($C_{P,D}$) está indicada por líneas punteadas. La diferencia entre estos valores define el cambio en la capacidad calorífica para la transición ($\Delta C_P = C_{P,D} - C_{P,N}$). El cambio en la entalpía (ΔH) está dado por el área bajo la curva de la endoterma delimitada por la traza misma y la línea base química que conecta a los estados nativos y desnaturalizados. La temperatura a la cual ocurre el máximo del pico de la endoterma se denota T_m . La capacidad calorífica del amortiguador (línea base física) se obtiene en un experimento donde ambas celdas están llenas sólo con amortiguador.

En la figura 2, la traza calorimétrica muestra un pico o transición. Esta endoterma es debida a la cantidad de energía absorbida al ocurrir la transición de la proteína del estado nativo al desplegado o desnaturalizado (para proteínas complejas, más transiciones sobrepuestas son a veces observadas). Los valores de C_P en la pre-transición y post-transición son las capacidades caloríficas de los estados nativo y desplegado de la proteína, respectivamente ($C_{P,N}$ y $C_{P,D}$, figura 2). La diferencia entre estos valores define el cambio en la capacidad calorífica (ΔC_P) para la transición.¹³ En la figura 2 también se muestra la línea base química que se obtiene al conectar los estados nativo y desnaturalizado; esta línea delimita el área bajo la curva que representa el cambio en la entalpía (ΔH) asociado a

la desnaturalización proteica. La temperatura a la cual el C_p es un máximo se denota como T_m . Si la transición es de dos estados (es decir, sin intermediarios estables) a esta temperatura la concentración de proteínas en el estado nativo es igual a la concentración de proteínas desplegadas. Por esta razón la T_m se conoce como la temperatura media de desnaturalización.^{14, 15} Es importante notar que la transición de nativo a desplegado ocurre en un rango de temperaturas. La magnitud de este rango está ligado a la cooperatividad de la transición: entre más pequeño sea este intervalo (un pico estrecho) la cooperatividad es mayor. Por último, la figura 2 también muestra el resultado del experimento donde ambas celdas contenían solo la disolución de amortiguador. Esta traza “amortiguador-amortiguador” es conocida como línea base física o instrumental que es comúnmente restada del perfil proteína-amortiguador (la endoterma), con el fin de corregir distorsiones instrumentales. Hay que notar que, en la mayoría de los rangos de temperatura utilizados, esta resta da valores negativos. Esto indica que la capacidad calorífica de la referencia es más alta que la de la muestra, lo cual es una consecuencia del bajo contenido de agua en la macromolécula (la proteína desplaza algunas moléculas de agua y la capacidad calorífica del agua líquida es más grande que la de la proteína).^{16, 17}

4. Aplicación de la Calorimetría Diferencial de Barrido al estudio de proteínas

4.1 Análisis de datos al equilibrio y cinéticos

Posiblemente una de las principales aplicaciones de la CDB es el análisis del desplegamiento de macromoléculas biológicas. Uno de los puntos importantes a considerar es si los valores de la capacidad calorífica en la figura 2 reflejan un proceso que está al equilibrio o no, es decir, si el desplegamiento está bajo un control termodinámico o bajo control cinético. Por “proceso al equilibrio” se entiende al equilibrio termodinámico entre los estados que visita la proteína en su ruta de plegamiento/desplegamiento (nativo, desplegado, parcialmente desplegado, etc.) y que es establecido a través de un barrido de temperatura.¹⁸ En otras palabras, las concentraciones de los estados poblados a cualquier temperatura están determinadas sólo por los valores de las constantes de equilibrio. Claro

está que las constantes de equilibrio cambian con la temperatura de acuerdo a relaciones termodinámicas simples, por lo que el movimiento de las poblaciones resulta en transiciones observadas en el termograma de CDB. Por ejemplo, para un proceso de dos estados al equilibrio con un valor de entalpía de desnaturalización suficientemente alta, la constante de equilibrio de desnaturalización cambia de un valor bajo que favorece el estado nativo a un valor alto que favorece el estado desnaturalizado en un rango de temperatura relativamente estrecho.¹⁹ Los procesos que se producen a través de una serie de estados al equilibrio se llaman reversibles. La reversibilidad de la transición se comprueba experimentalmente, como se explicó antes, realizando un segundo barrido (figura 3A). Cuando la transición del estado nativo al desplegado es reversible, es posible calcular todos los parámetros termodinámicos asociados a ella: T_m , ΔH y ΔC_p , además de parámetros derivados como ΔS y ΔG , que se obtienen mediante relaciones termodinámicas básicas.²⁰ Esta detallada descripción energética puede ser interpretada en términos estructurales usando correlaciones energético-estructurales que se han desarrollado con los años. Avances más recientes en los análisis al equilibrio incluyen los relacionados a la caracterización de los efectos por un ligando o el plegamiento de proteínas sin barreras energéticas,²¹ cuestiones que serán comentadas más adelante.

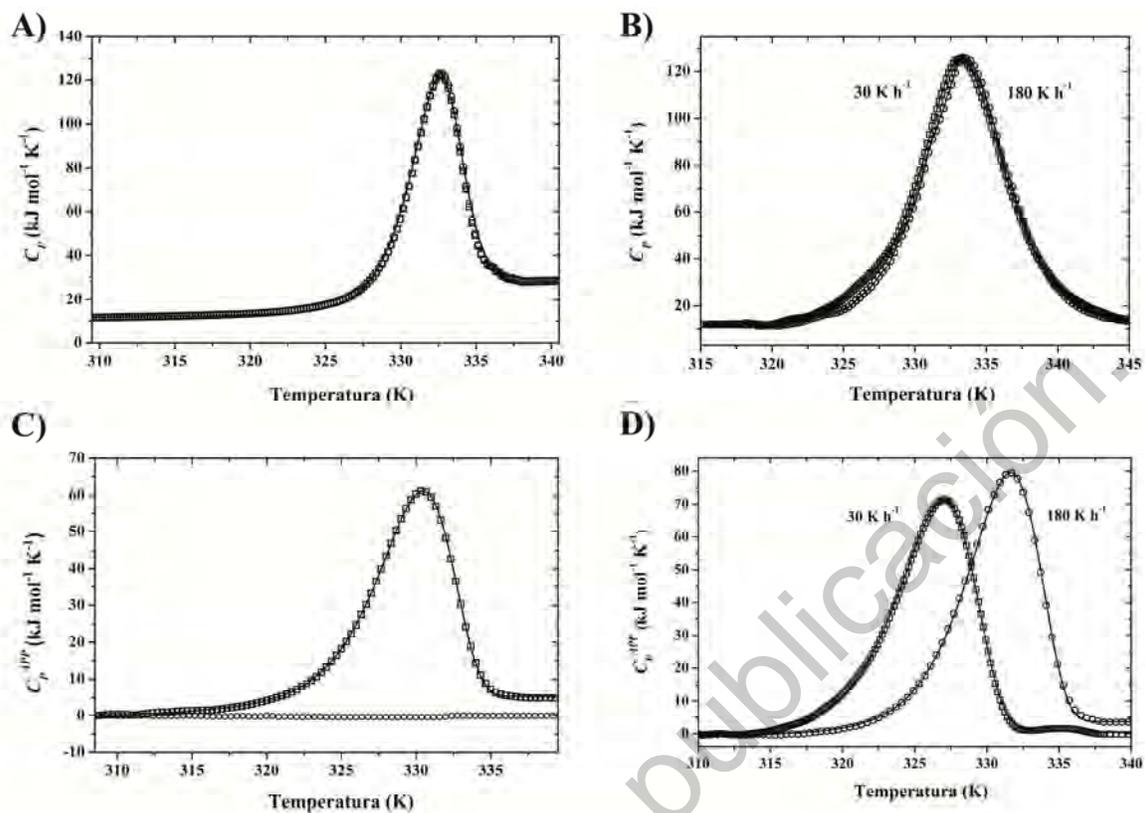


Figura 3. Ejemplos de trazas calorimétricas para distintas proteínas. **A)** Experimento con datos de desplegamiento al equilibrio. La segunda endoterma (círculos) es obtenida al volver a calentar la muestra tras la realización de un primer barrido (cuadros); ambas endotermas se superponen, es decir, son idénticas, indicando que el desplegamiento es un proceso reversible. El ajuste de la traza calorimétrica es empleando un modelo de dos estados reversible ($N \rightleftharpoons D$). **B)** Dependencia de la traza calorimétrica con la velocidad de barrido para una proteína que se despliega bajo un control termodinámico; nótese que la variación de T_m entre las endotermas ensayadas a dos velocidades de barrido extremas es de $< 1^\circ\text{C}$. **C)** Experimento con datos de desplegamiento bajo control cinético (cuadros). La segunda traza (círculos) no muestra ninguna transición, indicando un proceso irreversible. El ajuste de la endoterma es a un modelo de dos estados irreversible ($N \rightarrow F$). **D)** Dependencia de la traza calorimétrica con la velocidad de barrido para una proteína cuyo desplegamiento está controlado cinéticamente; nótese que la variación de T_m entre las trazas realizadas a dos velocidades de barrido extremas es $> 6^\circ\text{C}$. En todos los casos, los símbolos vacíos representan los datos experimentales y las líneas continuas los mejores ajustes a dichos datos. Las trazas mostradas corresponden a la proteína Triosafosfato Isomerasa.²²

En el caso de que experimentalmente se observe una única transición y que esta sea reversible, en muchos casos los datos se pueden ajustar a un modelo de dos estados reversible $N_n \rightleftharpoons nD$:

$$C_p(T) = B_0 + B_1T + f(T)\Delta C_p + \frac{\Delta H(T)}{RT_m^2} \left[\frac{1 - f(T)}{1 - n + \frac{n}{f(T)}} \right] \quad (\text{Ecuación 1})$$

donde n es el número de subunidades presentes en la macromolécula (1 en el caso de proteínas monoméricas, 2 para diméricas y así sucesivamente), B_0 y B_1 son constantes que describen las pre y post transiciones y $f(T)$ es la fracción de proteína en estado monomérico. El ajuste de los datos empleando la ecuación 1 proporciona los parámetros T_m , ΔH y ΔC_p .

Cuando el proceso del estado nativo al desplegado no está al equilibrio (en el sentido señalado arriba) sino que es dependiente del tiempo (es decir, cinético), se dice que el proceso está controlado cinéticamente y, por tanto, el análisis de la traza calorimétrica debe realizarse de manera diferente a la empleada cuando los datos de CDB reflejan un proceso al equilibrio. En una transición que es controlada cinéticamente las concentraciones de los estados involucrados en ella están determinadas por ecuaciones de velocidad (y no por constantes de equilibrio dependientes de temperatura).^{23, 24} Las dos fuentes más comunes de control cinético en proteínas son el equilibrio lento y la presencia de alteraciones irreversibles de las proteínas (agregación, autoproteólisis, etc.). Estas alteraciones irreversibles llevan a la proteína a un estado final que es incapaz de replegarse de regreso al estado nativo. En los experimentos de CDB, la presencia de alteraciones irreversibles es fácilmente detectable por la ausencia de una transición o traza calorimétrica en un barrido de recalentamiento (es decir, un segundo barrido después de calentar y enfriar la proteína por primera vez; figura 3C). La caracterización detallada del proceso controlado cinéticamente puede llevarse a cabo con experimentos CDB realizados a diferentes velocidades de barrido. Dado que el estado del sistema a una temperatura dada dependerá del tiempo requerido para alcanzar dicha temperatura, las trazas calorimétricas deben mostrar una dependencia importante con la velocidad de barrido: las trazas se desplazan a mayores temperaturas al aumentar la velocidad de calentamiento (figura 3D). Esto contrasta con el caso de un sistema controlado termodinámicamente, donde la dependencia de las transiciones con la velocidad de barrido es mínima (figura 3B).

En muchos casos, las transiciones irreversibles controladas cinéticamente obedecen a un modelo de dos estados irreversible ($N \rightarrow F$). Este modelo asume que sólo los estados nativo y desnaturalizado están significativamente poblados durante el experimento de CDB y proporciona la siguiente expresión para la capacidad calorífica (ecuación 2-4):

$$C_p = C_{p,N} + (C_{p,D} - C_{p,N}) * (1 - X_N) - \Delta H * \left(\frac{dX_N}{dT} \right) \quad (\text{Ecuación 2})$$

con:

$$X_N = \exp \left[- \exp \left(\frac{E_A * \Delta T}{R * T_m^2} \right) \right] \quad (\text{Ecuación 3})$$

$$\frac{dX_N}{dT} = - \frac{E_A}{R * T_m^2} * \exp \left(\frac{E_A * \Delta T}{R * T_m^2} \right) * \exp \left[- \exp \left(\frac{E_A * \Delta T}{R * T_m^2} \right) \right] \quad (\text{Ecuación 4})$$

donde X_N es la fracción mol del estado nativo, ΔH es la entalpía para la desnaturalización (para la transición de N a F), $C_{p,N}$ y $C_{p,D}$ son las capacidades caloríficas de los estados nativo y desnaturalizado (ver figura 2) que normalmente se toman como funciones lineales de la temperatura, $\Delta T = T - T_m$, y E_A es la energía de activación que separa al estado nativo del estado de transición. En la obtención de esta ecuación para C_p se ha supuesto que la conversión irreversible del estado nativo al estado final queda descrita por una ecuación de velocidad de primer orden,^{23, 25} cuya dependencia con la temperatura esta dado por la ecuación de Arrhenius (ecuación 5):²⁶

$$k = \exp \left[- \frac{E_A}{R} \left(\frac{1}{T} - \frac{1}{T^*} \right) \right] \quad (\text{Ecuación 5})$$

donde T^* es la temperatura a la cual $k = 1 \text{ min}^{-1}$. Al hacer el gráfico de $\ln k$ vs. $1/T$ (conocido como gráfico de Arrhenius) para todas las velocidades de barrido ensayadas para una proteína y ajustar dichos datos a una recta, la pendiente proporciona la E_A (figura 4).

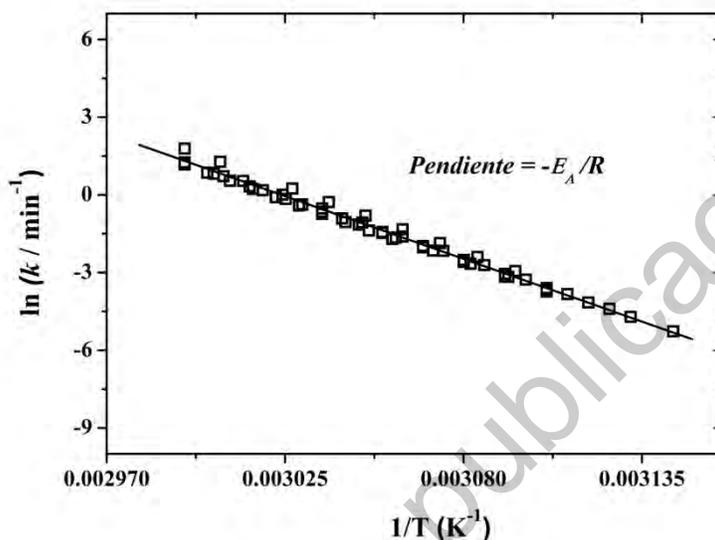


Figura 4. Gráfico de Arrhenius para el desplegamiento térmico irreversible de una proteína globular en solución. Los datos experimentales son derivados de seis experimentos a velocidades de barrido diferentes (30, 60, 90, 120, 150 y 180 K/h). La línea es el mejor ajuste a la ecuación de Arrhenius para la proteína Triosafosfato Isomerasa.²²

Se ha observado que, en general, las proteínas pequeñas se desnaturalizan reversiblemente mientras que las proteínas de mayor tamaño u oligoméricas lo hacen irreversiblemente o son parcialmente reversibles, además de que se agregan o se precipitan antes de que la condición desnaturalizante pueda ser modificada a su estado original.²⁷

4.2 Estabilidad cinética y estabilidad térmica

Uno de los objetivos más importantes de realizar experimentos de CDB para proteínas es poder determinar la estabilidad de estas macromoléculas. Existen dos tipos de estabilidad: la estabilidad cinética y la estabilidad térmica o termodinámica.^{28, 29} A continuación se discuten ambas.

Cuando la proteína en estudio tiene un desplegamiento que está bajo control cinético, evaluar su estabilidad -conocida como estabilidad cinética- consiste en conocer el tiempo durante el cual la proteína permanecerá en el estado nativo a una cierta temperatura T' . Este tiempo puede estimarse a partir del valor de la constante k a T' empleando el gráfico de Arrhenius obtenido de los experimentos de CDB (figura 4). Cuando la temperatura de interés T' (por ejemplo temperatura ambiente o fisiológica) está fuera del intervalo donde se construyó el gráfico de Arrhenius (que corresponde al intervalo en el que aparece la endoterma en la traza calorimétrica) la evaluación del tiempo de estabilidad implica realizar una extrapolación, de tal manera que a esa T' la estabilidad cinética de la proteína es sólo una estimación (entre más cerca esté T' del intervalo, más confiable será el valor obtenido).^{30, 31}

La termodinámica proporciona un criterio claro para determinar si un proceso es o no espontáneo: si el cambio de energía libre de Gibbs es negativo ($\Delta G < 0$), el proceso ocurre de manera espontánea. Entonces, para el proceso de desplegamiento de una proteína ($N \rightarrow D$) cuando $\Delta G > 0$ el estado estable es el nativo. Dado que ΔG depende de la temperatura, la estabilidad de una proteína será distinta a diferentes temperaturas: en aquellas temperaturas en las que la desnaturalización de la proteína tenga un $\Delta G > 0$, la proteína será estable (permanecerá en su estado nativo), mientras que fuera de ese intervalo el estado estable será el desnaturalizado (figura 5). A la temperatura a la cual el $\Delta G > 0$ es máximo se le conoce como la temperatura de máxima estabilidad térmica o termodinámica de la proteína. A la representación de ΔG vs. T mostrada en la figura 5 se le conoce como la curva de estabilidad. Cada proteína tiene su propia curva de estabilidad y, al igual que la traza calorimétrica, esta es diferente de acuerdo a las condiciones experimentales usadas en la determinación de los parámetros termodinámicos (pH, solutos, amortiguador, etc.) La curva de estabilidad muestra que la T_m a alta temperatura no es única, sino que además existe un valor de T_m a bajas temperaturas (T_m').^{32, 33} A estas dos temperaturas la proteína se desnaturaliza y, por ello, se denominan coloquialmente las temperaturas a las cuales ocurre la desnaturalización en caliente (T_m) y desnaturalización en frío (T_m'), respectivamente. Para muchas proteínas es imposible medir la segunda, debido a que esta T_m' se encuentra por debajo de 273.15 K y, por lo tanto, la muestra se congela antes de alcanzarla.

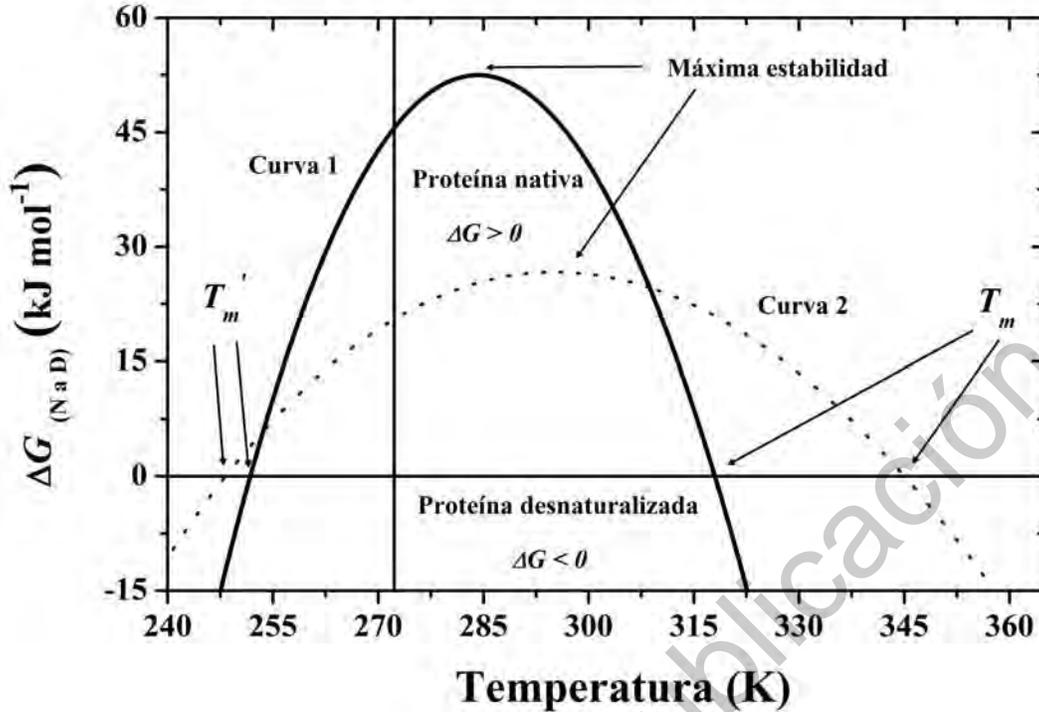


Figura 5. Curvas de estabilidad para dos proteínas. Esta curva se obtiene empleando los valores de ΔH , ΔC_p y T_m que se obtienen de los experimentos de CDB. En la zona del diagrama donde $\Delta G (N \rightarrow D) > 0$ el estado nativo es estable. Fuera de esa zona, donde el $\Delta G (N \rightarrow D) < 0$, el estado estable es el desnaturalizado. T_m y T_m' indican las temperaturas medias de desnaturalización en caliente y en frío, respectivamente. La línea vertical indica la temperatura a la cual la muestra se congela.

La evaluación del ΔG requiere que el proceso de desplegamiento esté en equilibrio y sea reversible (ver sección 4.1). La expresión de la termodinámica (ecuación 6) que permite el cálculo de la curva de estabilidad (figura 5) es:

$$\Delta G(T) = \Delta H \left(1 - \frac{T}{T_m} \right) - \Delta C_p \left[T_m - T + T \ln \left(\frac{T}{T_m} \right) \right] \quad (\text{Ecuación 6})$$

donde los valores de ΔH , ΔC_p y T_m se obtienen de los experimentos de CDB. Para obtener esta ecuación³¹ se ha empleado como temperatura de referencia a la T_m , es decir $\Delta G = 0$ a la T_m . Es interesante hacer notar que la curva de estabilidad sería una recta si el ΔH y el ΔS para el proceso fueran independientes de la temperatura,³⁴ es decir, si fuesen constantes. Sin embargo, hace más de cuatro décadas, se encontró que el proceso de

desplegamiento (y por tanto su inverso, el plegamiento) está asociado con cambios importantes en la capacidad calorífica, esto es, que el ΔH y el ΔS no son independientes de la temperatura. Además, con el descubrimiento del efecto hidrofóbico y su singular dependencia de la temperatura,^{35, 36} quedó claro que la simplificación de que $\Delta C_P = 0$ era una representación muy pobre de lo que en realidad sucedía en el desplegamiento de las proteínas. Por otro lado, fue el desarrollo de la CDB lo que permitió la evaluación experimental del ΔC_P . Con diversos estudios pioneros en el campo de las proteínas, se encontró que el ΔC_P de desplegamiento de una proteína es grande y positivo, y que, dentro del error experimental, éste puede tomarse como una constante para una proteína determinada, es decir es independiente de la temperatura.³⁷⁻³⁹ Debido a la ausencia de reversibilidad en el proceso de desplegamiento térmico de muchas proteínas, las curvas de estabilidad se han obtenido para relativamente pocas proteínas, principalmente para proteínas monoméricas pequeñas.⁴⁰ Es también importante mencionar que es común en la literatura encontrar la aseveración que establece que una proteína A es más estable que otra proteína B cuando su T_m es mayor ($T_{mA} > T_{mB}$). En vista de la figura 5, este tipo de aseveraciones no son necesariamente correctas (la proteína A puede tener un ΔG menor aunque su T_m sea mayor) por no ser suficientemente claras o precisas. El lenguaje correcto debiera ser que la proteína A es más termo-resistente que la proteína B (compárense las curvas 2 y 1 en la figura 5).

Entre las propiedades termodinámicas, la capacidad calorífica a presión constante goza de una situación privilegiada. Esto es debido a la relación que existe entre el ΔC_P y la entropía ΔS y la energía libre de Gibbs ΔG : $\Delta C_P = T(d\Delta S/dT)_P = -T(d^2\Delta G/dT^2)_P$. La relación con ΔG indica que el ΔC_P es el responsable de la curvatura en la curva de estabilidad (figura 5). Considerando la interpretación microscópica de la entropía, la relación con el ΔS indica que el ΔC_P es un reflejo del nivel de organización o estructura en el sistema. Más precisamente, el ΔC_P es una medida de la variación con temperatura de esa organización o estructura. Así, la relación entre ΔC_P y ΔS permite establecer -de manera cualitativa- un vínculo entre las propiedades termodinámicas macroscópicas determinadas experimentalmente y las características o cambios a nivel microscópico que ocurren durante el proceso, es decir explicar o racionalizar el proceso de desplegamiento (o del

plegamiento) en términos microscópicos. Por ejemplo, el $\Delta C_P > 0$ para la desnaturalización es resultado de un $\Delta S > 0$, es decir de que la entropía del estado final (desnaturalizado) es mayor que la del estado inicial (nativo). A nivel microscópico, el $\Delta C_P > 0$ refleja la exposición al agua (hidratación) de los grupos hidrofóbicos ocultos en el estado nativo de la proteína que quedan expuestos al disolvente durante el desplegamiento.³⁴

Los valores de ΔC_P también contienen información relevante referente a los estados estructurales de la proteína. Por ejemplo, la capacidad calorífica de una proteína en su estado desplegado ($C_{P,D}$) puede ser calculado a partir de la secuencia de aminoácidos (ecuación 7): si todos los aminoácidos están expuestos al agua, entonces el $C_{P,D}$ puede aproximarse a ser la suma de las contribuciones individuales de los aminoácidos constituyentes (cadenas laterales más cadena principal):⁴¹

$$C_{P,D} = \left(\sum_{i=1}^{20} n_i * C_{P,i} \right) + (N_{AA} - 1) * C_{P,bb} + C_{P,NH_2} + C_{P,COOH} \quad (\text{Ecuación 7})$$

donde la sumatoria refiere a los 20 aminoácidos, n_i es el número de aminoácidos de cada tipo y $C_{P,i}$ es la capacidad calorífica molar de esa cadena lateral; N_{AA} es el número de aminoácidos totales en la proteína, $C_{P,bb}$ es la capacidad calorífica de la unidad básica (esqueleto) de la cadena principal (-CHCONH-), C_{P,NH_2} la capacidad calorífica del amino terminal y $C_{P,COOH}$ la capacidad calorífica del carboxilo terminal.

Por otro lado, la capacidad calorífica del estado nativo ($C_{P,N}$) es una función lineal de la temperatura (ver figura 2) dentro del rango en el cual este puede ser medida. Dentro de ese rango, la capacidad calorífica de los estados nativos es muy similar para todas las proteínas cuando las trazas son normalizadas por el peso molecular. Se ha descrito que el ΔC_P entre los estados nativo y desplegado es directamente proporcional al cambio de las superficies polares y no polares accesibles al disolvente (agua) entre esos dos estados (ecuación 8):^{42, 43}

$$\Delta C_P = \Delta C_{P,ap}^\circ * \Delta A_{ap} + \Delta C_{P,pol}^\circ * \Delta A_{pol} \quad (\text{Ecuación 8})$$

donde $\Delta C_{P,pol}^\circ$ y $\Delta C_{P,ap}^\circ$ son las contribuciones de los residuos polares y no polares a la capacidad calorífica, cuyos valores han sido obtenidos utilizando resultados de ΔC_P

experimentales ($\Delta C_{P,ap}^{\circ} = 1.88 \pm 0.08 \text{ J/K/[mol}\cdot\text{\AA}^2]$ y $\Delta C_{P,pol}^{\circ} = -1.09 \pm 0.13 \text{ J/K/[mol}\cdot\text{\AA}^2]$), y ΔA_{ap} y ΔA_{pol} son las diferencias entre las superficies polares y no polares accesibles al solvente de los dos estados calculadas a partir de su estructura.

4.3 Criterio calorimétrico y procesos de dos estados

El modelo de equilibrio de dos estados ha sido ampliamente empleado para estudiar el plegamiento/desplegamiento de proteínas y evaluar su estabilidad. Este modelo asume que el desplegamiento es reversible y que sólo dos estados de la proteína son significativamente poblados en el equilibrio: nativo y desplegado. El estado desplegado es considerado un macroestado o ensamble de conformaciones al equilibrio, todas las cuales son capaces de volver al estado nativo. Así, frente a datos experimentales del desplegamiento de una proteína, la primera interrogante a contestar es si el proceso queda satisfactoriamente descrito por un mecanismo o proceso de sólo esos dos estados. Una parte de la evidencia para contestar esta pregunta es la concordancia de las transiciones experimentales obtenidas por diferentes técnicas biofísicas (Dicroísmo Circular cercano y lejano, Fluorescencia, entre otras). Sin embargo, mientras que una clara discrepancia entre dichas técnicas efectivamente descarta el comportamiento de dos estados, la coincidencia de las mismas no constituye una evidencia definitiva, ya que siempre existe la posibilidad de que en la traza calorimétrica estén “escondidas” otras transiciones (entre otros estados significativamente poblados).^{44, 45} En este contexto, los experimentos de CDB proveen de un criterio confiable para saber si el modelo de dos estados describe adecuadamente el proceso: el llamado “criterio calorimétrico”, el cual se describe a continuación.

La entalpía de desnaturalización (la diferencia de energía entre los estados nativo y desnaturalizado) puede evaluarse de dos formas diferentes en un experimento de CDB.²¹ Por una parte, puede ser obtenida calculando el área bajo la endoterma como se explicó en la sección 3; a esta entalpía se le conoce como entalpía calorimétrica (ΔH_{cal}). Por otra parte, puede ser obtenida a partir del conocimiento de que la entalpía de desplegamiento determina la dependencia de la constante de equilibrio con respecto a la temperatura a través de la ecuación de van't Hoff (ecuación 9):⁴⁶

$$\frac{d \ln K_D}{dT} = \frac{\Delta H_{vH}}{RT^2} \quad (\text{Ecuación 9})$$

donde K es la constante de equilibrio para el proceso ($K = [D]/[N]$) y ΔH_{vH} es conocida como entalpía de van't Hoff. Esta ecuación indica que un valor grande de la entalpía implica que la constante de equilibrio de desplegamiento es fuertemente dependiente de la temperatura, produciendo una traza calorimétrica estrecha; contrariamente, cuando la entalpía es pequeña, la constante de equilibrio cambia sólo moderadamente con la temperatura y la endoterma es más ancha. Para el caso donde el desplegamiento ocurre en un proceso de dos estados, se obtiene que (ecuación 10):^{14, 47}

$$\Delta H_{vH,1/2} = 4RT_{1/2}^2 \frac{C_{P,1/2}}{\Delta H_{cal}} \quad (\text{Ecuación 10})$$

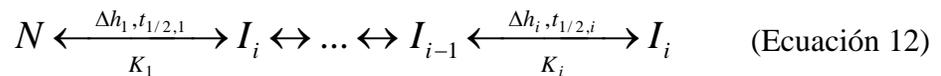
donde $T_{1/2}$ es la temperatura a la cual ocurre la mitad del área bajo la endoterma y $C_{P,T1/2}$ es el valor de la capacidad calorífica a esa temperatura. Así, para un proceso de dos estados la endoterma es simétrica, las poblaciones de ambos estados son iguales ($[N] = [D]$) a la $T_{1/2} = T_m$, y $C_{p,T1/2} = C_{p,Tm}$ es el máximo de la endoterma. Esta ecuación permite el cálculo de la ΔH_{vH} a partir de la traza calorimétrica obtenida de un experimento de CDB, que proporciona $T_{1/2}$, $C_{P,T1/2}$ y ΔH_{cal} . Es importante hacer notar que mientras que la evaluación de la entalpía calorimétrica es independiente del modelo, el cálculo de la entalpía de van't Hoff se basa en un modelo de dos estados. De esta forma, la coincidencia de ambos valores, es decir $\Delta H_{cal} = \Delta H_{vH}$, provee la evidencia más importante de que el proceso es de dos estados. Por otra parte, si $\Delta H_{cal} \neq \Delta H_{vH}$ la premisa de que el proceso es de dos estados es falsa. A la comparación entre ΔH_{cal} y ΔH_{vH} se le conoce como el criterio calorimétrico para evaluar si la transición es o no de dos estados. Este criterio puede ser generalizado para, otra vez a partir de la información en la traza calorimétrica, determinar si existen diferentes unidades cooperativas de plegamiento.¹⁴ Si una proteína tiene dominios estructurales, pero estos interactúan tan fuertemente que se despliegan o pliegan simultáneamente, entonces se sigue conservando que $\Delta H_{cal}/\Delta H_{vH} = 1$. Por tanto, no necesariamente el número de dominios estructurales corresponde al número de unidades cooperativas de plegamiento.

4.4 Transiciones múltiples

Como se mostró en la sección anterior, la CDB no sólo proporciona las características energéticas del proceso de plegamiento/desplegamiento sino que también es muy valiosa para acercarse a conocer el mecanismo o ruta por el cual este proceso ocurre. Para muchas proteínas, el mecanismo de dos estados regularmente es el adecuado para describir el desplegamiento. Sin embargo, la desnaturalización de muchas otras proteínas es más compleja y no puede ser descrita por este modelo. El hecho que el $\Delta H_{cal} \neq \Delta H_{vH}$, indica que en el proceso existen intermediarios termodinámicamente estables.⁴⁷ Normalmente, esta situación corresponde al caso en que ciertas regiones de la proteína, o dominios cooperativos, se desnaturalizan uno del otro de manera independiente o cercana a ello (cuasi-independiente). Para el caso de una proteína que presente dominios distintos, de forma general se pueden plantear dos mecanismos para el desplegamiento.⁴⁸ El primer modelo (ecuación 11) considera que los distintos dominios se despliegan de forma independiente y cada uno de ellos lo hace en una ruta de dos estados:



donde el subíndice i denota el i -ésimo dominio y X al resto de la proteína (los otros dominios, diferentes al i) tanto en el estado nativo como en el desnaturalizado. En este caso, la capacidad calorífica a cada temperatura es la suma de las contribuciones de cada dominio. El segundo modelo (ecuación 12), conocido como el modelo secuencial, contempla varias etapas o pasos secuenciales con intermediarios entre el estado nativo y desnaturalizado:



donde el subíndice i denota el i -ésimo paso e I_i el intermediario correspondiente. En este caso, la capacidad calorífica no es la suma simple de las contribuciones de cada intermediario. Para las proteínas en las que es posible elucidar cuál de estos dos mecanismos describe mejor su desnaturalización, es posible cooperar a responder la

pregunta básica de cómo las estructuras principales de la proteína son construidas, es decir distinguir entre el escenario donde dichas estructuras se forman paso a paso de forma secuencial o se forman por separado y luego se ensamblan para producir la estructura nativa final.²¹

Las trazas calorimétricas para proteínas cuyo mecanismo de desplegamiento sea más complejo que el indicado por el modelo de dos estados pueden tener varias formas. En algunos casos la traza tiene varios máximos o picos (figura 6A), indicando la existencia de varias transiciones. En otros, la traza tiene un solo máximo pero en realidad está compuesta de varias transiciones (figura 6B). En este último caso, la existencia de más de una transición puede deducirse de la asimetría de la endoterma, o más claramente de la aparición de un hombro en ella; también es posible que la endoterma sea simétrica y sin embargo sea el producto de más de una transición. En la mayoría de estos casos, la evidencia de la CDB no es contundente sino que es necesario obtener información adicional empleando otras técnicas, que puedan confirmar o negar el mecanismo propuesto a partir del análisis de la traza calorimétrica.

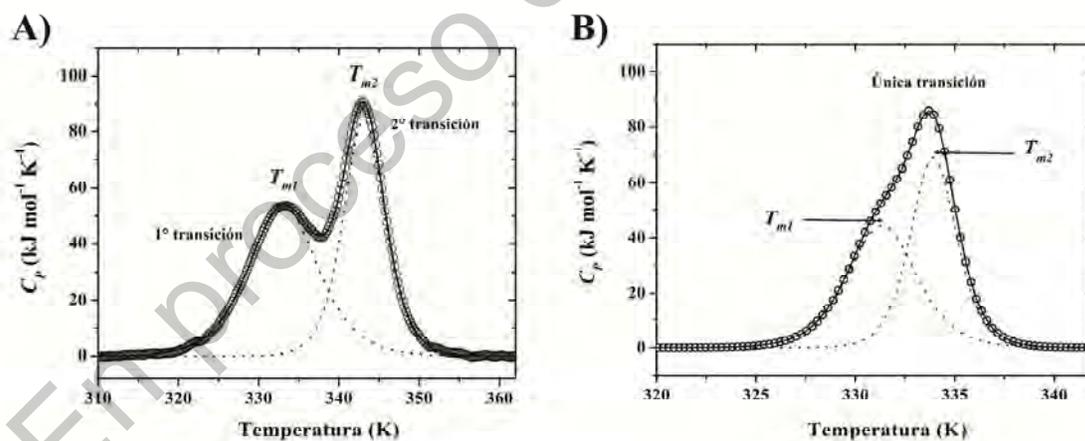


Figura 6. Ejemplos de perfiles de CDB para proteínas cuyo mecanismo de desplegamiento es diferente al de dos estados. Los símbolos vacíos representan los datos experimentales y las líneas continuas el ajuste de esos datos a un modelo. **A)** Endoterma cuyo desplegamiento térmico sigue un proceso secuencial; nótese las distintas transiciones en la traza calorimétrica. **B)** Endoterma que responde a un desplegamiento térmico de tres estados; nótese que la endoterma está formada por la contribución de dos transiciones Datos simulados en el programa Origin 7.0 (Microcal Inc.)

4.5 Transiciones continuas o plegamiento cuesta abajo

Desde años atrás se han reportado que proteínas pequeñas con un plegamiento rápido se comportan conforme al modelo de dos estados, pero con una barrera energética baja (o nula, en términos termodinámicos) entre los estados plegados y desplegados.⁴⁹⁻⁵² En general se asume que las barreras energéticas que separan los dos estados son lo suficientemente grandes en magnitud como para poder ser detectadas experimentalmente (figura 7A). Sin embargo, no hay ninguna razón física para que la barrera sea grande, de tal manera que es posible que existan barreras energéticas pequeñas; de hecho, sin la existencia de este tipo de barreras no sería claro cómo sucedería el plegamiento de una forma rápida.²¹ Si la barrera termodinámica está en el orden de magnitud al de la energía térmica (unos pocos kJ/mol), es posible que microestados con distintos grados de desplegamiento lleguen a estar poblados durante el desplegamiento al equilibrio. El análisis de datos muestra que las endotermas pueden describir energéticamente lo que sucede en el proceso de desplegamiento.⁵⁰ Una escala de entalpía (tomando como referencia el estado nativo) es usada como una medida del grado de desplegamiento siguiendo la teoría de Landau de las transiciones críticas.⁵³ Esto es, la energía libre vs. a la entalpía (H) (relacionada con el grado de desplegamiento) mostraría esencialmente un solo mínimo que podría desplazarse al cambiar la temperatura, y entonces el proceso de plegamiento-desplegamiento sería continuo e involucraría un único macroestado (figura 7A). A este mecanismo de plegamiento continuo en ausencia de barreras se le conoce como desplegamiento cuesta abajo (downhill en inglés).⁴⁹ La diferencia entre un mecanismo de dos estados con una barrera grande y uno continuo cuesta abajo se ve reflejado en el ancho y área bajo la curva de las trazas calorimétricas (figura 7B).

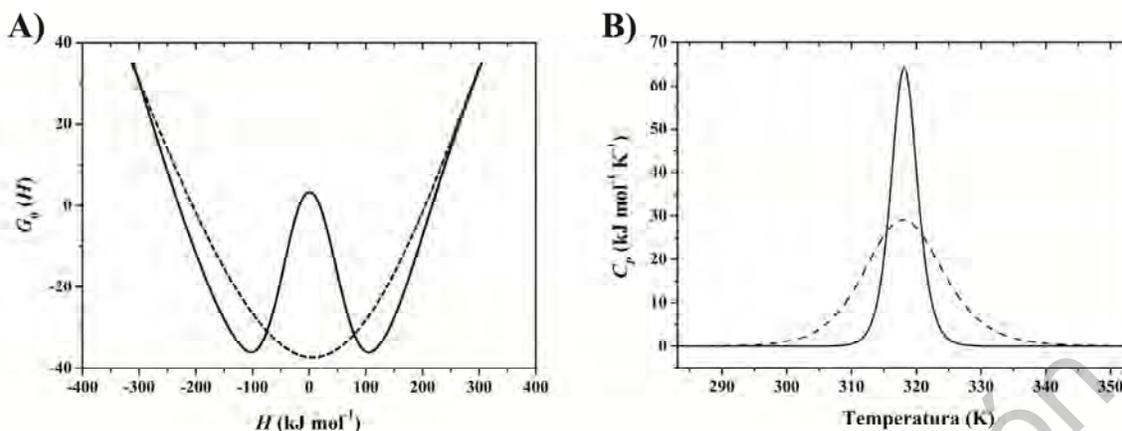


Figura 7. Comparación entre dos proteínas con mecanismos de plegamiento de dos estados y cuesta abajo. **A)** Se muestran los perfiles de energía libre vs. entalpía para una proteína con comportamiento de dos estados (línea continua) y una con plegamiento continuo cuesta abajo (con una barrera energética pequeña, comparable a la energía térmica) (línea discontinua). Nótese que en caso del plegamiento cuesta abajo sólo existe un mínimo de energía entre estados plegados y desplegados dado que no existe una barrera para llegar a ambos estados. **B)** Se muestran también las trazas calorimétricas obtenidas usando CDB para las mismas proteínas en el desplegamiento de dos estados (línea continua) y en el desplegamiento cuesta abajo (línea discontinua). Figura modificada de: Muñoz y Sánchez-Ruiz, 2004.

5. Aplicación de la Calorimetría Diferencial de Barrido al estudio de otros sistemas

Además del estudio de proteínas que se ha descrito en este capítulo, la CDB se ha utilizado en otras muchas áreas con interesantes aplicaciones y resultados. A continuación se describen brevemente ejemplos de estas, proporcionando algunas de las referencias más importantes en ellas:

- i. **Ácidos nucleicos.** Una de las aplicaciones iniciales de la CDB fue la caracterización fisicoquímica del ADN.⁵⁴ Actualmente, la técnica se utiliza como una herramienta para detectar la interacción de ciertos fármacos con esta macromolécula: una pérdida de la transición calorimétrica después de un tratamiento con un fármaco en estudio está correlacionado con la inhibición de la división celular inducida por ciertos agentes intercalantes, indicando que el fármaco interactúa con el ADN.⁵⁵ Además, en otros trabajos se han analizado nuevos procesos como el silenciamiento génico inducido por virus, técnicas nuevas de transfección, entre otros.⁵⁶

- ii. Lípidos.* La CDB puede determinar la temperatura, entalpía, entropía y cooperatividad de una amplia gama de transiciones de fase lipídicas y cómo estos parámetros dependen de las variaciones en la hidratación, pH y fuerza iónica experimentales.⁵⁷ Bajo ciertas condiciones, la CDB también puede caracterizar la cinética de algunas transiciones lipídicas.⁵⁸ Por lo tanto, los datos termodinámicos obtenidos a través de CDB pueden proveer información valiosa acerca de la organización de los ensamblajes lipídicos y cómo las propiedades físicas y estructurales de las membranas son moduladas por otros constituyentes de la bicapa o por su ambiente celular.⁵⁹
- iii. Pureza y estabilidad de los fármacos.* Las constantes físicas y los perfiles de pureza de algunos fármacos se han determinado mediante CDB.^{60, 61} La técnica también permite calcular las poblaciones del fármaco que pueden estar en distintas conformaciones o interaccionando con algún contaminante.⁵⁴ Por otra parte, la CDB es una herramienta útil para determinar la estabilidad de los fármacos y facilita el diseño de nuevas potenciales moléculas o el mejoramiento de las existentes.⁶²
- iv. Nanopartículas.* La CDB también puede ser usada para analizar la incorporación de fármacos dentro de nanopartículas, examinando los cambios de entalpía que suceden.^{63, 64} Por ejemplo, el análisis de las trazas calorimétricas de liposomas cargados con nanopartículas permite caracterizar el estado físico de estas, el polimorfismo de las partículas y la relación de incorporación del fármaco en ellas.
- v. Péptidos antimicrobianos.* La síntesis de moléculas nuevas para atacar infecciones bacterianas representa un área en constante desarrollo actual. La CDB ha proveído un análisis cuantitativo del efecto de la interacción de péptidos antimicrobianos con membranas bacterianas. Gracias a la caracterización energética de las interacciones entre las moléculas y sus blancos terapéuticos ha sido posible desarrollar modelos de unión para el diseño racional dirigido de nuevas moléculas terapéuticas.⁶⁵

6. Referencias

1. Sánchez-Ruíz, J. M., (2011), Probing free-energy surfaces with Differential Scanning Calorimetry, *Annu Rev Phys Chem*, 62, 231-255.
2. Gruebele, M., (2005), Downhill protein folding: evolution meets physics, *C R Biol*, 238, 701-712.
3. Sánchez-Ruíz, J. M., (2010), Protein kinetic stability, *Biophys Chem*, 148, 1-15.
4. O'Neill, M. J., Watson, E. S., (1962), Differential microcalorimeter. U.S. Patent: US3263484 A, Washington D.C., U.S.
5. Privalov, P. L., (1975), Results of Science and Technology, *Mol Biol*, 6, 7.
6. Cooper, A., Johnson, C. M., (1994), Differential Scanning Calorimetry, *Methods Mol Biol*, 22, 125-136.
7. Freire, E., (1995), Differential Scanning Calorimetry, en: B. A. Shirley (Ed), *Methods in Molecular Biology*, Vol. 40: Protein stability and folding: theory and practice, New Jersey: *Humana Press*, 191-218.
8. Privalov, P. L., (1980), Scanning microcalorimeters for studying macromolecules, *Pure Appl Chem*, 52, 479-497.
9. Plotnikov, V. V., Brandts, J. M., Lin, L. N., Brandts, J. F., (1997), A new ultrasensitive scanning calorimeter, *Anal Biochem*, 250, 237-244.
10. Spink, C. H., (2008), Differential Scanning Calorimetry, en: J. Correia, H. W. Detrich (Eds), *Biophysical tools for biologists*, Massachusetts: *Academic Press*, 115-142.
11. Halliday, D., Resnick, R., Walker, J., (2013), *Fundamentals of physics extended*, New Jersey: *Wiley*, 672 pp.
12. Quinn, C., (2011), Buffer compatibility with Nano DSC, Technical note: *TA Instruments*, 1-9.
13. Freire, E., Biltonen, R. L., (1978), Statistical mechanical deconvolution of thermal transitions in macromolecules. I. Theory and application to homogeneous systems, *Biopolymers*, 17, 463-479.
14. Freire, E., (1989), Statistical thermodynamic analysis of the heat capacity function associated with protein folding-unfolding transitions, *Comments Mol Cell Biophys*, 6, 123-140.
15. Freire, E., van Osdol, W. W., Mayorga, O. L., Sanchez-Ruiz, J. M., (1990), Calorimetrically determined dynamics of complex unfolding transitions in proteins, *Annu Rev Biophys Biophys Chem*, 19, 159-188.
16. Gomez, J., Hilser, V. J., Xie, D., Freire, E., (1995), The heat capacity of proteins, *Proteins*, 22, 404-412.
17. Makhatadze, G. I., Privalov, P. L., (1995), Energetics of protein structure, *Adv Prot Chem*, 47, 307-426.
18. Sturtevant, J. M., (1977), Heat capacity and entropy changes in processes involving proteins, *Proc Natl Acad Sci USA*, 74, 2236-2240.
19. Prabhu, N. V., Sharp, K. A., (2005), Heat capacity in proteins, *Annu Rev Phys Chem*, 56, 521-548.
20. Johnson, C. M., (2013), Differential Scanning Calorimetry as a tool for protein folding and stability, *Arch Biochem Biophys*, 531, 100-109.
21. Ibarra-Molero, B., Sánchez-Ruíz, J. M., (2006), Differential Scanning Calorimetry of proteins: an overview and some recent developments, en: J. L. R. Arrondo and A. Alonso (Eds), *Advanced techniques in Biophysics*, Berlin: *Springer Berlin Heidelberg*, 27-48.

22. Romero-Romero, S., Costas, M., Rodríguez-Romero, A., Fernández-Velasco, D. A., (2015), Reversibility and two state behaviour in the thermal unfolding of oligomeric TIM barrel proteins, *Phys Chem Chem Phys*, 17, 20699-20714.
23. Sánchez-Ruíz, J. M., (1992), A theoretical analysis of Lumry-Eyring models in Differential Scanning Calorimetry, *Biophys J*, 61, 921-935.
24. Plaza del Pino, I. M., Ibarra-Molero, B., Sánchez-Ruíz, J. M., (2000), Lower kinetic limit to protein thermal stability: a proposal regarding protein stability in vivo and its relation with misfolding diseases, *Protein*, 40, 58-70.
25. Lyubarev, A. E., Kurganov, B. I., (1998), Modeling of irreversible thermal protein denaturation at varying temperature. I. The model involving two consecutive irreversible steps, *Biochemistry*, 63, 434-440.
26. Arrhenius, S.A., (1889), Über die Dissociationswärme und den Einfluß der Temperatur auf den Dissociationsgrad der Elektrolyte, *Z Phys Chem*, 4, 96-116.
27. Creighton, T. E., (1990), Protein folding, *Biochem J*, 15, 1-16.
28. Dill, K. A., Shortle, D., (1991), Denatured states of proteins, *Annu Rev Biochem*, 60, 795-825.
29. Creighton, T. E., (1996), Proteins: structure and molecular properties, New York: *Freeman*, 507 pp.
30. Pain, R. H., (2000), Mechanisms of protein folding, New York: *Oxford University Press*, 325 pp.
31. Rees, D. C., Robertson, A. D., (2001), Some thermodynamic implication for the thermostability of proteins, *Prot Sci*, 10, 1187-1194.
32. Pace, N. C., Tanford, Ch., (1968), Thermodynamics of the unfolding of β -lactoglobulin A in aqueous urea solution between 5 and 55 °C, *Biochemistry*, 7, 198-208.
33. Privalov, P. L., (1990), Cold denaturation of proteins, *Crit Rev Biochem Mol Biol*, 25, 281-305.
34. Fersht, A., (1999), Structure and mechanism in protein science. A guide to enzyme catalysis and protein folding, New York: *Freeman*, 631 pp.
35. Frank, H. S., Evans, M. W., (1945), Free volume and entropy in condensed systems III. Entropy in binary liquid mixtures; partial molar entropy in dilute solutions; structure and thermodynamics in aqueous electrolytes, *J Chem Phys*, 13, 507-532.
36. Kauzmann, W., (1959), Some factors in the interpretation of protein denaturation, *Adv Prot Chem*, 14, 1-63.
37. Brandts, J. F., (1964), The thermodynamics of protein denaturation, *J Amer Chem Soc*, 86, 4291-4301.
38. Privalov, P. L., Khechinashvili, N. N., (1974), A thermodynamic approach to the protein problem of stabilization of globular protein structure: a calorimetric study. *J Mol Biol*, 86, 665-684.
39. Shiao, D. F., Lymry, R., Fahey, J., (1971), Studies of chymotrypsinogen family proteins. Heat capacity changes accompanying reversible unfolding of proteins. *J Amer Chem Soc*, 93, 2024-2035.
40. Doyle, C. M., Rumfeldt, J. A., Broom, H. R., Broom, A., Stathopoulos, P. B., Vassall, K. A., Almey, J. J., Meiering, E. M., (2013), Energetics of oligomeric protein folding and association, *Arch. Biochem. Biophys*, 531, 44-64.
41. Privalov, P. L., Makhataдзе, G. I., (1990), Heat capacity of proteins. I. Partial molar heat capacity of individual amino acid residues in aqueous solution: hydration effect, *J Mol Biol*, 213, 385-391.
42. Murphy, K. P., Bhakuni, V., Xie, D., Freire, E., (1992), Molecular basis of co-operativity in protein folding. III. Structural identification of cooperative folding units and folding intermediates, *J Mol Biol*, 227, 293-306.
43. Murphy, K. P., Freire, E., (1992), Thermodynamics of structural stability and cooperative folding

behavior in proteins, *Adv Protein Chem*, 43, 313-361.

44. Zhou, Y. Q., Hall, C. K., Karplus, M., (1999), The calorimetric criterion for a two-state process revisited, *Protein Sci*, 8, 1064-1074.
45. Kaya, H., Chan, H. S., (2000), Polymer principles of protein calorimetric two-state cooperativity, *Proteins*, 40, 637-661.
46. Atkins, P., De Paula, J., (2010), Physical Chemistry, New York: *Freeman*, 965 pp.
47. Robertson, A. D., Murphy, K. P., (1997), Protein structure and the energetics of protein stability, *Chem Rev*, 97, 1251-1267.
48. Microcal Inc., (2004), DSC data analysis in Origin. Tutorial guide, Massachusetts: *Microcal Inc.*, 1-110 pp.
49. García-Mira, M. M., Sadqui, M., Fischer, N., Sánchez-Ruíz, J. M., Muñoz, V., (2002), Experimental identification of downhill folding, *Science*, 298, 2191-2195.
50. Muñoz V., Sánchez-Ruíz, J. M., (2004), Exploring protein-folding ensembles: a variable-barrier model for the analysis of equilibrium folding experiments, *Proc Natl Acad Sci USA*, 101, 17646-17651.
51. Oliva, F. Y., Muñoz, V., (2004), A simple thermodynamic test to discriminate between two-state and downhill folding, *J Am Chem Soc*, 126, 8596-8597.
52. Naganathan, A. N., Pérez-Jiménez, R., Sánchez-Ruíz, J. M., Muñoz, V., (2005), Robustness of downhill folding: guidelines for the analysis of equilibrium folding experiments on small proteins, *Biochemistry*, 44, 7435-7449.
53. Landau, L., (1937), On the theory of phase transitions, *Zh Eksp Teor Fiz*, 7, 19-32.
54. Chiu, M. H., Prenner, E. J., (2011), Differential Scanning Calorimetry: an invaluable tool for a detailed thermodynamic characterization of macromolecules and their interaction, *J Pharm Bioallied Sci*, 3, 39-59.
55. Almagor, M., Cole, R. D., (1989), Differential Scanning Calorimetry of nuclei as a test for the effect of anticancer drugs on human chromatin, *Cancer Res*, 49, 5561-5566.
56. Litzinger, D. C., Huang, L., (1992), Phosphatidylethanolamine liposomes: drug delivery, gene transfer and immunodiagnostic applications, *Biochim Biophys Acta* (1113): 201-227.
57. Lewis, R., McElhaney, R. N., (2009), The physicochemical properties of cardiolipin bilayers and cardiolipin-containing lipid membranes, *Biochim Biophys Acta Biomembr*, 1788, 2069-2079.
58. Heerklotz, H., (2004), The microcalorimetry of lipid membranes, *J Phys Condens Matter*, 16, 441-467.
59. Ulkowski, M., Musialik, M., Litwinienko, G., (2005), Use of Differential Scanning Calorimetry to study lipid oxidation. 1. Oxidative stability of lecithin in linolenic acid, *J Agric Food Chem*, 53, 9073-9077.
60. Grady, L. T., Hays, S. E., King, R. H. Klein, H. R. Mader, W. J., Wyatt, D. K., (1973), Drug purity profiles, *J Pharma Sci*, 62, 456-464.
61. Giron, D., (2002), Applications of thermal analysis and coupled techniques in pharmaceutical, *J Thermal Anal Calorim*, 68, 335-357.
62. Capelle, M. A., Gurny, R., Arvinte, T., (2007), High throughput screening of protein formulation stability: practical considerations, *Eur J Pharma Biopharma*, 65, 131-148.
63. Castelli, F., Puglia, C., Sarpietro, M. G., Rizza, L., Bonina, F., (2005), Characterization of indomethacin-loaded lipid nanoparticles by differential scanning calorimetry, *Int J Pharma*, 304, 231-238.

64. Vasconcelos, T., Sarmiento, B., Costa, P., (2007), Solid dispersions as strategy to improve oral bioavailability of poor water soluble drugs, *Drug Discov Today*, 12, 1068-1075.
65. Lohner, K., Prenner, E. J., (1999), Differential Scanning Calorimetry and X-ray diffraction studies of the specificity of the interaction of antimicrobial peptides with membrane mimetic systems, *Biochim Biophys Acta*, 1462, 141-156.

En proceso de publicación.