



**UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO**

Facultad de Medicina

Licenciatura en Investigación
Biomédica Básica

**Caracterización fisicoquímica de un ancestro eucarionte
de la enzima *triosafosfato isomerasa***

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

LICENCIADA EN INVESTIGACIÓN BIOMÉDICA BÁSICA

P R E S E N T A:

MARIANA SCHULTE-SASSE JIMÉNEZ

Facultad de Medicina



Director de tesis:

Dr. Daniel Alejandro Fernández Velasco

Mayo, 2013



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



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 Abu, por tu apoyo, generosidad y cariño.
Siempre estás presente.

Agradecimientos académicos

A la Universidad Nacional Autónoma de México, por ser mi segunda casa durante los últimos 5 años.

Al pueblo de México que paga sus impuestos.

El desarrollo de este proyecto fue financiado parcialmente por los proyectos PAPIIT IN 206510, PAPIIT IN 219913, de los que también recibí apoyo económico para mi conclusión de estudios; y la Facultad de Medicina de la UNAM.

A la SEP, por la beca de titulación que me otorgó.

A los miembros del jurado, por tomarse el tiempo de revisar el presente trabajo:

Dr. Salvador Uribe Carvajal

Dr. D. Alejandro Fernández Velasco

Dr. León D. Islas Suárez

Dr. Fernando López Casillas

Dr. León P. Martínez Castilla

Al Dr. Miguel Costas Basín y al laboratorio de termofísica de la Facultad de Química por dejarnos usar generosamente su equipo.

A la Dra. Adela Rodríguez Romero y al Laboratorio Nacional de Estructura de Macromoléculas, por la dedicación mostrada.

Al Dr. Enrique García Hernández, por su paciencia, amabilidad y dejarnos usar el equipo de su laboratorio.

A las técnicas del LFQIP, M. en C. Laura Ileana Álvarez Añorve y la Dra. Ma. Isabel Velázquez, por su valioso trabajo.

Agradecimientos personales

A mi mamá. Eres mi mejor maestra, mi mejor amiga, mi fan número uno y la persona que más admiro.

A mi papá, porque tus abrazos son los mejores, porque me contagias tu curiosidad y por tu espíritu altruista, con personas como tú “otro mundo sí es posible”.

A Jonas, gracias por no dejarme ser una hija única odiosa. Admiro tu inteligencia, rebeldía y determinación.

A Daniel. Te encontré cuando menos te buscaba e iluminaste mi mundo. Sé que en tus brazos estoy segura.

A toda la familia Jiménez, por la solidaridad.

A mis *manas*: Nora, Ana y Andrea. Son la definición de amigas para toda la vida. Nuestras terapias de risa son las mejores.

A todos los integrantes de la *Simbiose*: Mariana, Adriana, Rosi, Olivia, Gaby y Jan. Porque me aceptan con todo y mi ñoñez y son amigos incondicionales.

A Dani, porque estuviste a mi lado en los momentos más difíciles y también en los mejores.

A todos mis amigos en la carrera: Manuel, Lalo, Alberto y Alicia. Tomar clases a su lado fue un honor.

A Diana y Silvana. Ustedes son mi equipo de trabajo ideal, no solamente por la dedicación e inteligencia, sino también por la amistad y el compañerismo. Las admiro.

A Hugo, Jaz y Juss. Son el ejemplo de que los mejores amigos se pueden encontrar en cualquier momento. Con ustedes no existe el tiempo; son mis hermanos.

A Alejandro Fernández. Sin duda, el mejor tutor que alguien puede tener. Me contagias tu pasión por las proteínas y por la ciencia. Gracias por la dedicación a tu oficio, es un ejemplo para la futuras generaciones.

A Fátima, porque ella empezó el trabajo y me contagió su entusiasmo.

A Jesús. Gracias por las pláticas sinceras y por guiarme en el mundo de biomédicas desde que nos topamos en mi primer laboratorio.

A Haven. Gracias por compartir esa visión única y diferente que tienes del mundo. Me has enseñado a nunca conformarme.

A Sergio. Porque no tenemos que hablar para entendernos. Eres mi mejor ejemplo a seguir como ser humano.

Por último, a todo el LFQIP. Formar parte de la comuna es una experiencia mágica, no hay mejor lugar de trabajo.

ÍNDICE

ABREVIATURAS	VII
RESUMEN	VIII
INTRODUCCIÓN	1
PLEGAMIENTO DE PROTEÍNAS	2
RECONSTRUCCIÓN ANCESTRAL	3
LA <i>TRIOSAFOSFATO ISOMERASA</i>	7
ANTECEDENTES	15
HIPÓTESIS	22
OBJETIVOS	22
OBJETIVO GENERAL	22
OBJETIVOS PARTICULARES	22
MATERIALES Y MÉTODOS	23
CLONACIÓN Y EXPRESIÓN	23
PURIFICACIÓN DE LAS PROTEÍNAS ANCESTRALES	24
PROPIEDADES ESPECTROSCÓPICAS	25
ESTABILIDAD TÉRMICA	28
CATÁLISIS	31
ESTABILIDAD A LA DILUCIÓN	32
RESULTADOS Y DISCUSIÓN	36
EXPRESIÓN Y SOLUBILIDAD DE LAS PROTEÍNAS RECOMBINANTES	36
ESTRUCTURAS PLEGADAS	37
ESTABILIDAD TÉRMICA DE LOS ANCESTROS	40
CATÁLISIS	48
CARACTERIZACIÓN DE LA DISOCIACIÓN	50
CONCLUSIONES Y PERSPECTIVAS	62
BIBLIOGRAFÍA	68
APÉNDICE I	73
CATH	73
SCOP	75

ABREVIATURAS

TIM:	<i>triosafosfato isomerasa</i>
G3P:	gliceraldehído-3-fosfato
DHAP:	dihidroxiacetona fosfato
TmTIM:	TIM de <i>Tenebrio molitor</i>
BsTIM:	TIM de <i>Bacillus subtilis</i>
OcTIM:	TIM de <i>Oryctolagus cuniculus</i>
GgTIM:	TIM de <i>Gallus gallus</i>
LmTIM:	TIM de <i>Leishmania mexicana</i>
HsTIM:	TIM de <i>Homo sapiens</i>
ScTIM:	TIM de <i>Saccharomyces cerevisiae</i>
EhTIM:	TIM de <i>Entamoeba histolytica</i>
TtTIM:	TIM de <i>Thermotoga maritima</i>
PfTIM:	TIM de <i>Plasmodium falciparum</i>
TbTIM:	TIM de <i>Trypanosoma brucei</i>
TcTIM:	TIM de <i>Trypanosoma cruzi</i>
GITIM:	TIM de <i>Giardia lamblia</i>
K_m:	constante de Michaelis-Menten
k_{cat}:	constante catalítica
DSC:	calorimetría diferencial de barrido
T_m:	temperatura media de desnaturalización
C_p:	capacidad calorífica
ITC:	calorimetría de titulación isotérmica
K_{dis}:	constante de disociación
ΔH_{dis}:	entalpía de disociación
ΔG_{dis}:	energía libre de Gibbs de disociación
ΔS_{dis}:	entropía de disociación

RESUMEN

La enzima glucolítica *triosafosfato isomerasa* (TIM) es un barril (β/α)₈ oligomérico responsable de catalizar la interconversión de D-gliceraldehído 3-fosfato (GAP) y dihidroxiacetona fosfato (DHAP), una reacción esencial en la glicólisis. La TIM es considerada un catalizador perfecto debido a que su cinética está solamente limitada por la difusión del sustrato al sitio activo de la enzima. Hasta el momento no se ha encontrado ni se ha podido crear una TIM monomérica activa, a pesar de que el sitio activo es independiente en cada subunidad, de hecho, se ha observado que la contribución mayoritaria a la estabilidad global de la enzima proviene del paso de dimerización en la vía de plegamiento. La TIM es ubicua en los tres dominios de la vida y su estructura tridimensional está sumamente conservada entre diversas especies, esto se debe a que el tipo de plegamiento es muy robusto y permite bastante variabilidad en la secuencia. Sin embargo, mutaciones en sitios específicos pueden afectar significativamente propiedades de la enzima como el patrón de plegamiento o la estabilidad. El estudio comparativo de proteínas ha sido útil para entender mejor la relación entre la secuencia y las propiedades fisicoquímicas, sin embargo, uno de los aspectos que carece este enfoque es el de poder dar una explicación integrativa y evolutiva al problema. Para dilucidar cómo se generaron las propiedades catalíticas, el estado de asociación y la estabilidad de estas enzimas, en esta tesis examinamos la historia molecular de la TIM mediante la reconstrucción de proteínas ancestrales y su posterior caracterización fisicoquímica.

En un trabajo previo se utilizaron métodos de máxima verosimilitud para inferir las secuencias de TIMs ancestrales de eucariontes. Se escogieron dos nodos que presentan una alta verosimilitud de reconstrucción (>0.8) para un alto porcentaje ($>70\%$) de los sitios, estos corresponden al ancestro común de animales y hongos así como para el nodo que incluye a las plantas. El objetivo de la presente tesis fue caracterizar a esas dos proteínas ancestrales de manera fisicoquímica y termodinámica. Los genes que codifican para estas proteínas ancestrales fueron

sintetizados y expresados en *E. coli*, las proteínas recombinantes se purificaron a homogeneidad a partir de la fracción soluble. El radio de Stokes observado corresponde a un dímero para ambos ancestros y sus características espectroscópicas sugieren estructura secundaria y terciaria. Su termoestabilidad fue analizada mediante técnicas de dicroísmo circular y calorimetría diferencial de barrido (DSC), ambas proteínas ancestrales presentan termoestabilidad mayor a la encontrada en las TIM extantes (proteínas que existen hoy en día); son además catalíticamente activas. Estudios de calorimetría de titulación isotérmica (ITC) y ensayos de dilución monitoreados por actividad permitieron observar que la superficie enterrada en la interfase del dímero es comparable a la de TIMs extantes, que la energía asociada al proceso de dimerización es baja y que la constante de disociación (K_{dis}) es muy alta resultando en una baja afinidad entre monómeros. Para poder entender mejor estos resultados y poder dar una explicación estructural, está en proceso la determinación estructural por cristalografía de rayos X de uno de los ancestros.

La reconstrucción ancestral es una metodología relativamente nueva que nos permite crear secuencias hipotéticas, en este trabajo solamente se sintetizaron secuencias que presentaron la mayor probabilidad de reconstrucción, sin embargo, es necesario tomar en cuenta que existen segundas posiciones de reconstrucción para cada residuo y que se pueden analizar y comparar con aquellos que están generando conflicto en la catálisis o en la interfase. Es un proceso iterativo que puede ser muy útil para la comprensión integrativa de la historia que generó las enzimas que existen hoy en día.

INTRODUCCIÓN

Miles de millones de organismos provenientes de millones de especies diferentes coexistimos en la tierra, desde microorganismos extremófilos que viven en ambientes con temperaturas por encima de los 100°C hasta bacterias que resisten temperaturas cercanas a cero grados centígrados, todo esto pasando por arqueas, eucariontes unicelulares, parásitos, hongos, plantas, etcétera etcétera. Además, existen eucariontes complejos como el ser humano cuya vida corre peligro cuando la temperatura corporal sale de un rango específico (entre 34 y 40°C). Pero a pesar de la gran diversidad, todos estamos conformados por células, la unidad básica de la vida.

Cada organismo puede diferenciarse en las particularidades de sus células, sin embargo, visto desde un punto de vista práctico, las células son sistemas químicos sumamente complejos que obedecen las leyes de la física y la química. Lo particular de estos sistemas es que están constituidos principalmente por compuestos de carbono que forman moléculas poliméricas a las que llamamos biomoléculas: lípidos, ácidos nucleicos y proteínas. De entre todas ellas, las proteínas son las que componen la mayoría de la masa seca de una célula. Si se mira a una célula en el microscopio o se analiza su actividad eléctrica o bioquímica, lo que realmente se está observando son sus proteínas. Regresando al ejemplo del inicio, la razón por la que un organismo vive o muere a cierta temperatura extrema se debe a ellas. Las proteínas efectúan prácticamente cualquier proceso que sucede en una célula. Desde un punto de vista químico son las moléculas estructuralmente más complejas y funcionalmente más sofisticadas que se conocen. Algunos ejemplos de las muchas funciones en las que participan son transporte, estabilidad estructural, expresión génica, señalización, biosíntesis de otras moléculas, reconocimiento molecular y metabolismo. La relación estructura-función en las proteínas es fundamental para entender el proceso de la vida, por ello, es importante estudiar a detalle estas biomoléculas.

Plegamiento de proteínas

Las proteínas son polímeros de aminoácidos, también llamados polipéptidos, de diferentes longitudes. Son 20 aminoácidos los que se usan en las células para sintetizar proteínas y se distinguen por la composición de su cadena lateral. Cada proteína es resultado de una combinatoria específica entre ellos, es decir, cada proteína tiene una secuencia característica de aminoácidos. Dado que el enlace entre residuos es siempre el mismo (enlace peptídico), lo que determina las diferencias en la estructura tridimensional que adoptan las proteínas son las cadenas laterales. Se le llama plegamiento al proceso mediante el cual una proteína adopta su conformación tridimensional funcional (conformación nativa), a partir de una secuencia lineal; todavía se desconoce específicamente a aquello que guía el proceso en la secuencia. Ya desde 1961 Anfinsen publicó el primer trabajo en el que demuestra que la Ribonucleasa A puede plegarse hasta llegar a la estructura nativa solo a partir de la secuencia lineal de aminoácidos (Anfinsen *et al.*, 1961).

Algunas preguntas que han intrigado a científicos por décadas con respecto al tema son las siguientes: ¿cómo dictan las cadenas laterales de los aminoácidos el proceso de plegamiento? ¿Cómo encuentra una secuencia lineal su estructura tridimensional en el paisaje conformacional posible? Podríamos pensar que secuencias similares llevan al mismo plegamiento o que secuencias muy diferentes codifican para plegamientos también diferentes, sin embargo, la segunda aseveración es falsa. Un ejemplo muy claro de esto es el del plegamiento tipo barril (β/α)₈, también llamado barril TIM, representado en aproximadamente el 10% de las estructuras depositadas en el banco de datos de proteínas (PDB) (Farber & Petsko, 1990). Bases de datos de proteínas como SCOP y CATH (ver Apéndice 1), que clasifican a las proteínas en función de su estructura, distinguen 33 y 29 superfamilias de barriles TIM, respectivamente. La mayoría son enzimas y hay al menos una representante de cinco de las seis clases de enzimas definidas por la comisión de enzimas (EC) (Nagano *et al.*, 2002; Wierenga, 2001). Las secuencias tienen una alta variabilidad entre ellas, algunas presentan una

identidad cercana a 12%, por ejemplo las enzimas *triosafosfato isomerasa* de *Plasmodium falciparum* y la de *Pyrococcus woesei*.

El plegamiento $(\beta/\alpha)_8$ consta de ocho repeticiones del motivo hebra β -hélice α que globalmente se acomodan para formar una hoja β cerrada con las hebras paralelas en el interior del barril rodeada por las hélices. Cada elemento de estructura secundaria está conectado entre sí por medio de asas, de manera que existen ocho asas $\beta\alpha$, que conectan el extremo carboxilo de las hebras con las hélices, y ocho asas $\alpha\beta$ que conectan el extremo carboxilo de las hélices con las hebras (fig. 1). Para todas las enzimas con este tipo de estructura tridimensional, sin importar la baja conservación de la secuencia, el sitio activo está localizado en la cara C-terminal del barril (denominada de esta manera por ser el extremo carboxilo de las hebras β), a pesar de que no presenta una evidente ventaja sobre la cara N-terminal; por ello se ha sugerido que los barriles tienen un origen evolutivo común (Nagano et al. 2002).

A partir de lo anterior, se debe tomar en cuenta que si bien las características fisicoquímicas de los aminoácidos que conforman una proteína determinan sus propiedades, las proteínas se encuentran en un ambiente celular que está sujeto al proceso evolutivo. Tanto las fuerzas evolutivas como las determinantes fisicoquímicas deben ser tomadas en cuenta si se quiere entender mejor el problema del plegamiento de proteínas.

Reconstrucción ancestral

Existen dos posibles caminos para realizar un análisis evolutivo en una familia de proteínas. El primero se basa en lo conocido, es decir, en las secuencias de las proteínas que existen hoy en día (secuencias extantes). A partir de alineamientos se identifican las diferencias en la secuencia que posiblemente generan diferencias en las propiedades funcionales. Por medio de mutagénesis sitio-dirigida se prueba el papel funcional de los residuos candidatos. Este enfoque "horizontal" encuentra en muchas ocasiones residuos importantes para una función, ya que al mutarlos se generan proteínas no funcionales, sin embargo, muy rara vez se encuentra el conjunto de aminoácidos necesarios para cambiar la

función de una proteína por el de otra en la misma familia. Esto se debe, por un lado, a que pueden existir diferencias funcionalmente irrelevantes que se acumulan en las secuencias y, por el otro, a la epistasia. Esto último es el fenómeno en el que el efecto de una mutación depende del contexto genético en que se genere, por ejemplo, hay mutaciones *permissivas* que propician el surgimiento de nuevas funciones y otras *restrictivas* que disminuyen la posibilidad de cambios. Si solamente se observan las diferencias en las secuencias extantes muchas veces se ignora la historia evolutiva (fig. 2) (Harms & Thornton, 2010).

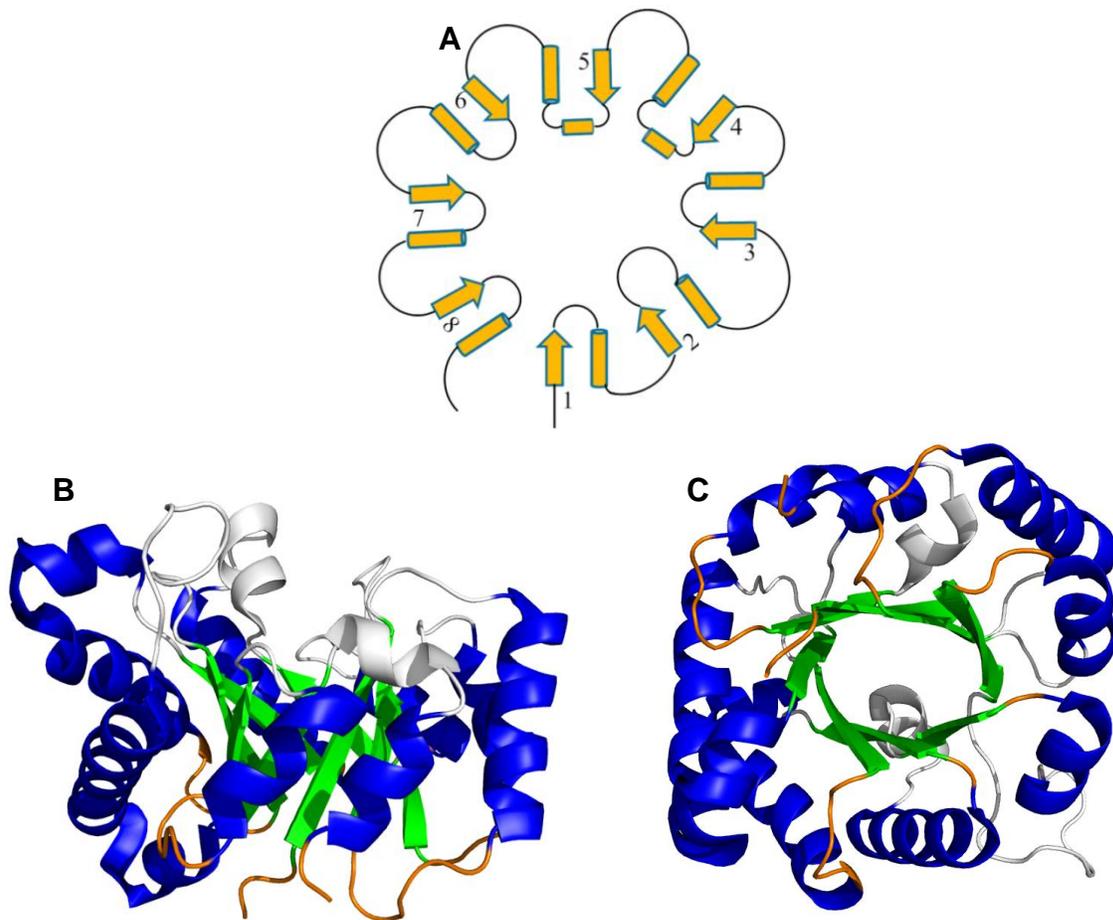


Figura 1. Plegamiento tipo barril TIM. A) Representación esquemática del plegamiento, las flechas representan las hebras β y los cilindros las hélices α . B) vista lateral y C) desde la base del barril. Código de colores, verde: hebras β , azul: hélices α , anaranjado: asas $\beta\alpha$, blanco: asas $\alpha\beta$.

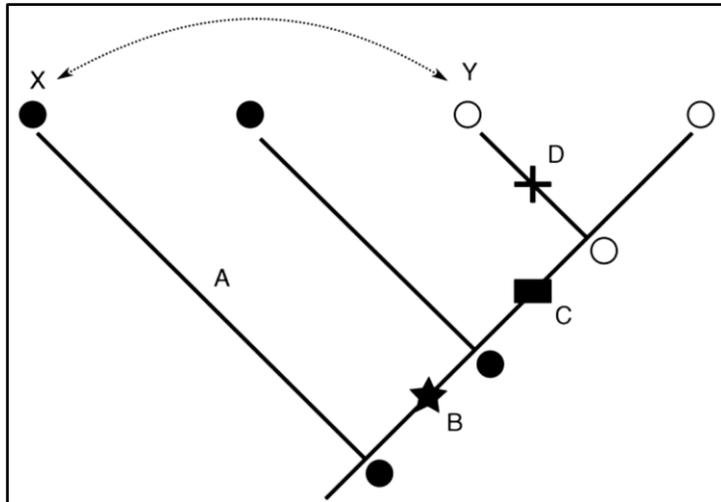


Figura 2. Diversificación funcional en una familia de proteínas. ●: función ancestral, ○: función derivada. El cambio funcional fue causado por un conjunto de cambios en la secuencia a lo largo de la rama C (■). En este escenario se requirieron mutaciones permisivas en la rama B (★) para que la proteína pueda tolerar las mutaciones necesarias para el cambio de función. En la rama D se acumularon mutaciones restrictivas incompatibles con la función ancestral (×). El intercambio de residuos entre las proteínas modernas (flecha) es ineficiente porque las secuencias difieren en las mutaciones en las ramas A, B, C y D. La proteína X no tiene las mutaciones permisivas, por lo mismo, no puede aceptar la función derivada. La proteína Y tiene mutaciones restrictivas que le impiden tolerar la función ancestral. Tomado de (Harms & Thornton, 2010).

El segundo camino corresponde a un análisis filogenético explícito de manera “vertical”. Al conocer el contexto evolutivo de la proteína, es posible identificar los cambios que llevan a epistasis y tomarlos en cuenta. Para este tipo de enfoque es necesario primero conocer la historia evolutiva de las proteínas, es decir, conocer sus ancestros, los cuales, por definición ya no existen. La única forma experimental de lograrlo es reconstruir a esos ancestros y “revivirlos” en el laboratorio. Para ello se utiliza una metodología relativamente nueva: la reconstrucción ancestral de secuencias (Thornton, 2004; Liberles, 2007).

El proceso evolutivo se genera a partir de mutaciones en la secuencia de un gen que por deriva génica y/o selección natural en algún momento se pueden fijar en una población. En caso de que la mutación sea en una región codificante para un carácter morfológico o fisiológico, los individuos de la población presentarán

dicho fenotipo. Para poder rastrear el origen de los caracteres que observamos y entender el mecanismo evolutivo por el que se originaron, es necesario establecer un árbol filogenético que nos hable de las relaciones evolutivas entre las secuencias extantes. De esta manera es posible inferir la secuencia de un gen o proteína ancestral. Los primeros científicos que plantearon esta idea fueron Zuckerkandl y Pauling en 1965 (Zuckerkandl & Pauling, 1965), sin embargo, en ese tiempo era metodológicamente imposible llevar a cabo el proceso. Gracias a la secuenciación masiva, el desarrollo de la teoría evolutiva molecular, la biología molecular y sobre todo el aumento en el poder de cómputo, esta idea se ha convertido en una realidad. El estudio comparativo de proteínas ancestrales y extantes genera una infinidad de preguntas que se pueden plantear, algunas se presentan a continuación.

En un primer nivel, se puede buscar simplemente el fenotipo de una proteína ancestral a partir de proteínas extantes que divergieron hace millones de años y presentan hoy en día fenotipos diferentes. El trabajo de Ugalde y colaboradores en 2004 logra no solamente resolver una pregunta de este tipo al descubrir el color del ancestro de proteínas similares a la proteína verde fluorescente con una edad aproximada de 20 millones de años, sino que también descifra cuáles son las mutaciones implicadas en cambiar el color de la fluorescencia (Ugalde et al., 2004).

Otro tipo de preguntas se pueden resolver de manera indirecta, tal es el caso del trabajo en el que se estudió la termoestabilidad de factores de elongación ancestrales (EF-TU) y por medio de la determinación de su temperatura media de desnaturalización (T_m) se logró hacer una aproximación de la temperatura de la tierra en el precámbrico (hace 3500 millones de años); los resultados coinciden con datos geológicos (Gaucher et al., 2008).

Por último, existen proteínas con la misma función, pero diferente especificidad por uno o más sustratos. Desde una aproximación no evolutiva se ha intentado hacer una comparación de las secuencias extantes y a partir de mutaciones puntuales cambiar la especificidad de la enzima por su sustrato. La mayoría de las

veces esto ha fracasado y se debe a que a lo largo del proceso evolutivo las mutaciones que ocurren son varias y en muchas ocasiones están ligadas entre sí, por lo que una mutación no es suficiente para cambiar la especificidad. Esto se demostró en un trabajo en el que la reconstrucción de receptores de corticoides ancestrales permitió identificar los aminoácidos que promovían el reconocimiento diferencial de ligandos en las proteínas extantes (Ortlund et al., 2007).

La *triosafosfato isomerasa*

El modelo de estudio en la presente tesis es la enzima glucolítica *triosafosfato isomerasa* (TIM), proteína en la que se describió por primera vez el plegamiento $(\beta/\alpha)_8$ y razón por la cual éste lleva su nombre. La TIM es responsable de catalizar la interconversión de D-gliceraldehído 3-fosfato (G3P) y dihidroxiacetona fosfato (DHAP), una reacción esencial en la glucólisis; probablemente por ello es que se ha encontrado en todos los organismos en los que se ha buscado en los tres dominios de la vida, casi siempre como dímero (fig. 3A). La TIM se localiza en citoplasma, cloroplastos y mitocondria y algunos organismos patógenos anaerobios son estrictamente dependientes de su actividad, por ello se ha usado como blanco en el diseño de drogas (Enríquez-Flores et al., 2011). Dada su participación en múltiples vías metabólicas, como son la glucólisis, la gluconeogénesis, la síntesis de triglicéridos y el ciclo de Calvin (D. Voet & J. G. Voet, 1995), es muy probable que la enzima se encuentre cercana físicamente a otras enzimas participantes en las vías, sin embargo, la mayoría de los estudios realizados sobre esta enzima en diferentes organismos han sido *in vitro*. Es muy poco lo que se puede inferir del contexto celular en el que normalmente está la proteína y se desconoce cuál es el papel que juegan los mecanismos de estabilización *in vivo*, como son las chaperonas, el fenómeno de “crowding” y la asociación en complejos proteicos mayores. Los estudios y resultados que a continuación se presentan se llevaron a cabo con proteínas aisladas y purificadas.

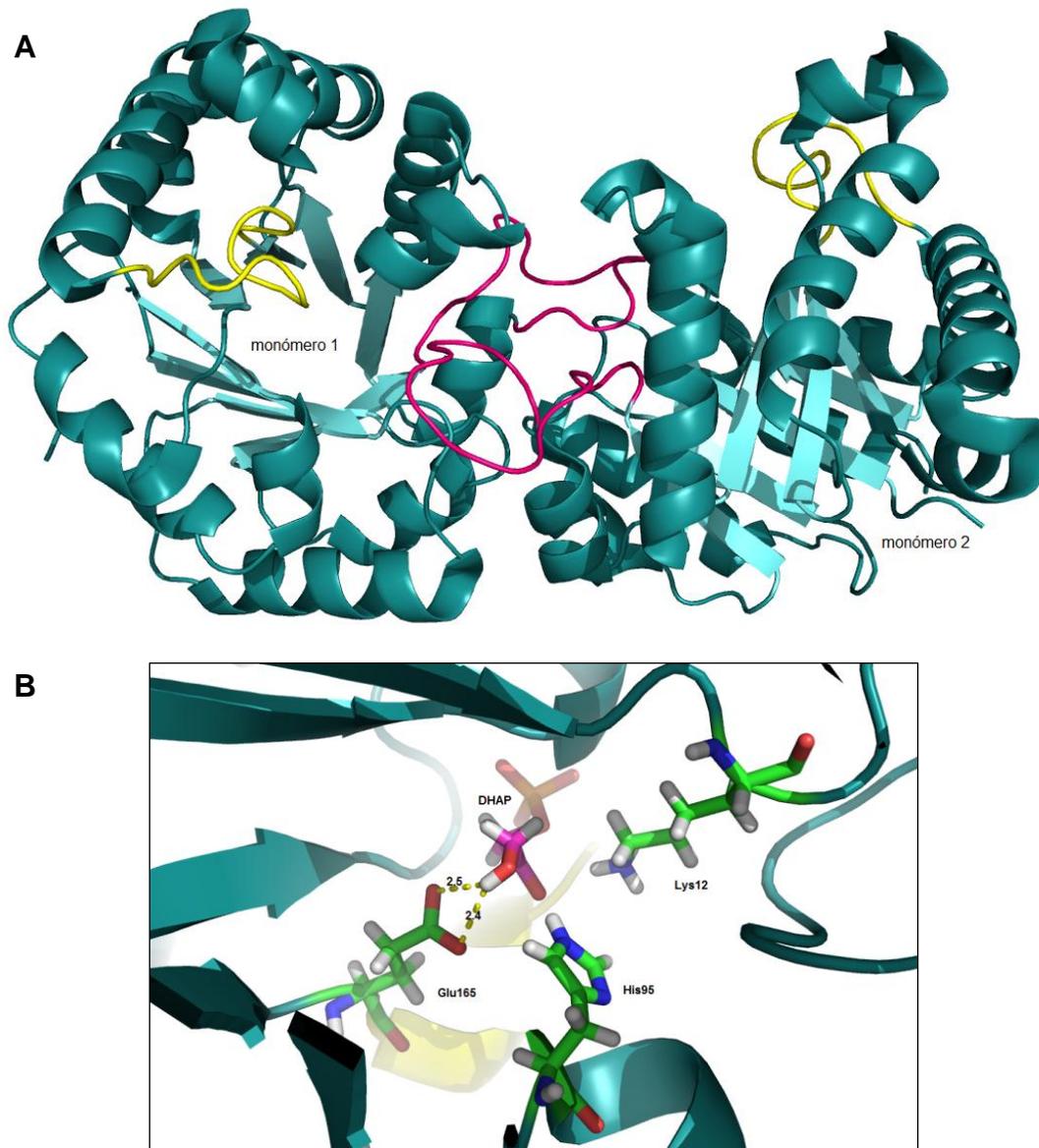


Figura 3. Características generales y del sitio activo de la TIM. A) Representación del dímero de la TIM de *Saccharomyces cerevisiae* (PDB ID: 1YPI). En amarillo el asa $\beta\alpha$ 6 y en rosa el asa $\beta\alpha$ 3. B) Arquitectura del sitio activo en presencia del sustrato DHAP para ScTIM, los números representan la distancia en Å para los puentes formados entre el carboxilo del glutámico catalítico hacia DHAP (PDB ID: 1NEY).

La reacción de isomerización ocurre sin cofactores y la enzima aumenta la reacción en una tasa de 10^9 en relación a la reacción no enzimática y la eficiencia catalítica (k_{cat}/K_m) es del orden de $10^9 \text{ M}^{-1}\text{s}^{-1}$ (Knowles & Albery, 1977). Este valor es cercano al estimado teórico para una reacción limitada por difusión, lo que

quiere decir que la cinética está solamente limitada por la difusión del sustrato al sitio activo de la enzima, es por ello que la TIM se clasifica como un “catalizador perfecto” (Albery & Knowles, 1976).

El sitio activo está sumamente conservado entre organismos y se conforma por la triada catalítica Lys12 (asa $\beta\alpha$ 1), His95 (asa $\beta\alpha$ 4) y Glu165 (asa $\beta\alpha$ 6) (numeración según la TIM de *Saccharomyces cerevisiae*) (fig. 3B). El perfil de energía para la formación de los complejos TIM-G3P, TIM-enediol (intermediario de la reacción) y TIM-DHAP fue caracterizado por Knowles, Albery y colaboradores hace ya casi 40 años (Albery & Knowles, 1976) y presenta los siguientes pasos en el mecanismo de reacción (fig. 4) 1) unión del sustrato, 2) formación del enediolato por deprotonación del carboxilato de Glu165, 3) formación del producto por reprotonación y 4) liberación del producto. Algunas de las características claves del sitio activo son la estabilización del intermediario por la lisina y la histidina, estabilización del fosfato del sustrato por puentes de hidrógeno con la cadena principal de la proteína y la exclusión del solvente. Para lograr este último rasgo, se ha observado que el asa 6 funciona como una tapa que presenta un movimiento de aproximadamente 7 Å al unir el ligando y ocultar al sitio activo, así como posicionar al Glu165 para la catálisis. La ausencia de dicha asa genera la liberación temprana del intermediario y la transformación del sustrato a metilglioxal fosfato, una molécula sumamente tóxica para las células (Pompliano et al., 1990).

Otra asa que participa de manera indirecta en la estabilización del sitio activo es el asa 3, que se asoma de su subunidad y se inserta en el monómero vecino entre las asas 1 y 4 en la interfase. Una red compleja de puentes de hidrógeno son los que se encargan de mantener a los aminoácidos catalíticos en la posición correcta (Maithal et al., 2002).

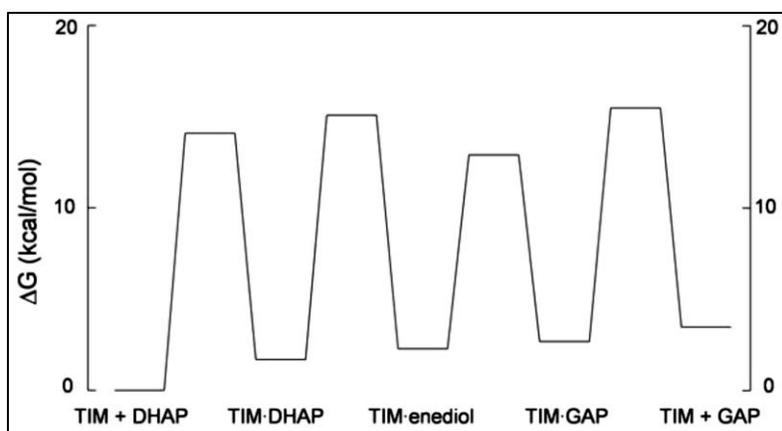


Figura 4. Diagrama de Energía para la reacción catalizada por la TIM. GAP: Glicelaldehído-3-fosfato, modificado de (Knowles & Albery, 1977).

Para estudiar el acto catalítico de la enzima se han estudiado extensivamente dos inhibidores competitivos que son análogos de los intermediarios de la reacción: fosfoglicolohidroxamato (PGH) y ácido 2-fosfoglicólico (PGA). A pesar de las similitudes en los parámetros catalíticos con los sustratos endógenos para todas las TIMs de diferentes organismos, se ha encontrado que su afinidad por el PGA varía ampliamente (tabla 1).

Tabla 1. Parámetros cinéticos para TIMs de diferentes organismos.

Especie	K_i PGA (mM)	K_m G3P (mM)	k_{cat} (min^{-1})
<i>Vibrio marinus</i> ^a	89	1.9 ± 0.2	$4.2 \pm 0.1 \times 10^5$
<i>Escherichia coli</i> ^a	6	1.03 ± 0.1	$5.4 \pm 0.1 \times 10^5$
<i>Giardia lamblia</i> ^b	0.043	0.53 ± 0.03	$2.9 \pm 0.2 \times 10^5$
<i>Saccharomyces cerevisiae</i> ^b	0.03	1.27	2.6×10^5
<i>Trypanosoma brucei</i> ^b	0.024	0.35	1×10^6
<i>Homo sapiens</i> ^c	0.0074	0.49	2.7×10^5
<i>Tenebrio molitor</i> ^d	27	0.84	2.6×10^5
<i>Oryctolagus cuniculus</i> ^d	14	0.64	2.2×10^5

^a(Álvarez, Zeelen et. al., 1998), ^b(López-Velázquez, 2004), ^c(Mainfroid, 1996), ^d(Knobeloch, 2010).

K_i : constante de inhibición, K_m : constante de Michaelis-Menten, k_{cat} : constante catalítica ($V_{max}/[E]$).

Fuera de las mutaciones que afectan el sitio activo o la interfase, el plegamiento es sumamente estable y permite una gran variabilidad en la secuencia, por ejemplo, en 2001 Silverman y colaboradores lograron generar mutantes puntuales en 142 sitios distintos de la TIM de levadura que resultaron en enzimas con actividad comparable a la silvestre (Silverman et al., 2001).

Hasta el momento (marzo de 2013) hay 140 estructuras cristalográficas de TIMs depositadas en el PDB, 117 de eucariontes, 20 de bacterias y 3 de arqueas, todas ellas pertenecientes a 28 especies distintas. La identidad de secuencia varía desde 98.4 % entre la TIM de humano (HsTIM) y la TIM de conejo (OcTIM) hasta 12 % entre PftTIM (TIM de *Plasmodium falciparum*) y PwTIM (TIM de *Pyrococcus woesei*), sin embargo, el alineamiento estructural revela que la topología tiene una alta conservación y a excepción de las TIMs de arqueas, que son tetraméricas, el resto presentan una estructura cuaternaria dimérica. No se han encontrado hasta ahora monómeros silvestres activos.

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. Existen varios trabajos en los que se ha intentado generar monómeros activos al realizar ingeniería de proteínas en la interfase del dímero para desestabilizarlo; desde mutaciones puntuales hasta cambiar la longitud y composición del asa 3, sin embargo, la eficiencia catalítica del monómero es aproximadamente 1000 veces menor que la de una enzima dimérica nativa (tabla 2) aun cuando el monómero está correctamente plegado (Schliebs et al., 1996).

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., 1996). La segunda hipótesis presupone que la asociación genera una estabilidad conformacional máxima y que los monómeros son termodinámicamente inestables (Mainfroid et al., 1996). Con respecto a esta última hipótesis, se ha observado que

aproximadamente el 70% de la estabilidad global de la enzima proviene del paso de dimerización en la vía de plegamiento (fig. 5).

Tabla 2. Parámetros catalíticos de TIMs monoméricas. Tomado de (Schliebs et al., 1997)

		TbTIM	MonoTIM	RE-TIM
k_{cat}	(G3P)	3.7 x 10 ⁵	3.1 x 10 ²	1.3 x 10 ²
K_m	(G3P)	0.25 ± 0.05	4.1 ± 0.6	1.4 ± 0.2
k_{cat}/K_m	(G3P)	1.5 x 10 ⁶	75	93
k_{cat}	(DHAP)	6.5 x 10 ⁴	12	28
K_m	(DHAP)	1.2 ± 0.1	12.2 ± 1.9	17.2 ± 5.8
k_{cat}/K_m	(DHAP)	5.4 x 10 ⁴	1	1.6
K_i	(PGH)	0.008 ± 0.001	0.5 ± 0.07	1.75 ± 0.3
K_i	(PGA)	0.026 ± 0.007	0.052 ± 0.007	0.034 ± 0.005

TbTIM: TIM de *Trypanosoma brucei*, monoTIM: mutante monomérica de TbTIM con reemplazo en el asa 3, RE-TIM: doble mutante monomérica de TbTIM en el asa 3.

k_{cat} en min⁻¹, K_m y K_i en mM, k_{cat}/K_m en min⁻¹mM⁻¹

Por último, a pesar del alto grado de conservación estructural entre las enzimas de las diferentes especies, las vías de plegamiento que llevan al estado nativo son muy diversas. Para las TIMs podemos encontrar desde el comportamiento más sencillo de dos estados; en los que un monómero desplegado pasa a un dímero plegado sin intermediarios al equilibrio (TIMs de conejo y humano, entre otras); hasta vías complejas con intermediarios monoméricos y/o diméricos, como es el caso de las TIMs de los tripanosomátidos *T. cruzi* y *T. brucei*. La tabla 3 presenta un resumen de la variedad de vías que se han encontrado.

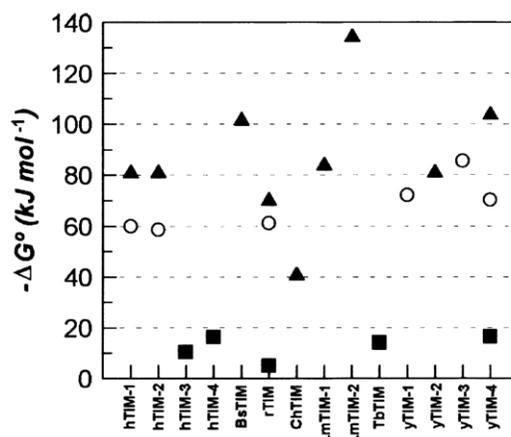


Figura 5. Contribución de la asociación a la estabilidad global de las TIMs. Tomada de (Nájera et al., 2003). $\Delta G^{\circ}_{\text{plegamiento}}$ (■), $\Delta G^{\circ}_{\text{asociación}}$ (○), $\Delta G^{\circ}_{\text{total}}$ (▲). hTIM-1 y -2: TIM de humano, hTIM-3 y -4: mutantes monoméricas, Bs: *Bacillus stearothermophilus*, rTIM: TIM de conejo, ChTIM: TIM de pollo, Lm: *Leishmania mexicana*, LmTIM-2: mutante puntual, TbTIM: mutante monomérica de TIM de *Trypanosoma brucei*, yTIM: TIM de levadura.

Tabla 3. Variedad en las vías de plegamiento para TIMs de diferentes especies. Modificado de (Pardo-Ávila, 2011).

Vía de plegamiento	TIMs que la presentan
$2D \leftrightarrow N_2$	BsTIM, OcTIM, GgTIM, LmTIM, HsTIM
$2D \leftrightarrow 2M \leftrightarrow N_2$	ScTIM, EhTIM
$2D \leftrightarrow A \leftrightarrow N_2$	TtTIM
$2D \leftrightarrow A \leftrightarrow N_2^* \leftrightarrow N_2$	PfTIM
$2D \leftrightarrow 2M \leftrightarrow N_2^* \leftrightarrow N_2$	TbTIM, TcTIM

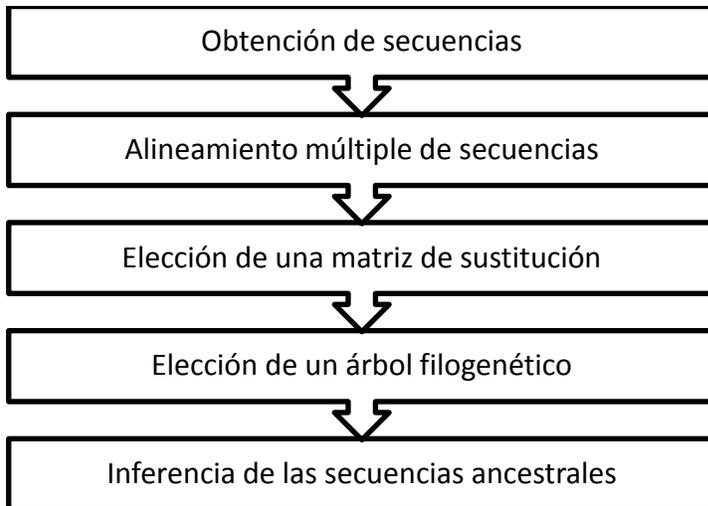
D: estado desplegado, M: monómero, N_2 : dímero nativo, N_2^* : intermediario dimérico, A: agregados. Bs: *Bacillus subtilis*, Oc: *Oryctolagus cuniculus*, Gg: *Gallus gallus*, Lm: *Leishmania mexicana*, Hs: *Homo sapiens*, Sc: *Saccharomyces cerevisiae*, Eh: *Entamoeba histolytica*, Tt: *Thermotoga maritima*, Pf: *Plasmodium falciparum*, Tb: *Trypanosoma brucei* y Tc: *Trypanosoma cruzi*.

El estudio comparativo de proteínas ha sido útil para entender mejor la relación entre la secuencia y las propiedades fisicoquímicas, sin embargo, este enfoque no ofrece una explicación integrativa y evolutiva al problema. Para dilucidar cómo se generaron las propiedades catalíticas, el estado de asociación y la estabilidad de

estas enzimas, en esta tesis examinamos la historia molecular de la TIM mediante la reconstrucción de proteínas ancestrales y su posterior caracterización fisicoquímica.

ANTECEDENTES

En un trabajo previo en el Laboratorio de Fisicoquímica e Ingeniería de Proteínas de la Facultad de Medicina de la UNAM, se realizó la reconstrucción ancestral para las TIMs de eucariontes, a continuación se presenta un resumen general de los resultados obtenidos por Fátima Pardo en su tesis de licenciatura (Pardo-Ávila, 2011):



Se obtuvieron 44 secuencias de nucleótidos de la base de datos del NCBI correspondientes solamente a organismos del dominio *Eukarya* y se tradujeron al código de aminoácidos en el programa SE-AL. Este mismo programa sirvió para refinar un alineamiento inicial de las secuencias en el programa Clustal-X. A partir del alineamiento se calculó el índice de conservación para cada residuo en la proteína y se observó que las regiones cercanas a los aminoácidos catalíticos tienen los mayores índices de conservación.

Para poder inferir la secuencia de un ancestro a partir de las secuencias extantes, es necesario primero calcular el tiempo evolutivo transcurrido, el cual depende de la tasa de sustituciones por residuo. Para el caso de los aminoácidos, existen matrices empíricas que describen las tasas de cambio de un residuo a otro y es importante analizar cuál es la matriz que se ajusta mejor a los datos. A partir del análisis de los datos con el programa Prottest, se eligió la matriz de sustitución LG, porque mostró una clara mejora en la verosimilitud de la reconstrucción para los datos. Esta matriz fue creada a partir de 3912 alineamientos de la base de

datos Pfam, y contiene la información de las probabilidades de sustitución de aminoácidos al analizar ~50,000 secuencias y ~6.5 millones de residuos.

Normalmente, cuando se realiza una reconstrucción ancestral de proteínas se infiere un árbol filogenético con los datos. El árbol filogenético es una representación de las relaciones evolutivas que existen entre los organismos extantes. Las secuencias extantes están localizadas en los extremos de las ramas y la distancia entre ellas representa la distancia evolutiva entre organismos; por ejemplo, secuencias parecidas estarán en ramas adyacentes y se unirán en una rama común debajo de ellas mientras que secuencias distantes tendrán ramas alejadas. Cada unión de dos ramas (nodo) corresponde al último ancestro común de las mismas y la longitud de cada una está determinada por el número de cambios por sitio (distancia evolutiva). El objetivo del análisis filogenético es descubrir el arreglo de las ramas, así como la longitud de éstas que mejor represente la relación entre las secuencias (Page & Holmes, 1998).

Debido a que existen eventos como transferencia horizontal de genes, la historia evolutiva de una proteína no siempre representa la historia evolutiva del organismo. Este es el caso de la TIM. Dado que no es un buen marcador evolutivo no se puede usar para inferir un árbol filogenético de especies, sin embargo, como es parte del metabolismo central, se supuso que la historia evolutiva de la TIM es la misma que la de los organismos que la contienen. Por ello, se trabajó con el árbol propuesto por Ciccarelli en 2006 (Ciccarelli et al., 2006), el cual fue reconstruido a partir de 36 genes que no presentan transferencia horizontal de 191 organismos. Con ayuda del programa MacClade se acoplaron la topología del árbol y los alineamientos previamente realizados (fig. 6).

La reconstrucción de estados ancestrales fue realizada por métodos de máxima verosimilitud. Este modelo probabilístico asigna un valor de verosimilitud a cada residuo, en cada posición, en una secuencia ancestral dependiendo de los datos observados y la matriz de sustitución elegida. La proteína ancestral fue reconstruida a partir de los aminoácidos que tienen mayor probabilidad de reconstrucción por sitio.

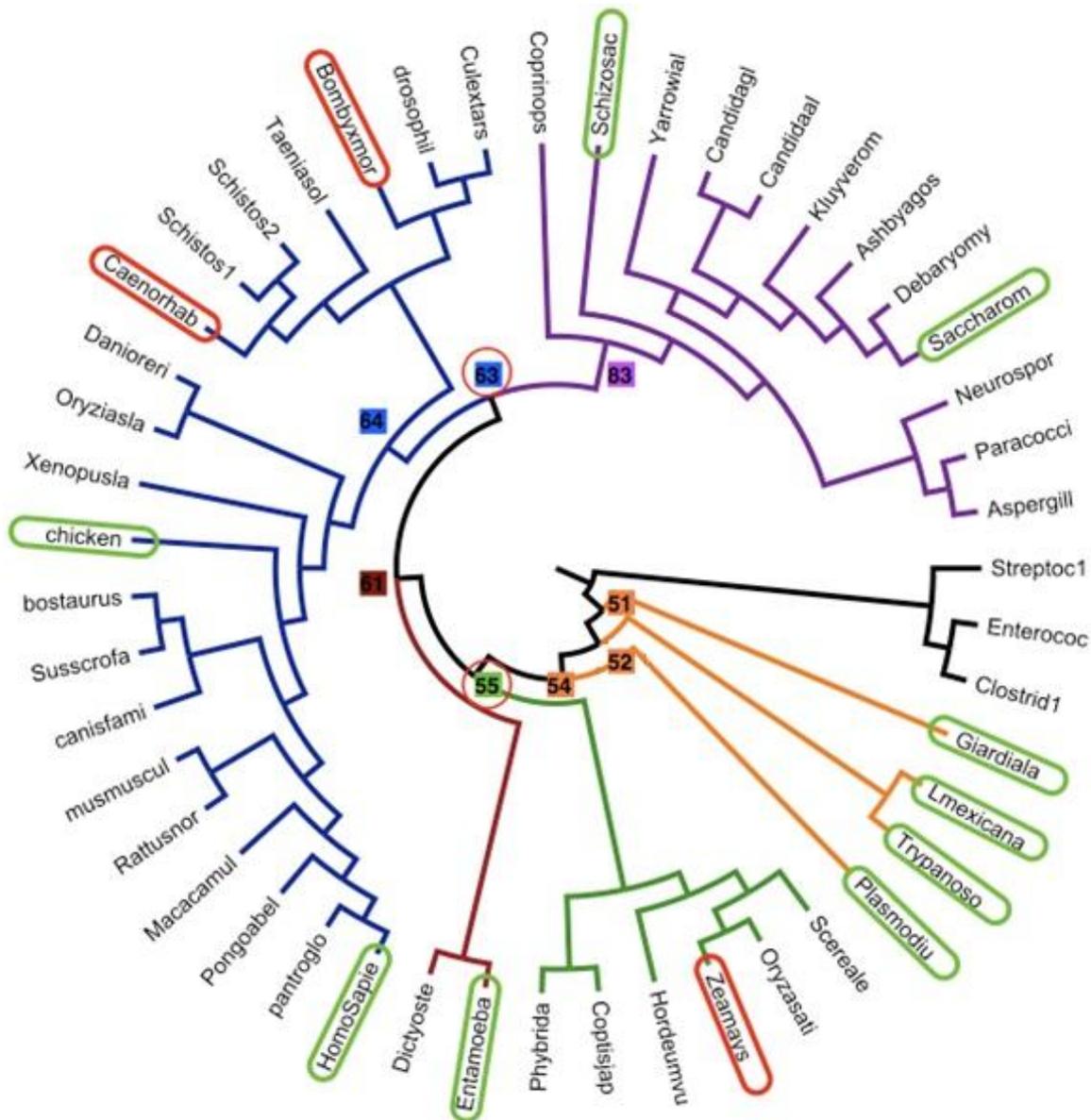


Figura 6. Árbol filogenético de eucariontes usado en la reconstrucción ancestral de la *triosafosfato isomerasa*. Los números representan secuencias ancestrales, en círculos verdes están los organismos en los que la TIM se ha estudiado y en círculos rojos secuencias que se mandaron a sintetizar para su estudio en nuestro grupo de trabajo. Los ancestros estudiados en la presente tesis son los números 63 y 55. (Modificado de Pardo-Ávila, 2011).

Para cada nodo en el árbol se obtiene una secuencia ancestral hipotética. En la presente tesis se eligió “revivir” y estudiar a dos ancestros en el laboratorio, cuyas secuencias tienen un alto porcentaje de sitios con alta probabilidad de reconstrucción (fig. 7). El primero corresponde al ancestro común más reciente de animales y hongos (nodo número 63) con una edad de 1.37 miles de millones de años (Perez-Jimenez et al., 2011) y el segundo corresponde al ancestro común que incluye además a las plantas (número 55) (fig. 6).

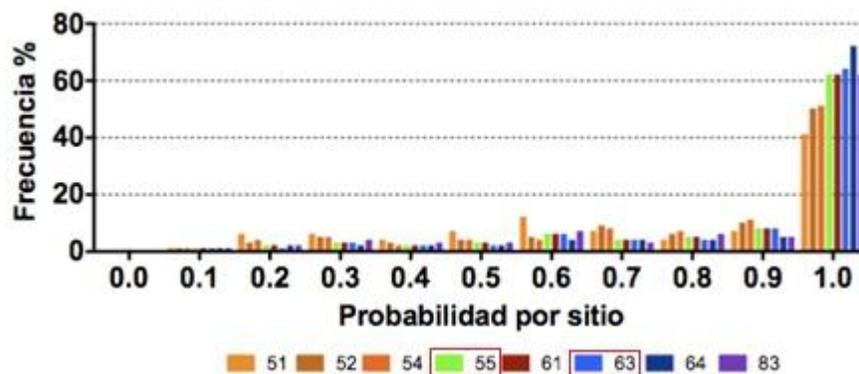


Figura 7. Distribución de las probabilidades de reconstrucción por sitio. Los números corresponden a los nodos en la figura 6. (modificado de Pardo-Ávila, 2011)

La comparación entre las secuencias de los nodos reveló que tienen 96% de identidad entre ellos, que equivale a 11 aminoácidos diferentes, 8 de los cuales son similares (con respecto a sus características químicas). En la figura 8 se presenta el alineamiento entre ellas y con la TIM de *Tenebrio molitor* (TmTIM, comúnmente conocido como gusano de la harina) que corresponde a la que presenta mayor similitud de secuencia (72% para TIM63 y 69% para TIM55) y además está disponible su estructura cristalográfica.

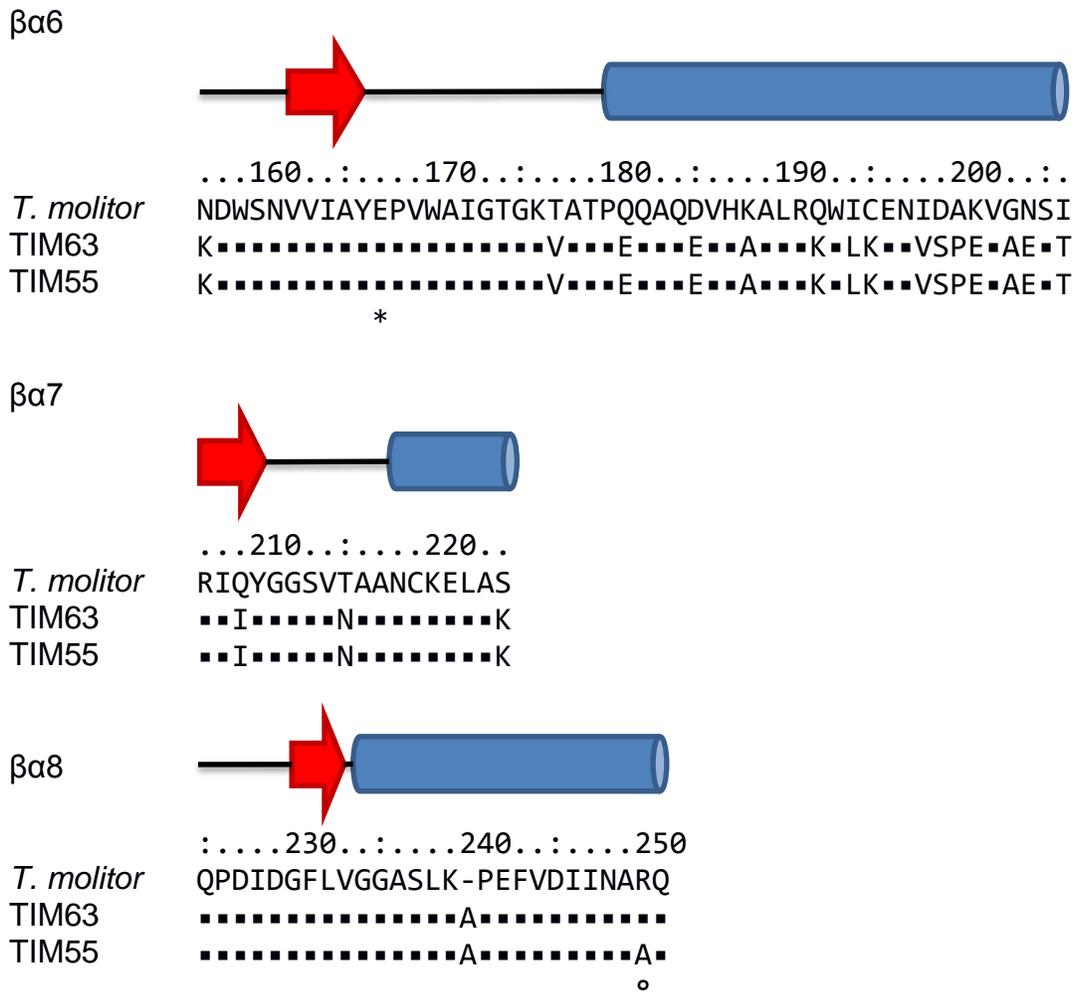


Figura 8. Alineamiento de TIMs ancestrales y la TIM de tenebrio molitor. La numeración corresponde a la de los ancestros, el alineamiento está dividido en ocho motivos $\beta\alpha$ de acuerdo a la estructura de TmTIM, las flechas rojas representan las hebras betas y los cilindros azules las hélices alfa. Con un * se marcan los aminoácidos catalíticos. ■ indica que el residuo en esa posición es idéntico al que está en la TIM de *T. molitor* y con ° están marcadas las 11 diferencias entre las secuencias de TIMs ancestrales.

La caracterización de proteínas ancestrales en los estudios realizados a la fecha se basa principalmente en la determinación de la temperatura media de desnaturalización, como parámetro que describe la estabilidad, o en el estudio de la función endógena de la proteína y su comparación con las proteínas extantes. Sin embargo, el estudio a profundidad de proteínas ancestrales reconstruidas puede arrojar mucha más información que ésta, por ejemplo, sobre las determinantes estructurales que llevan a un tipo de plegamiento o las

características fisicoquímicas que estabilizan una conformación específica. Este trabajo es el primero en el que se realiza una caracterización fisicoquímica y termodinámica de una proteína ancestral.

HIPÓTESIS

La caracterización fisicoquímica y funcional de las proteínas reconstruidas es una herramienta eficaz para validar la calidad de la reconstrucción ancestral de secuencias.

OBJETIVOS

Objetivo general

Caracterizar las propiedades catalíticas, estructurales y de estabilidad termodinámica de dos reconstrucciones ancestrales de la *triosafosfato isomerasa* de eucariontes.

Objetivos particulares

- Subclonar, expresar y purificar TIM63 y TIM55.
- Determinar las propiedades espectroscópicas (fluorescencia y dicroísmo circular) de las enzimas ancestrales.
- Estudiar la termoestabilidad de los ancestros.
- Cuantificar la actividad enzimática de ambas enzimas y determinar sus parámetros catalíticos.
- Caracterizar la asociación monómero-monómero de las enzimas y determinar los parámetros termodinámicos que rigen el proceso.

MATERIALES Y MÉTODOS

Clonación y expresión

Los genes correspondientes a las dos TIMs ancestrales de eucariontes (TIM63 y TIM55) se mandaron sintetizar en el plásmido pBluescript II SK (-) (Epoch life sciences). Posteriormente, cada gen se subclonó en el vector de expresión pET28 b (+), el cual coloca una etiqueta de histidinas en el residuo amino terminal de las proteínas y contiene un sitio de corte para la trombina que permite la remoción de la secuencia de histidinas después de la purificación de las proteínas, dejando solamente los aminoácidos Gly - Ser - His, que nos son parte de la secuencia original de la proteína.

Los genes de TIM63 y TIM55 se digirieron con las enzimas NdeI y XhoI o NdeI y BamHI (New England Biolabs® Inc.), respectivamente, durante 6 h en una reacción de 50 ul (10 ul de plásmido, 5 ul de amortiguador número 4 10X (New England Biolabs® Inc.), 5 ul de una solución de albúmina sérica bovina (BSA) 10X, 3 ul NdeI, 3 ul BamHI o XhoI y 24 ul de agua). El plásmido pET28 se digirió durante 2.5 h en una reacción de 50 ul (20 ul de vector, 5 ul de amortiguador número 4, 5 ul de BSA 10X, 2 ul NdeI, 2 ul BamHI o XhoI y 16 ul de agua). Los productos se purificaron a partir de gel por kit de Qiagen®. La relación inserto: vector en las reacciones de ligación fue de 5:1 para TIM63 y 15:1 para TIM55 en un volumen de 10 ul; se incubaron durante 2 días a 16 °C. La nueva construcción se mandó secuenciar a la compañía Laragen Inc. Todas las purificaciones de DNA se llevaron a cabo en células quimiocompetentes DH5α. Las construcciones se transformaron en células quimiocompetentes BL21plis. A partir de una colonia aislada se inocularon 10 ml de medio LB + kanamicina y se incubaron toda la noche a 37 °C y 250 rpm. A partir de 1 ml de los precultivos se inocularon 250 ml de medio LB + kanamicina y se crecieron a 37 °C y 250 rpm hasta obtener una DO a 600 nm de ~ 0.6 - 0.8. Estos cultivos se indujeron con 1 mM de IPTG y se incubaron toda la noche (~13 h) a 37 °C y 250 rpm.

Purificación de las proteínas ancestrales

Columna de afinidad

Las células inducidas se cosecharon mediante centrifugación a 8 000 rpm durante 20 min; se resuspendieron en un amortiguador de fosfatos (35 mM NaH_2PO_4 , 300 mM NaCl, 5 mM imidazol, pH 8) y se lisaron por sonicación (17 pulsos de 45 segundos con una amplitud de 25%). Se centrifugaron durante 15 min a 15000 rpm y se recuperó el sobrenadante, que se filtró para pasarlo por una columna de afinidad a las etiquetas de histidinas (Protino[®] NiTED) equilibrada previamente con el amortiguador en el que se resuspendieron las células. La columna está conectada a un equipo Äkta FPLC. La muestra se inyectó en el equipo y posteriormente se aplicó un gradiente de 0 a 60% del amortiguador de fosfatos con 500 mM de imidazol en un volumen de 200 ml. Las proteínas eluyen a partir de 80 mM de sal (fig. 9A). Las fracciones correspondientes al pico de absorbancia se concentraron en tubos Centricon (Millipore[®]) hasta un volumen de aproximadamente 4 ml, después se agregaron 20 ml de amortiguador de corte (Tris 50 mM, pH 8) y se centrifugó a 5000 rpm por 20 minutos. Este proceso se repitió 5 veces y se realizó con el fin de cambiar de amortiguador a la proteína y poder pasar al proceso del corte de la etiqueta de histidinas.

Corte con trombina

Una vez que las proteínas ya estaban disueltas en el amortiguador de corte, se le agregó CaCl_2 para llegar a una concentración de 10 mM y se centrifugaron por 15 minutos a 15 000 rpm. Se utilizó el kit CleanCleave[™] de Sigma[®] con la trombina fusionada a una resina de agarosa. La reacción de corte se llevó a cabo en un volumen final de 10 ml, con una concentración de proteína de aproximadamente 3 mg/ml, con 1 ml de resina, 1 ml de amortiguador para el corte 10X. La reacción duró 4 horas a temperatura ambiente, después de este tiempo se recuperó a la proteína en el sobrenadante de una centrifugación a 2 500 rpm por 5 min. Este sobrenadante se dializó primero contra un amortiguador de Tris 50 mM sin CaCl_2 y posteriormente contra el amortiguador de fosfatos con 5 mM de

imidazol, pH 8, para pasarlo por la columna de afinidad de nuevo y así eliminar a la población de enzima recombinante que no fue cortada. La proteína sin etiqueta de histidinas (TIM63sc o TIM55sc) eluye en el frente de elución (fig. 9B), éste se recuperó y se dializó contra un amortiguador de trietanolamina (TEA) 10 mM, pH 7.6 y se concentró en tubos centríficos (Millipore[®]) para posteriormente pasarlo por una columna de intercambio aniónico Mono Q[®] 10/10 (Pharmacia Biotech) previamente equilibrada con el mismo amortiguador. Se corrió un gradiente de 0 a 50% del amortiguador TEA 10 mM, NaCl 1 M, pH 7.6 por la columna; la proteína eluye a una concentración de NaCl de aproximadamente 80 mM. Por último se dializó a la proteína contra el amortiguador necesario para cada experimento.

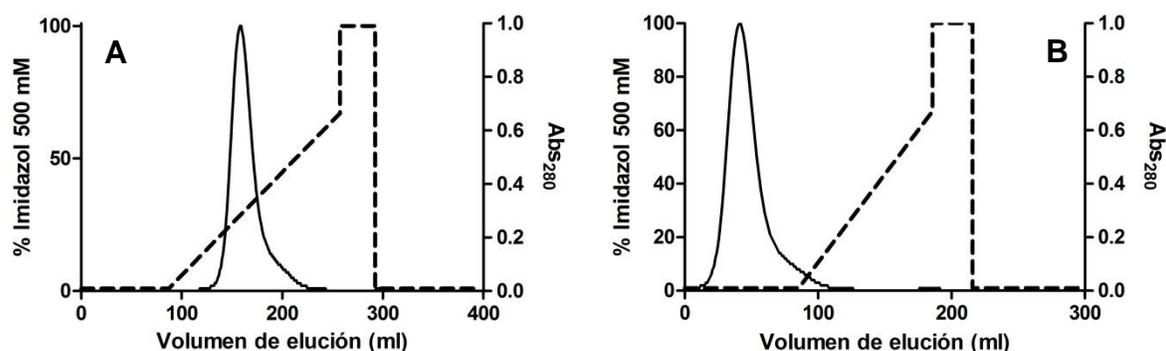


Figura 9. Purificación de proteínas recombinantes. Elución en la columna de afinidad de níquel de proteínas recombinantes A) con etiqueta de histidinas y B) sin etiqueta después del corte con trombina. La línea punteada corresponde al gradiente de sal aplicado, eje izquierdo; la línea continua representa la absorbancia a 280 nm normalizada de las fracciones eluidas, eje derecho.

Propiedades espectroscópicas

Para comparar los espectros de las TIMs ancestrales con los de alguna TIM extante, también se determinaron las propiedades de la TIM de *Saccharomyces cerevisiae* (ScTIM) que se compró pura y liofilizada de Sigma[®] y se diluyó en el amortiguador de TEA 10mM, pH 7.6, para llegar a las concentraciones necesarias de cada experimento.

Dicroísmo circular (DC)

El dicroísmo circular es una técnica espectroscópica en la que se aprovecha la naturaleza quiral de los carbonos en la cadena polipeptídica de las proteínas y el arreglo específico de los enlaces formados en estructuras secundarias para absorber luz polarizada circularmente. Este haz de luz cuenta con dos componentes circulares que pueden ser absorbidos de manera diferencial por moléculas ópticamente activas. Esta técnica es utilizada principalmente en dos rangos de longitudes de onda en experimentos con proteínas: UV lejano (180-260 nm) y UV cercano (>250). La señal de DC en el UV lejano se debe principalmente a la absorción de los aminoácidos aromáticos de las proteínas como son el triptófano, las tirosinas y la fenilalanina, así como los puentes disulfuro presentes. Por otra parte, cada estructura secundaria absorbe de manera específica luz polarizada en el UV lejano. Con ayuda de algoritmos es posible calcular la contribución de hélices α o hebras β en un espectro, además de que ciertos plegamientos cuentan con patrones característicos. Los datos crudos están en unidades de elipticidad de la muestra, para poder compararlos es necesario convertirlos a una escala normalizada: la elipticidad molar. Este parámetro está normalizado por la concentración de proteína, por el ancho de la celda, por la masa molecular y el número de residuos.

Los espectros de dicroísmo circular de las proteínas se tomaron en un equipo ChirascanTM de Applied Photophysics a una concentración de proteína de 250 $\mu\text{g/ml}$ en un amortiguador de fosfatos 10 mM, pH 8 en una celda de cuarzo de 0.1 mm de paso de luz. Se tomaron tres espectros para cada proteína con un ancho de banda de 0.5 nm desde 185 a 250 nm, el espectro presentado es un promedio.

Fluorescencia

Para poder observar fluorescencia es necesario primero excitar electrones para que pasen a un nivel energético elevado por medio de la absorción de un fotón a una longitud de onda específica. Una vez en el estado energético elevado, este electrón necesita regresar a su nivel basal de energía y lo hace liberando un fotón de una longitud de onda mayor al fotón que absorbió en un inicio (menor energía).

En las proteínas se puede observar el fenómeno de fluorescencia, los aminoácidos aromáticos, principalmente los triptófanos y las tirosinas, son los responsables de la fluorescencia intrínseca. Normalmente lo que se mide es la intensidad de la longitud de onda de emisión, entre 300 y 350 nm. Una longitud de onda de 280 nm excita tirosinas y triptófanos, mientras que una de 295 nm excita principalmente a los triptófanos, el análisis comparativo de ambas emisiones puede sugerir cambios terciarios específicos en la proteína. La emisión de fluorescencia de los triptófanos depende mucho del ambiente en el que éstos se encuentren. Por ejemplo, cuando uno de estos residuos se encuentra en el núcleo hidrofóbico de la proteína, la emisión estará corrida hacia el azul (más cercano a 300 nm); en el proceso de desplegamiento de proteínas seguidas por la emisión de fluorescencia de triptófanos, el espectro se corre hacia el rojo (más cercano a 350 nm) cuando los triptófanos pasan a un ambiente acuoso.

Las proteínas se excitaron a longitudes de onda de 280 y 295 nm en un espectrofluorómetro ISS PC1 en celdas de cuarzo de 0.5 x 0.5 cm y se midió la emisión en el rango de 310 a 400 nm. La concentración de proteína fue de 50 µg/ml.

En el estudio de proteínas también se puede utilizar la fluorescencia de manera indirecta, por ejemplo, al usar moléculas fluorescentes que tengan afinidad por determinadas regiones proteicas. Tal es el caso del 1-anilino-8-naftalenosulfonato (ANS), una molécula con anillos aromáticos que fluoresce en ambientes apolares; uno de sus usos es comprobar la exposición de regiones hidrofóbicas en una proteína.

Se añadió ANS 100 µM a la muestra de proteína, se excitó a 360 nm y se midió la emisión de fluorescencia en el rango de 400-600 nm. Las concentraciones de proteína variaron desde 25 a 500 µg/ml para las TIMs ancestrales y ScTIM y desde 1 a 25 µg/ml para BSA, que se utilizó como control positivo.

Estabilidad térmica

Una manera de comprobar la estabilidad de una proteína es ponerla en contacto con algún agente desnaturante y evaluar hasta qué punto ésta puede mantener su estructura plegada. A diferencia de los agentes desnaturantes químicos que mueven el equilibrio del proceso de desnaturación, la temperatura afecta directamente a las interacciones proteicas, por ello se eligió usarla como agente desnaturante en este trabajo. Al elevar la temperatura paulatinamente, se aumenta la energía cinética de los átomos en la proteína hasta generar la ruptura de enlaces estabilizantes y provocar el desplegamiento de la estructura nativa hasta llegar a un estado desnaturado.

Rampa de temperatura

Los experimentos de desnaturación por temperatura se realizaron en el ChirascanTM. Se desnaturó a las proteínas elevando la temperatura de la celda con una tasa de 1 °C/min desde 20 hasta 90 °C. La concentración de la muestra fue de 250 µg/ml en un amortiguador de TEA 10 mM, pH 7.6. La temperatura de la muestra fue monitoreada por un termopar localizado en el interior de la celda. Se determinó el grado de desnaturación siguiendo la señal de dicroísmo circular a 222 nm. La temperatura reportada es la que se midió directamente dentro de la celda.

Suponiendo un proceso simple de dos estados, el valor de la variable observada en cada punto del proceso es en realidad la resultante de la fracción de proteína que se encuentra en estado nativo y la que se encuentra en estado desnaturado (1):

$$y_{obs} = y_N * f_N + y_D * f_D \quad (1)$$

Donde y_{obs} es la elipticidad molar en cada punto, y_N es la elipticidad molar observada para el estado nativo, f_N es la fracción nativa, y_D es la elipticidad molar observada para el estado desnaturado y f_D es la fracción en estado desnaturado.

Dado que

$$f_N + f_D = 1 \quad (2)$$

es posible poner la ecuación 1 solamente en términos de f_D y despejar para obtener el valor de esta variable:

$$f_D = \frac{y_{obs} - y_N}{y_D - y_N} \quad (3)$$

Los datos obtenidos de esta manera se ajustaron a la ecuación de Boltzmann (4), que describe una curva sigmoide con pendiente para obtener la temperatura media de desnaturalización aparente ($T_{m_{app}}$):

$$f_D = y_{min} \frac{y_{max} - y_{min}}{1 + \exp \frac{T_{m_{app}} - T}{m}} \quad (4)$$

siendo $y_{min} f_D=0$, $y_{max} f_D=1$ y m la pendiente en la transición.

No se utilizó la ecuación de Gibbs-Helmholtz ya que el desplegamiento de TIM55 y TIM63 es irreversible.

Calorimetría diferencial de barrido (DSC)

La desnaturalización seguida por técnicas espectroscópicas arroja información indirecta sobre los cambios energéticos en el proceso de plegamiento y desplegamiento de las proteínas, ya que los parámetros termodinámicos calculados se basan en un modelo.

La única técnica que mide directamente la energética del proceso sin importar el patrón de plegamiento es la calorimetría diferencial de barrido (DSC). Los calorímetros de barrido se componen de dos celdas, una de muestra (proteína) y otra de referencia (amortiguador). El aparato calienta ambas celdas a una velocidad constante, de manera que la diferencia de temperatura entre ambas sea igual a cero. Dado que la energía requerida para calentar a la proteína no es equivalente a la requerida para calentar el amortiguador, las celdas requieren diferente cantidad de energía para elevar su temperatura, por ello, el calorímetro aplica una potencia determinada a cada celda y mide la diferencia. La división de

la potencia entre la velocidad de barrido da como resultado el valor de la capacidad calorífica (C_p), es decir, la capacidad de almacenar o absorber energía de la muestra. Dado que el valor de C_p varía en el proceso, es necesario calcular este parámetro para el estado nativo (C_p^N) y desnaturalizado (C_p^D). Esto se conoce como la línea base química del proceso.

Normalmente se realiza una corrida calorimétrica con amortiguador en ambas celdas para poder calcular la línea base instrumental. Las trazas calorimétricas se normalizan a partir de la concentración de proteína y restando la línea base instrumental y química del proceso para obtener así el C_p en exceso.

Los experimentos de DSC con las TIMs ancestrales se llevaron a cabo en el VP-DSC del laboratorio de termofísica (FQ-UNAM). Se hicieron experimentos variando la concentración de proteína (0.4, 0.6, 1 y 3 mg/ml), así como la velocidad de barrido (30, 60, 90, 120, 150 y 180 °C/hora). El amortiguador en todos los casos fue TEA 10 mM, NaCl 100 mM, pH 7.6, previamente filtrado. Se corrieron líneas base (amortiguador-amortiguador) a la velocidad de barrido del experimento en curso. El rango de temperaturas usado fue de 20°C a 90 °C. Las trazas se analizaron en el software de Origin 7 DSC, en donde primero se restó la línea base del amortiguador (línea base química), se corrigió el tiempo de respuesta del equipo y los datos se normalizaron por la concentración de proteína.

Para obtener el ΔH total de proceso (ΔH_{cal}) se calculó el área bajo las curvas al que se les resto la línea base química y por lo mismo representan la capacidad calorífica en exceso (C_p^{exc}):

$$\Delta H_{cal} = \int C_p^{exc} dt \quad (5)$$

Cada traza se ajustó a un modelo de dos estados irreversible, descrito con una cinética de primer orden:



en donde la proteína nativa N pasa al estado final F, que puede ser desplegado y/o agregado, y k es la constante de velocidad que describe al proceso. Este modelo se conoce como el de Lumry-Eyring (J M Sanchez-Ruiz, 1992).

La ecuación 7 supone ese modelo y describe el comportamiento de los datos:

$$C_p^{AP} = C_p^{Pre} + (C_p^{Post} - C_p^{Pre})(1 - x_N) - \Delta H \frac{dx_N}{dT} \quad (7)$$

donde x_N es la fracción molar del estado nativo, ΔH es la entalpía de desnaturalización y C_p^{Pre} y C_p^{Post} son las líneas base de pre y post- transición, respectivamente (Costas et al., 2009).

A partir del ajuste de las trazas a diferentes velocidades de barrido, se obtuvieron los valores de las constantes de velocidad del proceso (k) para poder realizar los gráficos de Arrhenius:

$$\ln k = \ln km - \frac{1}{T_m} - \frac{E_{act}}{R} * \frac{1}{T} \quad (8)$$

En donde km está dada por

$$km = VB * \frac{E_{act}}{T_m} \quad (9)$$

VB es la velocidad de barrido en grados por minuto, E_{act} es la energía de activación, R la constante de los gases y T_m la temperatura media de desnaturalización.

Catálisis

Determinación de la actividad específica

En la reacción que cataliza la TIM es posible medir la cantidad de G3P que se convierte a DHAP mediante un ensayo acoplado a otra reacción enzimática catalizada por la glicerol-3-fosfato-deshidrogenasa (GDPH). La GDPH utiliza el cofactor NADH, cuya desaparición puede ser monitoreada siguiendo la caída en absorbancia a 340 nm:



Las reacciones se llevan a cabo en un amortiguador de TEA 100 mM, EDTA 10 mM, pH 7.6, en un volumen final de 700 μl por celda de reacción con una concentración final de enzima de 30 ng/ml y 150 ng para TIM63 y TIM55, respectivamente, 2 mM de G3P, 2 μM de NADH y 7 $\mu\text{g/ml}$ de GDPH (todos los reactivos son de la marca Sigma[®]). Los experimentos se realizaron a 25 °C en un equipo espectrofotómetro Beckman Coulter DU-7500. Para llegar a la concentración final de TIM, en la celda se hizo una dilución previa justo antes de la medición, de la cual se tomaron 10 μl y se diluyeron directamente en la celda.

Determinación de los parámetros cinéticos de la reacción

Las reacciones se llevaron a cabo en un amortiguador de TEA 100 mM, EDTA 10 mM, pH 7.6, en un volumen final de 700 μl por celda de reacción con una concentración final de enzima de 5 ng/ml de TIM63sc, 2 μM de NADH y 7 $\mu\text{g/ml}$ de GDPH. Las concentraciones de G3P varían entre 0.1 y 10 mM. A partir de la pendiente (m) se calculó la actividad específica de la enzima (a. e.):

La actividad se graficó contra la concentración de sustrato y se ajustó a una hipérbola de Michaelis-Menten en el programa Prisma para obtener los parámetros de K_m y V_{max} .

Estabilidad a la dilución

Cromatografía de exclusión molecular

Se calibró una columna analítica Superdex 75 (Pharmacia Amersham) con proteínas de peso molecular conocido según el protocolo de Uversky (Uversky, 1993) para estimar el radio de Stokes de TIM63 y TIM55 a diferentes concentraciones para observar el efecto de la dilución sobre la dimerización de la enzima. La columna se equilibró con un amortiguador de TEA 10 mM, EDTA 1 mM, pH 7.6 con 1.5 volúmenes de columna previo a la corrida, la cual se llevó a cabo a un flujo de 1 ml/min.

Inactivación por dilución

La proteína se diluyó en el amortiguador de TEA 10 mM, NaCl 100 mM, pH7.6 para obtener concentraciones de 100 µg/ml a 9 mg/ml. Estas proteínas se incubaron 24 horas a 25 °C para posteriormente medir su actividad.

La dilución de la enzima propicia la disociación de la misma y genera una fracción activa f_N (dímero) y una inactiva f_D (monómero). Las concentraciones de cada población están determinadas por la constante de disociación al equilibrio. Siguiendo la misma lógica que en las ecuaciones 1-3, se puede obtener la fracción activa en cada punto y normalizar los datos. En este caso se supuso que y_N es el valor de actividad de N obtenido a concentraciones altas de proteína y y_D es la actividad del monómero, que es cero.

Los datos se ajustaron a un modelo de disociación descrito por la ecuación:

$$f_N = \frac{K_{dis} + 4Pt - (K_{dis}^2 + (8K_{dis} * Pt)^2)}{4Pt} \quad (10)$$

donde K_{dis} es la constante de disociación del proceso a 25°C y Pt es la concentración de proteína a la que se incubaron la muestras durante 24 horas (Cisneros Armas, 2003).

Calorimetría de titulación isotérmica (ITC)

La calorimetría de titulación isotérmica, al igual que la DSC es una técnica que permite medir de manera directa cambios en la energía de un sistema. A diferencia de DSC, ITC no aumenta la temperatura de las celdas de muestra y referencia, sino que las mantiene constantes al igual que la diferencia de temperatura ($\Delta T=0$) entre ambas. Para los experimentos de disociación, se coloca amortiguador en la celda de muestra y se inyecta un volumen determinado de proteína. En principio, la proteína se disocia al diluirse en el buffer, esto va acompañado de una liberación o absorción de calor, el equipo suministra o deja de suministrar energía a la celda para regresar a $\Delta T=0$.

Los experimentos de calorimetría de titulación isotérmica para TIM63 se realizaron en un equipo ITC 200 de MicroCal (IQ-UNAM). Las proteínas se concentraron por medio de centricones hasta llegar a una concentración entre 0.35 y 0.45 mM (10 y 12 mg/ml) y se centrifugaron para eliminar agregados. Los ensayos de dilución se realizaron a diferentes temperaturas y constaron en inyecciones de 1 ul de muestra cada 150 segundos en una celda con 200 ul de amortiguado TEA 10 mM, NaCl 100 mM, pH 7.6 y agitación constante hasta alcanzar la saturación.

Los experimentos de calorimetría de titulación isotérmica para TIM55 se realizaron en un equipo VP-ITC de MicroCal (IQ-UNAM). Las proteínas se concentraron por medio de centricones hasta llegar a una concentración de 0.45 mM (14 mg/ml) y se centrifugaron para eliminar agregados. Los ensayos de dilución se realizaron a diferentes temperaturas y constaron en inyecciones de 5 ul de muestra cada 300 segundos en una celda con 2 ml de amortiguador de actividad (TEA 10 mM, NaCl 100 mM, pH 7.6) y agitación constante hasta alcanzar la saturación.

Suponiendo un modelo de disociación de dos estados en donde el dímero (P_2) se disocia en dos monómeros ($2P$):



$$K_{dis} = (P)^2/(P_2) \quad (12)$$

las trazas se ajustaron a la ecuación de disociación:

$$q_i = \Delta H_{dis} * (P_2)_{syr} * dV_i - \Delta H_{dis} [(P_2)_i - (P_2)_{i-1}][V_0 + \frac{dV_i}{2}] \quad (13)$$

donde q_i es el calor liberado en la inyección número i , V_i el volumen agregado por inyección, (P_2) la concentración de dímero en la jeringa (syr), en la inyección i o en la inyección $i-1$, V_0 el volumen de la celda, ΔH_{dis} el calor de disociación y K_{dis} la constante de disociación.

De esta manera se obtuvieron ΔH_{dis} y K_{dis} . ΔG_{dis} y ΔS_{dis} se calcularon a partir de las relaciones termodinámicas básicas:

$$\Delta G_{dis} = -RT * \ln K_{dis} \quad (14)$$

$$\Delta G_{dis} = \Delta H_{dis} - T * \Delta S_{dis} \quad (15)$$

Se utilizó la ecuación de Gibbs-Helmholtz modificada para la simulación de la curva de ΔG_{dis} para TIM63, en las que se usaron los parámetros obtenidos experimentalmente a una temperatura de referencia T^{ref} :

$$\Delta G_{dis} = \Delta H_{dis}^{ref} - T^{ref} \frac{\Delta H_{dis}^{ref} - \Delta G_{dis}^{ref}}{T^{ref}} + \Delta Cp (T - T^{ref} - T \ln \frac{T}{T^{ref}}) \quad (16)$$

ΔH_{dis}^{ref} y ΔG_{dis}^{ref} son la entalpía y la energía libre, respectivamente (Cisneros Armas, 2003).

RESULTADOS Y DISCUSIÓN

Expresión y solubilidad de las proteínas recombinantes

Tanto TIM63 como TIM55 se sobreexpresan notablemente y se encuentran en la fracción soluble obtenida después de sonicar las células (fig. 10A). El corte de la etiqueta de histidinas por la trombina fue monitoreado en un gel de acrilamida con la resolución necesaria para diferenciar entre bandas de 27 y 29 kDa, que son los pesos de las proteínas sin y con etiquetas, respectivamente (fig. 10B). La purificación de ambas proteínas ancestrales procede exitosamente y resulta en la obtención de proteínas recombinantes con un grado de pureza de aproximadamente 95% (fig. 10C y D).

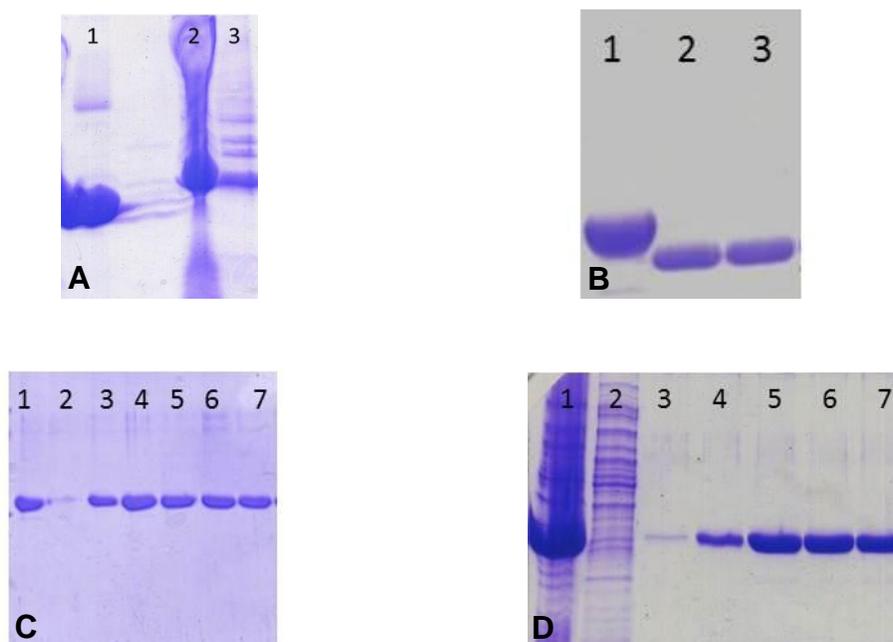


Figura 10. Purificación de TIMs ancestrales. Geles desnaturizantes de SDS para demostrar (A) solubilidad de TIMs recombinantes, (B) corte exitoso de la etiqueta de histidinas, (C) y (D) purificación de proteínas recombinantes. Carriles: (A) 1: TIM de *E. histolytica* (control); 2: sobrenadante celular después de sonicar; 3: pellet celular después de sonicar. (B) 1: TIM63 con etiqueta de histidinas (29 kDa), 2 y 3: TIM63 sin etiqueta (27 kDa). (C) 1-7: fracciones después de la primera columna de purificación por afinidad para TIM63. (D) 1: extracto celular, 2: frente de elución, 3-7: fracciones de después de la primera columna de purificación para TIM55.

Estructuras plegadas

Como primera aproximación a las características de las TIMs ancestrales, se corrió un blastp para TIM63 y TIM55 y se encontró que las secuencias con mayor similitud corresponden a la TIM del trigo y del maíz, respectivamente, y la identidad corresponde al 75% para ambas. Tomando esto en cuenta, una de las primeras dudas que surgieron al sintetizar, clonar y expresar a las proteínas era si éstas contarían con un plegamiento definido y en específico con el plegamiento característico de los barriles TIM. Para poder resolver esta incógnita se obtuvieron los espectros de dicroísmo circular en el UV lejano y de fluorescencia intrínseca de los ancestros (fig. 11).

Dicroísmo circular. El patrón de DC de los ancestros presenta un comportamiento típico de plegamiento tipo barril TIM, con una máximo de absorción aproximadamente a 190 nm y dos mínimos a ~210 y ~220 nm, las hélices alfa son la mayor contribución en estos últimos puntos. Se utilizó a la TIM de levadura (ScTIM) como control y se observa una alta coincidencia entre los patrones. Estas características nos hablan de la presencia de estructura secundaria α y β definida en ambos ancestros. Para confirmar la presencia de un núcleo hidrofóbico se obtuvieron los espectros de emisión de fluorescencia.

Fluorescencia. Ambas proteínas presentan un máximo de emisión cercano a 335 nm, que indica triptófanos escondidos en un ambiente hidrofóbico. Si bien este valor está corrido un poco hacia el rojo en comparación al de ScTIM, el de proteínas desnaturalizadas está alrededor 355 nm. Estos resultados nos sugieren que los ancestros cuentan con estructura secundaria así como con terciaria (presencia de un núcleo hidrofóbico). La diferencia en la posición del máximo de emisión entre los ancestros y ScTIM se puede deber al número de triptófanos en cada una de las proteínas. TIM63 y TIM55 cuentan con dos triptófanos extras en comparación a ScTIM y posiblemente están expuestos al solvente.

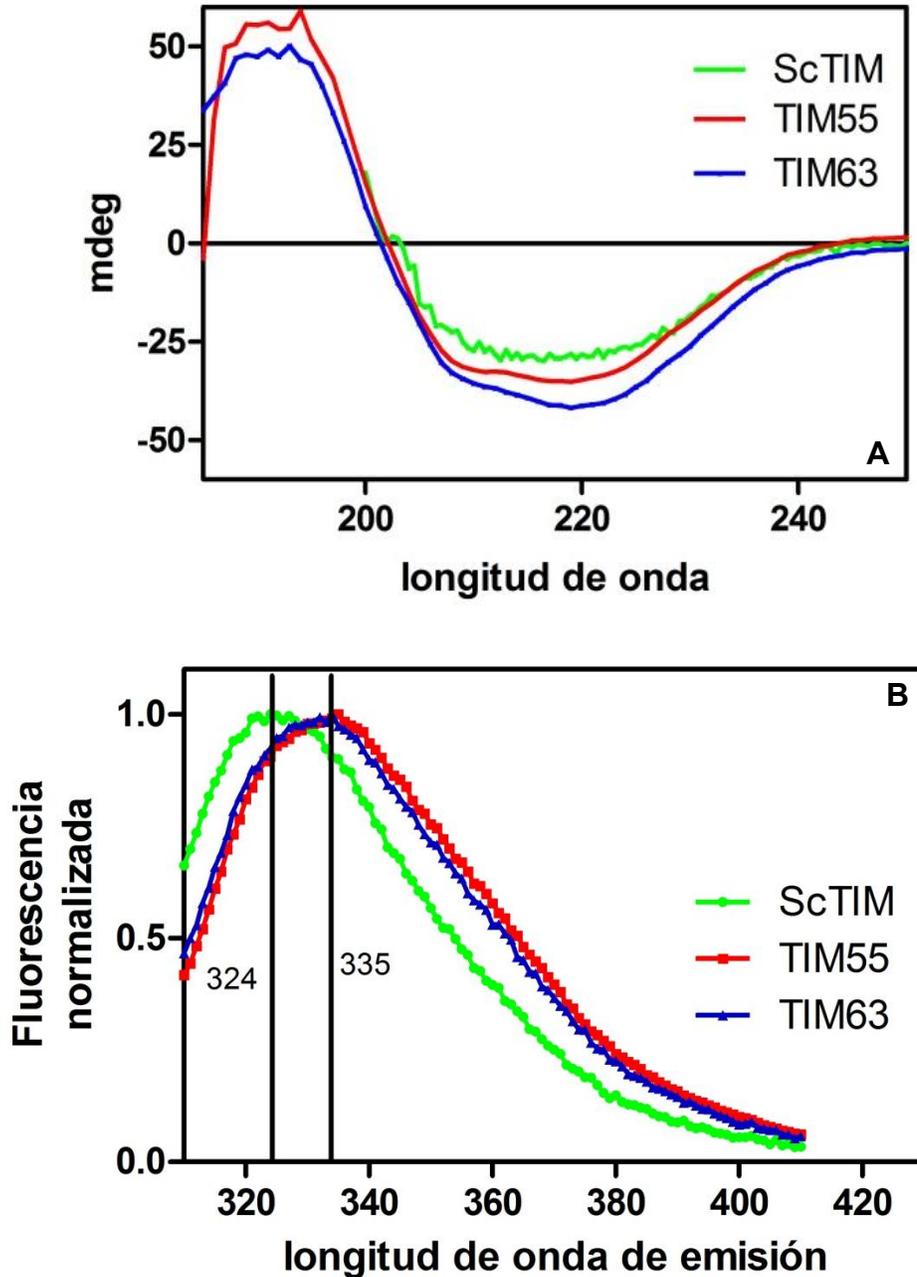


Figura 11. Propiedades espectroscópicas para TIM63sc y TIM55sc. (A) Espectros de dicroísmo circular en el UV lejano. (B) Espectros de emisión de fluorescencia excitando a 295 nm.

Unión de ANS. Una evidencia extra de que nuestros ancestros cuentan con una estructura plegada fue analizar la unión de ANS por fluorescencia. El ANS es una molécula aromática que se une a regiones principalmente hidrofóbicas y fluoresce

en ambientes apolares. En caso que TIM63 y TIM55 estuvieran “parcialmente” plegadas y presentaran regiones hidrofóbicas expuestas, entonces la unión de ANS aumentaría considerablemente al aumentar la concentración de proteína como pasa con la albúmina (BSA). BSA es un proteína con parches hidrofóbicos, por lo que sirvió en este experimento como control positivo de la unión de ANS. Para TIM63, su comportamiento es semejante al de ScTIM (fig. 12).

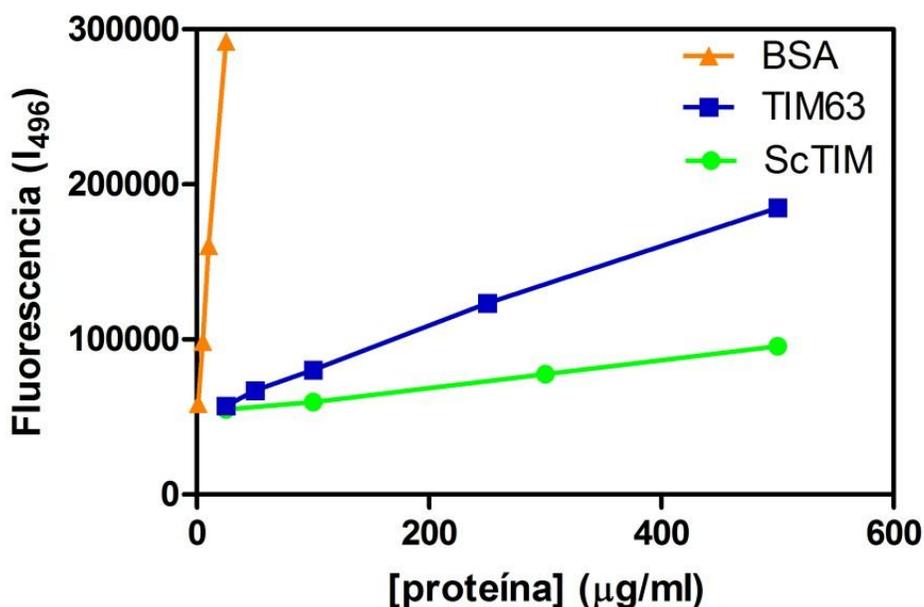


Figura 12. Áreas hidrofóbicas en TIMs ancestrales. Unión de ANS para TIM63, ScTIM (control negativo) y BSA (control positivo) a diferentes concentraciones de proteína.

Cromatografía de exclusión molecular. Para poder analizar la estructura cuaternaria de los ancestros y verificar si éstos son dímeros como el resto de las TIMs de eucariontes, se utilizó una columna de exclusión molecular. Con ayuda de proteínas de referencia, se calculó el radio hidrodinámico (radio de Stokes: R_S) de los ancestros a partir de los perfiles de elución (fig. 13). Los valores de R_S para TIM63 y TIM55 son de 28.1 Å y 27.7 Å, respectivamente, valores cercanos a 29.6 y 30 Å, que son los valores para ScTIM en estado dimérico observado por NMR y cromatografía, respectivamente (Morgan, 2000; Nájera, 2003).

En conjunto estos resultados demuestran que los ancestros cuentan con una estructura plegada, compuesta de elementos de estructura secundaria en la que los aminoácidos hidrofóbicos están ocultos del solvente. Además presentan una estructura cuaternaria dimérica compacta, con dimensiones hidrodinámicas semejantes a las encontradas en todas las TIMs extantes de eucariontes.

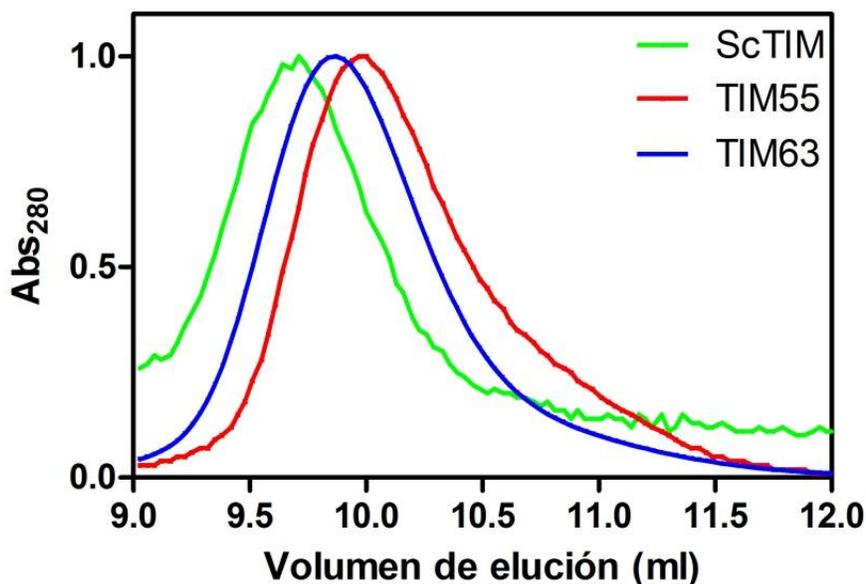


Figura 13. Determinación del radio de Stokes. Perfiles de elución normalizados. [ScTIM]=1mg/ml, [TIM55] y [TIM63]=3 mg/ml.

Estabilidad térmica de los ancestros

Dicroísmo circular. Una vez confirmada la presencia de una estructura plegada, se analizó la estabilidad de la misma. Para ello se utilizó la temperatura como agente desnaturizante. En primer lugar se realizó un experimento en el que se siguió el cambio en elipticidad molar a 222 nm de ambas proteínas al aplicarles una rampa de temperatura (60 °C/hora) (fig. 14). El proceso de desnaturización térmica para ambos ancestros fue irreversible y se observó agregación, sin embargo, siguió una sola transición sigmoide y se ajustó usando la ecuación de

Boltzmann (materiales y métodos). A pesar de que con este tipo de análisis solamente se obtienen parámetros aparentes, ya que la presencia de agregados impide realizar el análisis termodinámico formal del desplegamiento, es posible subrayar algunas observaciones. TIM63 presentó una mayor pendiente en la transición, lo que indica que el proceso de desnaturalización es más cooperativo para este ancestro que para TIM55. Las temperaturas medias aparentes de desnaturalización ($T_{m_{app}}$) de ambas proteínas son semejantes (TIM63: 68.7 °C y TIM55: 67.4 °C). Estos valores de T_m son cercanos a los valores más elevados observados en TIMs silvestres o diseñadas (tabla 4). Este primer parámetro obtenido sugiere una alta termoestabilidad de la estructura de las TIMs ancestrales, que a su vez respalda la validez de la reconstrucción ancestral, ya que no sólo se generaron proteínas que se pliegan correctamente, sino que éstas permaneces plegadas a temperaturas elevadas

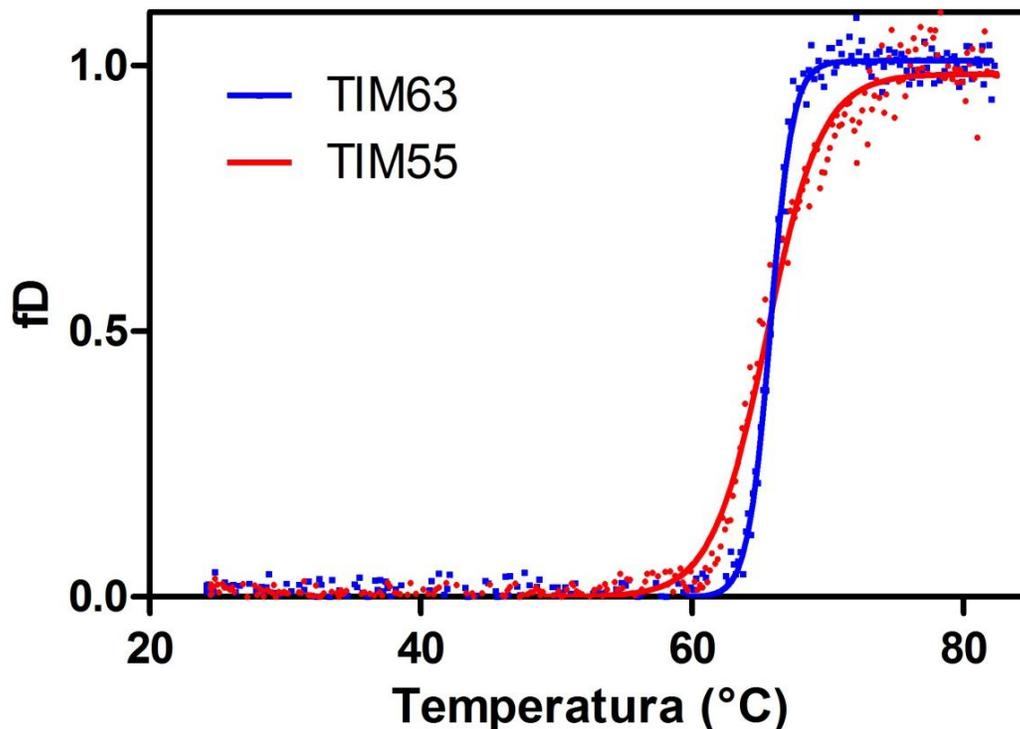


Figura 14. Estabilidad térmica por DC. Desnaturalización térmica seguida por el cambio en la elipticidad a 222 nm (velocidad de calentamiento: 60 °C/hora) Las líneas muestran el ajuste a la ecuación de Boltzmann en datos normalizados (fD: fracción desnaturalizada).

Tabla 4. Tm's de diferentes organismos.

TIM	Tm_{app} (°C)
LFFFL^a	69.9
LFFYA^a	69.8
TIM63	68.7 ¹ , 69.5 ²
TIM55	67.4 ¹ , 73.6 ²
ScTIM^a	65.8
PLATA^a	44.5 (dis) 60.2 (des)
AFAAS^a	43.3 (dis) 58.9 (des)
EhTIM^b	41.5 (dis) 57.0 (des)
PfTIM^c	58
monoTIM^d	57
TbTIM^e	53.6
RE-TIM^d	52

^a (Peimbert et. al., 2008), ^b (Téllez, et.al., 2008), ^c (Gokhale et. al., 1999), ^d (Schliebs et. al., 1996), ^e (Guzmán-Luna, 2006),

¹ obtenida por DC, ² obtenida por DSC. dis: disociación, des: desplegamiento. LFFFL, LFFYA: TIMs diseñadas para estabilizar la interfase con el programa Rosetta. PLATA, AFAAS: TIMs modificadas por evolución dirigida.

Calorimetría diferencial de barrido (DSC). Otra técnica que se usó para estudiar la termoestabilidad de los ancestros fue la calorimetría diferencial de barrido. Se calentó a la muestra a la misma velocidad que en el estudio por dicroísmo circular (60 °C/hora) y a una velocidad mayor (90 °C/hora), pero a concentraciones de proteína mayores. Un detalle que hay que destacar es que las primeras corridas de DSC fueron a 0.4 mg/ml de proteína y se observó mucho ruido experimental, la concentración de proteína a la que las trazas se ven bien definidas es a 1mg/ml. Dicho esto, hay que recordar que en DC no se puede utilizar alta concentración de proteína debido a la alta absorción de la muestra, por lo que resultó imposible realizar los experimentos a la misma concentración de proteína por ambas técnicas.

Dependencia de la Tm con la concentración de proteína. La presencia de una sola transición en las curvas de la figura 14 sugiere que el desplegamiento y la disociación son eventos concertados y por ello solamente se observan el estado nativo y el estado desnaturalizado sin intermediarios. Ya que la TIM es dimérica en

el estado nativo, la fracción de proteína que se encuentre en este estado depende de la concentración total de la misma (Pt) (ec. 10). En la tabla 5 se presentan la $T_{m_{app}}$ para las diferentes concentraciones y velocidades de calentamiento por DC y DSC. Por problemas en la función del calorímetro, se perdieron todas las trazas obtenidas a una tasa de calentamiento de 60°C/hora, por lo que los datos reportados para esa velocidad de calentamiento corresponden solamente a los obtenidos por DC. Para TIM55 las T_m 's por DC son significativamente menores que las obtenidas por DSC, lo cual podría indicar la presencia de intermediarios de plegamiento. Para TIM63 la T_m obtenida por las dos técnicas es semejante excepto por el punto a 250 $\mu\text{g/ml}$. Otra diferencia entre las técnicas es que por DSC, sin importar la velocidad de barrido, TIM55 se despliega a más altas temperaturas que TIM63. La única coincidencia entre ambas técnicas tiene que ver con la cooperatividad del proceso para ambas proteínas. Las curvas de desnaturalización por DSC para TIM63 presentan un pico más angosto (fig. 15) que está correlacionado con mayor cooperatividad. Independientemente de las diferencias, los datos de la tabla 5 muestran que a mayores concentraciones de proteína, la T_m tiene una tendencia a disminuir.

Tomando en cuenta que la TIM es dimérica, lo que uno espera es que la T_m aumente a mayores concentraciones de proteína debido a que la población de dímero también debería de ser mayor. El comportamiento opuesto por las TIMs ancestrales es atribuible a la formación de agregados; a mayor concentración de proteína se generan más agregados y por lo mismo se desestabiliza a las poblaciones de dímero y monómero, esto resulta en la disminución de la T_m aparente. Esta misma razón puede ser la explicación para las diferencias observadas entre las técnicas.

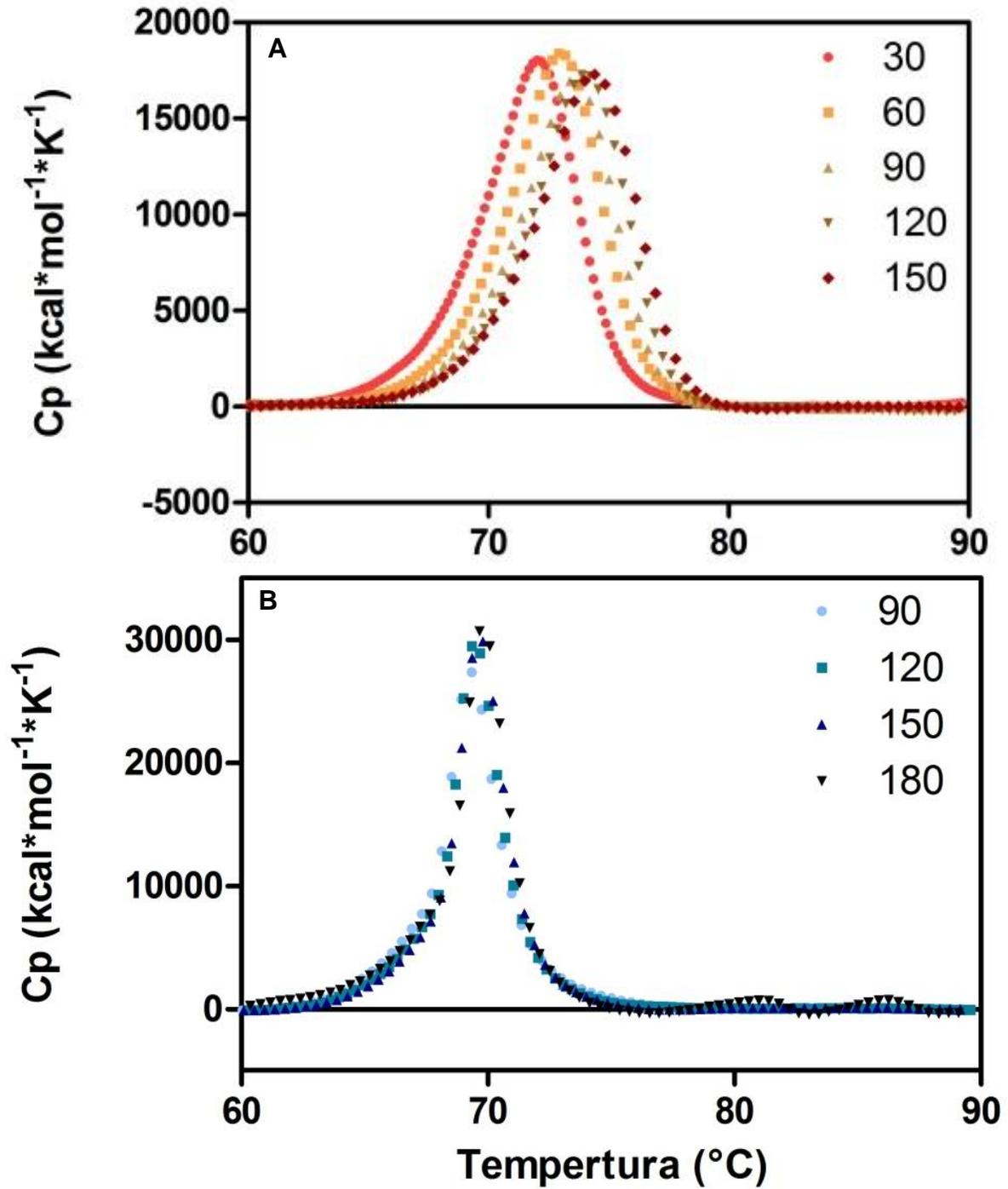


Figura 15. Dependencia de la T_m con la velocidad de barrido. Endotermas para el desplegamiento de TIM55 (A) y TIM63 (B) a diferentes velocidades de barrido en $^{\circ}\text{C}/\text{hora}$.

Tabla 5. Tm vs concentración de proteína.

VB	60 °C/hora		90 °C/hora	
	[TIM] (mg/ml)	Tm (°C)	[TIM] (mg/ml)	Tm (°C)
TIM63	0.05 ^a	68.72	0.40 ^b	67.69
	0.25 ^a	65.74	0.60 ^b	67.34
			1.00 ^b	69.52
			3.30 ^b	65.53
TIM55	0.05 ^a	67.42	0.40 ^b	70.71
	0.25 ^a	65.48	1.00 ^b	73.17
	1.00 ^b	72.62	2.78 ^b	69.64

^a datos obtenidos por DC

^b obtenidos por DSC

Dependencia de la Tm con la velocidad de barrido (DSC). Dado que el proceso de desplegamiento, tanto para TIM63 como para TIM55, es irreversible, a partir de los datos de DC solamente podemos obtener parámetros aparentes, ya que las especies en el proceso no están en equilibrio, si no que se están desviando hacia un estado de agregación. Para poder continuar con el análisis calorimétrico de las proteínas se tomó un camino alternó y se probó si el desplegamiento térmico presenta control cinético. La figura 15 presenta las endotermas obtenidas a diferentes velocidades de barrido (VB) para ambas proteínas. Cada traza fue ajustada al modelo de dos estados irreversible (ver materiales y métodos), mediante el cual se obtuvieron las Tm's a cada velocidad de barrido (fig. 16, tabla 6). En el caso de TIM55 se observa una evidente dependencia de la Tm con la velocidad de barrido, mientras que para TIM63 las trazas a las diferentes velocidades de barrido coinciden bastante bien y la Tm cambia menos de 1 grado entre las velocidades más altas y más bajas. La ausencia de un cambio significativo en este parámetro puede interpretarse como que las especies de proteína están en equilibrio a cada temperatura independientemente de la velocidad de barrido y por lo mismo la barrera energética que necesita ser superada para pasar de un estado a otro es pequeña. Para profundizar en ello se construyó un gráfico de Arrhenius tomando en cuenta todas las velocidades de barrido (fig. 17). La coincidencia entre todos los puntos es indicador del buen ajuste del modelo usado. A partir de la pendiente se obtiene el valor de la energía

de activación (E_{act}). Siguiendo la idea de que la T_m no cambia con la velocidad de barrido y por lo mismo la barrera energética debía ser pequeña, se esperaba una E_{act} muy pequeña, sin embargo, fue una sorpresa encontrar que ambos ancestros tienen una energía de activación elevada y dentro del rango alto de los valores reportados para diferentes TIMs (Tabla 7). La energía de activación es la energía necesaria para pasar del estado nativo al estado de transición antes de llegar al estado desplegado. Una energía de activación alta habla de un estado de transición energéticamente muy costoso, posiblemente uno en el que las interacciones internas que estabilizan al estado nativo ya se rompieron pero todavía no se han solvatado.

Tabla 6. T_m contra velocidad de barrido.

VB (°C/hr)	T_m (°C)	
	TIM63	TIM55
30	--	72.18 ± 0.037
60	--	73.11 ± 0.05
90	69.52 ± 0.009	73.62 ± 0.037
120	69.63 ± 0.089	74.04 ± 0.026
150	69.88 ± 0.009	74.42 ± 0.035
180	69.97 ± 0.150	74.70 ± 0.03

Los errores presentados corresponden a los del ajuste.

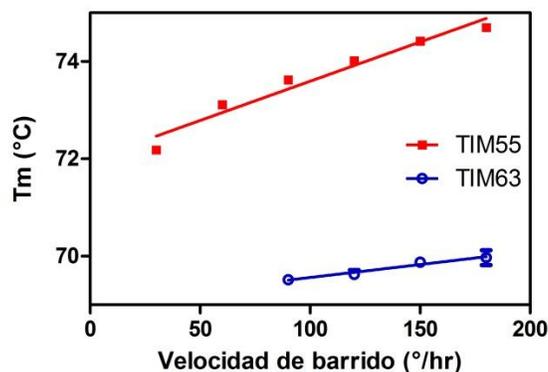


Figura 16. T_m vs velocidad de barrido.

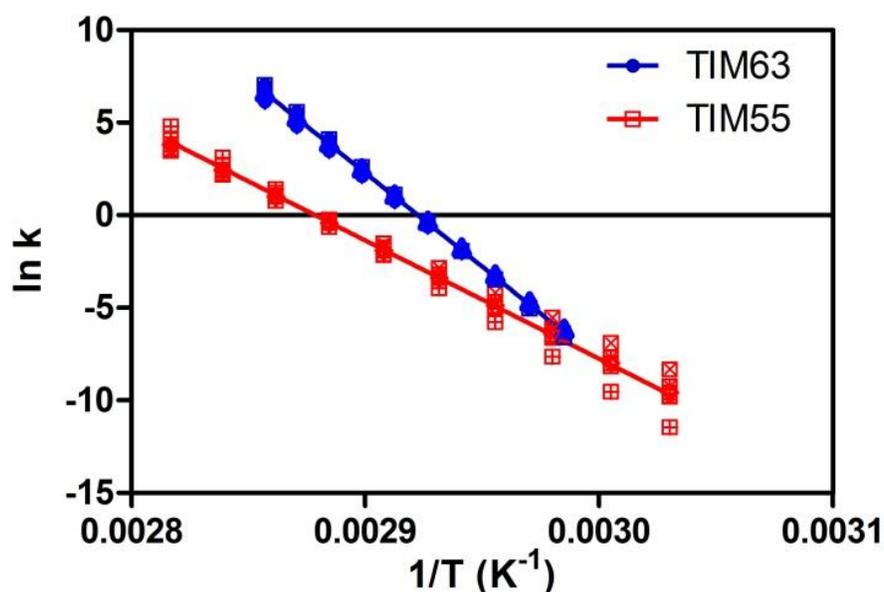


Figura 17. Energías de activación. Gráfico de Arrhenius para ambos ancestros a diferentes velocidades de barrido (símbolos superpuestos). La pendiente está definida por la $-E_{act}/R$. Cada símbolo representa una velocidad de barrido diferente.

Tabla 7. Energías de activación (Kcal/mol)

	TIM63	TIM55	LmTIM ^a	TbTIM ^a	TcTIM ^a
Arrhenius	201.1 ± 1.5	127.0 ± 1.9	75.3 ± 0.8	95.1 ± 1.1	189.5 ± 4.2
Lumry-Eyring¹	200.3 ± 6.3	120.7 ± 2.9	74.6 ± 1.1	94.9 ± 1.3	193.5 ± 1.7

^a (Costas et. al., 2009).

¹Promedio a partir las trazas individuales a cada velocidad de barrido ajustadas al modelo.

Determinación de la entalpía total de desplegamiento. La entalpía de desplegamiento ΔH_{tot} se calculó por el área bajo la curva de las trazas de DSC. Los valores son 91.5 y 95.8 Kcal/mol de monómero para TIM63 y TIM55, respectivamente. Estos valores son el promedio de lo obtenido a cada velocidad de barrido a una concentración de proteína de 1 mg/ml. El valor calculado a partir de estudios del desplegamiento de proteínas globulares está en aproximadamente 240 kcal/mol de monómero para una proteína de 250 aminoácidos (Robertson &

Murphy, 1997). Tomando esto en cuenta, los valores obtenidos para las TIMs ancestrales son muy bajos.

Catálisis

La presencia de una estructura plegada estable en TIM63 y TIM55 es por sí misma un gran éxito en nuestro proceso de reconstrucción ancestral de proteínas, sin embargo, la estructura global no está directamente ligada a la catálisis enzimática, ésta requiere de mucha más precisión en la posición de los aminoácidos catalíticos, por ello la presencia de actividad enzimática es un criterio más astringente sobre la calidad de la reconstrucción. Las cinéticas con el sustrato (G3P) se ajustaron a la ecuación de Michaelis-Menten para obtener los parámetros catalíticos (fig. 18). En el caso de TIM55, las actividades registradas son muy pequeñas y varían de lote a lote, por lo mismo, son difíciles de cuantificar con precisión. En la figura 18 se muestran los mejores valores obtenidos para ese ancestro. En el caso de TIM63, la actividad es comparable a la de TIMs extantes (tabla 8). Incluso la velocidad máxima (V_{max}) para TIM63 es la más alta obtenida para TIMs diseñadas o modificadas mediante ingeniería de proteínas.

Normalmente para poder realizar la determinación de los parámetros cinéticos de las TIMs se utilizan 5 ng de enzima en la celda de reacción, sin embargo, en nuestros experimentos a esa cantidad de proteína se observaba una curvatura inusual que representaba una disminución en la pendiente; dicho de otra manera, una disminución en la velocidad de reacción. Atribuimos este comportamiento a que la enzima se disocia y por lo mismo, pierde la capacidad catalítica observada sólo en el dímero. Para poder contrarrestar el efecto, se usó una mayor concentración de proteína en la celda. Esto tiene dos implicaciones importantes, la primera es que se está suponiendo que TIM63 se comporta como todas las TIMs extantes y solamente tiene actividad cuando está en su forma dimérica; la segunda es que la enzima se disocia en tiempos muy rápidos, tanto así que es posible ver la disociación en la celda de reacción.

Por otra parte, también notamos que las curvas a altas concentraciones de sustrato tienen una menor inactivación que aquellas con poco G3P, por lo que creemos que la unión del sustrato estabiliza al dímero. Es común que la unión de algún ligando estabilice a una estructura proteica, ya que los contactos creados provocan menor movilidad de residuos, así como posicionamiento específico de los mismos; además, por el principio de Le Chatelier, se genera un equilibrio entre las especies unidas al ligandos y las especies “libres”, dando lugar a una menor población de enzima inactiva (monómeros). A pesar de las complicaciones que implica la disociación de las enzimas en el proceso de determinación de actividad, decidimos aprovechar esta característica en los ancestros para caracterizar la disociación por dilución.

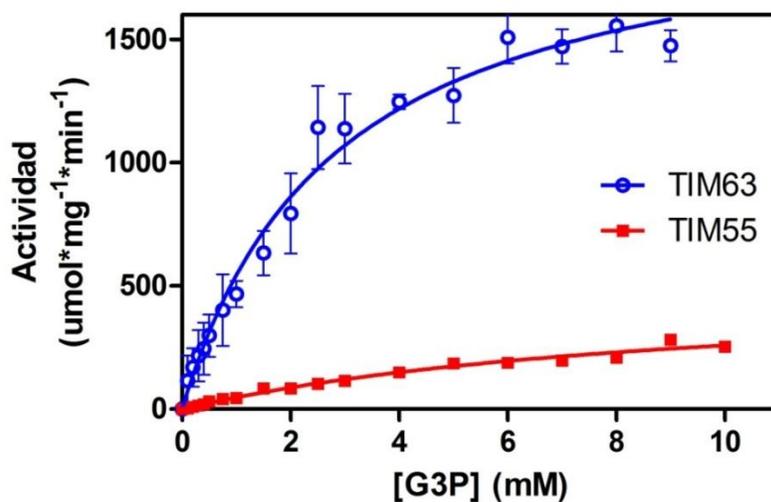


Figura 18. Actividad enzimática. Curvas de la cinética enzimática para ambos ancestros y ajuste a la ecuación de Michaelis-Menten. Las barras de error muestran la desviación del promedio de 3 experimentos para TIM63. Los valores reportados para TIM55 corresponden a un solo experimento, el que arrojó velocidades mayores.

Tabla 8. Parámetros catalíticos de TIMs extantes, diseñadas y ancestrales.

Proteína	k_{cat} (s ⁻¹)	K_m (mM)	k_{cat}/K_m (s ⁻¹ M ⁻¹)*
ScTIM ^a	6.40×10^3	1.4	1.3×10^8
HsTIM ^b	6.20×10^3	0.25 ± 0.05	7.2×10^8
PfTIM ^c	4.30×10^3	0.35 ± 0.16	3.6×10^8
GitIM ^d	4.80×10^3	0.53 ± 0.03	2.6×10^8
TmTIM ^e	4.30×10^3	0.84	1.5×10^8
OcTIM ^e	3.69×10^3	0.62	1.7×10^8
LFYAA ^a	1.04 ± 0.04	1.26 ± 0.12	2.3×10^4
LFFFL ^a	0.69 ± 0.05	1.20 ± 0.21	1.6×10^4
PLATA ^a	$4.16 \times 10^3 \pm 220$	2.44 ± 0.20	5.0×10^7
AFAAS ^a	$7.02 \times 10^3 \pm 210$	1.09 ± 0.07	1.9×10^8
monoTIM ^b	3.10×10^2	4.10 ± 0.6	2.1×10^6
TIM63	$2.08 \times 10^3 \pm 101$	2.80 ± 0.34	2.2×10^7
TIM55	$5.04 \times 10^2 \pm 60$	9.49 ± 1.90	1.5×10^6

^a (Peimbert et. al., 2008), ^b (Schliebs et. al., 1997), ^c (Gokhale et. al., 1999), ^d (Reyes-Vivas et. al., 2007), ^e (Knobeloch et. al., 2010).

* k_{cat}/K_m : eficiencia catalítica realizando una 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 1:29).

Caracterización de la disociación

Cromatografía de exclusión molecular. El primer acercamiento al proceso fue de modo cualitativo. Los ancestros se pasaron por una columna de exclusión molecular a diferentes concentraciones de proteína y se determinó el radio de Stokes de cada muestra (fig. 19A y B). Es posible observar un aumento en el volumen de elución al disminuir la concentración de proteína, lo que corresponde a una disminución del radio hidrodinámico (fig. 19C) hasta llegar a un valor cercano a 23 Å, lo que correlaciona con el radio hidrodinámico de un intermediario monomérico detectado por NMR en el desplegamiento de ScTIM (Morgan, 2000). Estos resultados sugieren que la proteína pasa de un estado dimerico a uno monomérico a bajas concentraciones. La cromatografía es un tipo de experimento cualitativo, ya que la concentración de proteína determinada antes de pasarla por la columna cambia una vez dentro de la misma y no se puede obtener de manera

exacta la constante de disociación (K_{dis}). Aprovechando la ausencia de actividad en el monómero, decidimos hacer experimentos acoplando la dilución con la inactivación de la catálisis.

Inactivación. Después de incubar toda la noche a la proteína a diferentes concentraciones, se midió su actividad. El principio detrás del experimento consiste en diluir a la proteína suponiendo que las poblaciones de dímero y monómero se equilibran dependiendo de la constante de disociación del proceso. Después de 24 horas, se mide la actividad de cada dilución a una concentración de proteína constante en la celda de reacción. Se espera que en concentraciones de incubación menores la población de monómeros sea mayor y por lo mismo, se observe menos actividad que en las incubaciones a alta concentración de proteína. Este resultado fue justamente el obtenido en el experimento. Los datos se normalizaron y se ajustan a un modelo de disociación de dos estados, a partir del cual se puede obtener un estimado de la constante de disociación (fig. 20). Los valores de K_{dis} representan la afinidad entre monómeros, el valor mayor para TIM55 en comparación al de TIM63 refleja una menor afinidad entre los monómeros de TIM55. En comparación a la TIM de humano, estos valores son mucho mayores y más bien se acercan a los reportados para una mutante parcialmente desestabilizada en la interfase (M14Q) que se presenta parcialmente como monómero (tabla 9).

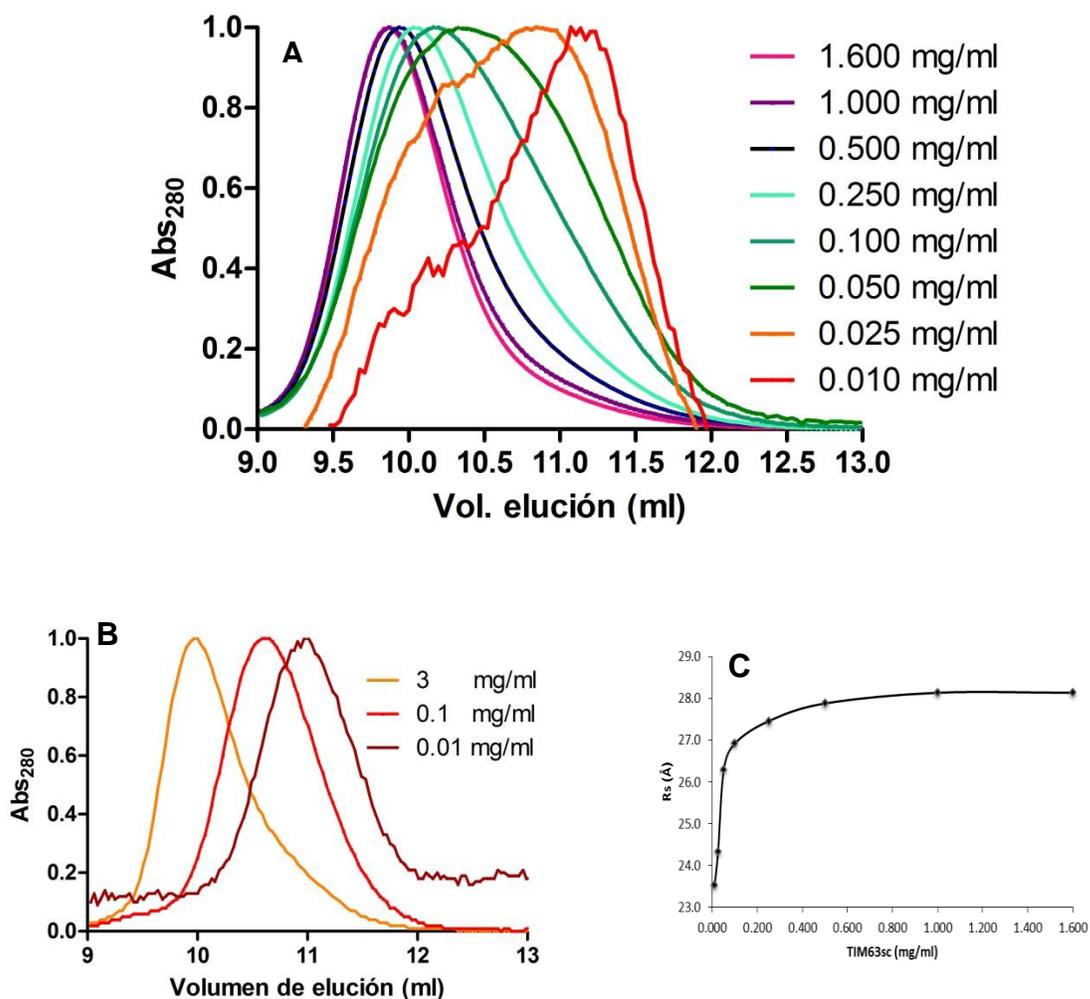


Figura 19. Disociación y cromatografía de exclusión molecular. Perfiles de elución de las proteínas ancestrales después de pasar por una columna de exclusión molecular a diferentes concentraciones de proteína para TIM63 (A) y TIM55 (B). Absorbancia a 280 nm normalizada. (C) Disminución en el radio hidrodinámico de TIM63 al reducir la concentración de proteína.

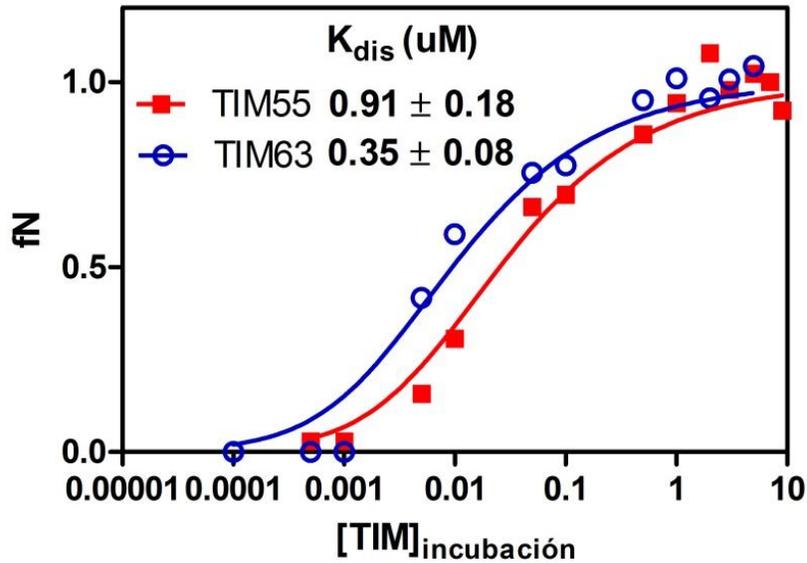


Figura 20. Pérdida de la actividad catalítica por dilución. Se estudió la presencia de actividad catalítica en los ancestros después de haberlos incubado por 24 horas a concentraciones bajas de proteína. La línea representa el ajuste de los datos a un modelo de disociación de 2 estados (materiales y métodos). El eje de las ordenadas está en escala logarítmica y las unidades son mg/ml. (fN: fracción nativa).

Tabla 9. Constantes de disociación para diferentes TIMs.

TIM	K_{dis} (M)
HsTIM ^a	32.1×10^{-12}
HsTIM (M14Q) ^a	45.2×10^{-8}
TIM63	34.7×10^{-8}
TIM55	92.9×10^{-8}

^a (Mainfroid et. al., 1996)

ITC para TIM63. Normalmente, es difícil analizar el proceso de disociación de las TIMs por la alta afinidad de los monómeros o la constante de velocidad lenta que acompaña al proceso, sin embargo, nuestros ancestros demostraron ser buenos candidatos para este análisis termodinámico debido a su rápida disociación ligada a la dilución. Para ello se realizaron experimentos de calorimetría de disociación isotérmica. En la figura 21 se pueden observar las

isotermas de disociación de los experimentos con TIM63 a seis temperaturas diferentes.

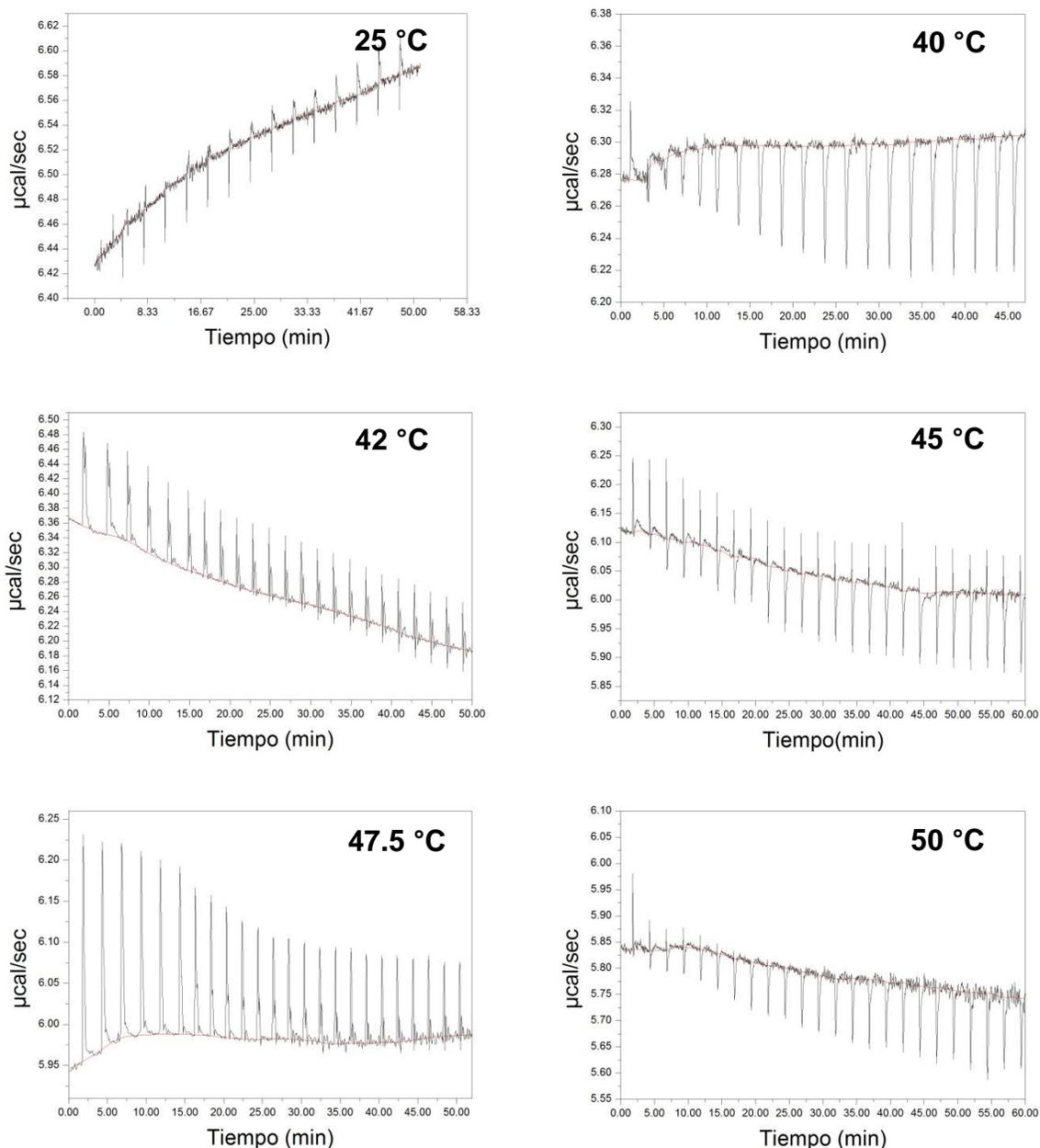


Figura 21. Isotermas de la disociación de TIM63. Datos crudos, sin normalizar con la concentración de proteína o restar la línea base. En cada imagen se presenta la temperatura a la que se realizaron los experimentos. Las concentraciones de proteína usadas varían para cada experimento y están en el rango de 10-12 mg/ml. Cada pico corresponde a una inyección de 1 µl de muestra de proteína en la celda.

Es posible integrar cada pico y obtener el calor asociado a cada evento. La figura 22 presenta las gráficas obtenidas a partir de dicha integración. Cada una de las gráficas fue ajustada al modelo de disociación (materiales y métodos) para obtener el valor de la entalpía de disociación (ΔH_{dis}). La pendiente resultante al graficar el cambio de ΔH_{dis} con respecto a la temperatura corresponde a la capacidad calorífica de disociación ($\Delta C_{\text{p}_{\text{dis}}}$) (fig. 23A). Se ha observado que este parámetro está directamente relacionado con el cambio en la superficie accesible al solvente (ΔASA), para un proceso de disociación corresponde al área que se expone en la interfase del dímero. Es importante mencionar que este análisis se realizó suponiendo la disociación de cuerpos rígidos, es decir, donde no hay cambios conformacionales asociados al proceso y el área que se expone corresponde únicamente a la que estaba escondida en la interfase. El valor experimental se comparó con valores teóricos calculados con diferentes modelos a partir del promedio de área polar y no polar presente en la interfase de diferentes TIMs extantes (tabla 10). El ΔC_{p} experimental es ligeramente mayor que el teórico, la diferencia de los valores puede tener dos implicaciones importantes: 1) el área en la interfase de los dímeros ancestrales es mayor que la de los dímeros extantes, lo cual es poco probable, o 2) el área expuesta en la disociación de los dímeros corresponde no solamente al área de contacto entre monómeros, sino que el proceso está acoplado a cierto desplegamiento de los mismos. Lo más probable es que suceda lo propuesto en la segunda opción y que los monómeros no se comporten como cuerpos rígidos. Dado que en la literatura no existen valores de $\Delta C_{\text{p}_{\text{dis}}}$ para TIMs, lo más cercano son valores de ΔC_{p} totales ($\Delta C_{\text{p}_{\text{tot}}} = \Delta C_{\text{p}_{\text{dis}}} + 2\Delta C_{\text{p}_{\text{unf}}}$) de TIMs de bacterias, en las que se encontró por primera vez un proceso reversible de desnaturalización y renaturalización por DSC (Romero-Romero, 2013). Los valores de $\Delta C_{\text{p}_{\text{tot}}}$ son 6 veces más grandes que los obtenidos para las TIMs ancestrales ($\sim 6 \text{ kcal}\cdot\text{mo}^{-1}\cdot\text{K}^{-1}$), lo que indica que la cantidad de estructura desplegada asociada a la disociación es poca.

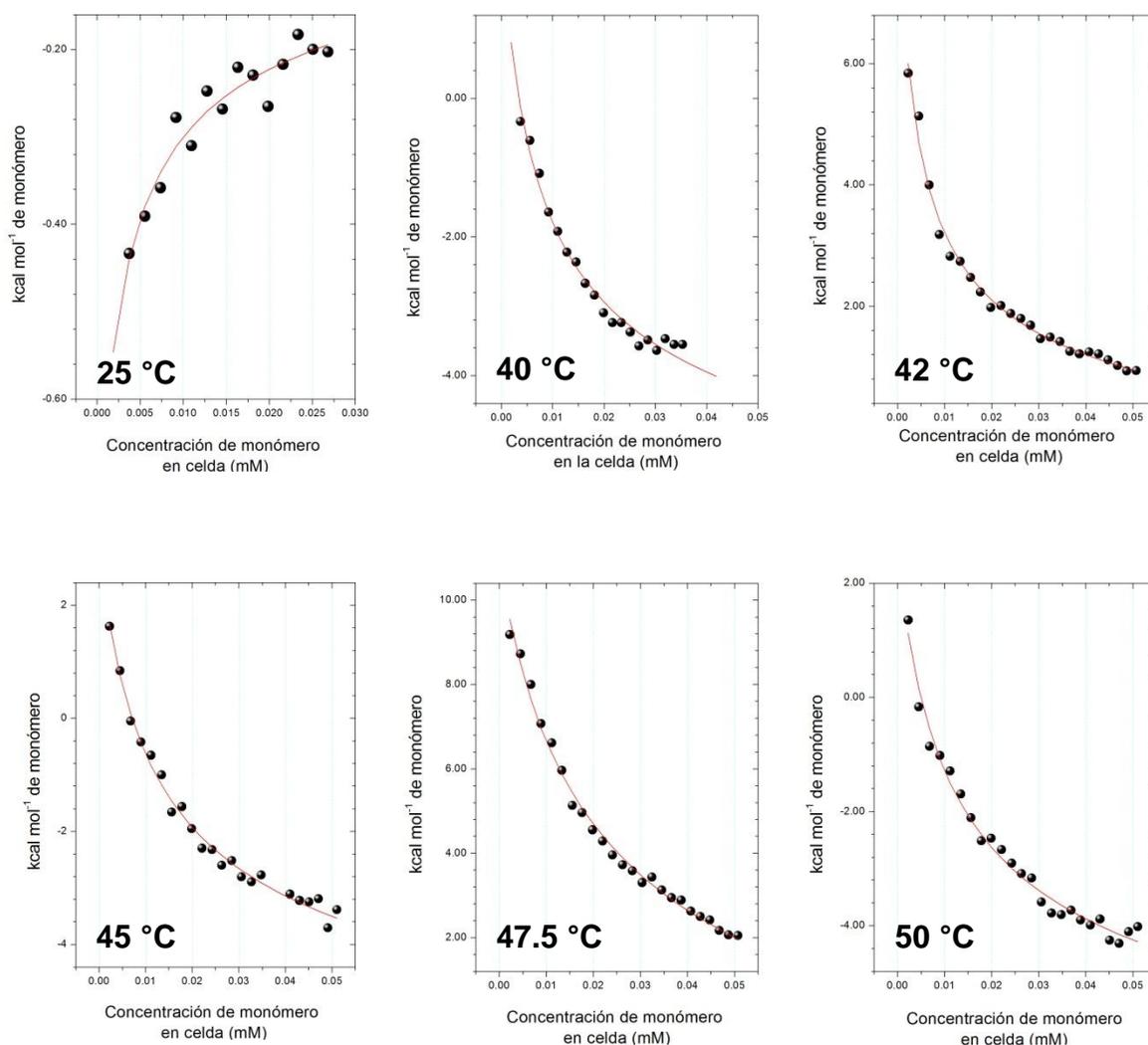


Figura 22. Caracterización termodinámica de la disociación de TIM63. Ajuste al modelo de disociación (línea continua) de los datos normalizados, restando la línea base e integrando cada pico. Se muestra la temperatura de cada experimento.

La curvatura hacia abajo de los datos en la figura 22 corresponde a un proceso de disociación endotérmico ($\Delta H_{\text{dis}} > 0$), mientras que lo contrario nos habla de un proceso exotérmico ($\Delta H_{\text{dis}} < 0$). El único experimento que presenta el segundo comportamiento es el que se realizó a 25°C, esto quiere decir que el proceso de asociación es endotérmico a esta temperatura, es decir, que los enlaces formados en la interfase del dímero son energéticamente más costosos que los que se

forman entre esos mismos grupos y el solvente. Para un dímero estable se espera un proceso de asociación exotérmico (comportamiento que sigue TIM63 en las temperaturas de 40°C o mayores), en donde los enlaces formados en el proceso de dimerización liberan energía y favorecen la estabilidad del oligómero. A alguna temperatura entre 25 y 40 °C el proceso de disociación es isoentálpico, esto quiere decir que a esa temperatura la energía de las interacciones que se rompen entre residuos es equivalente a la de las nuevas interacciones que se forman. Ese valor se puede obtener en el gráfico de ΔH contra la temperatura (fig. 23A) y corresponde a 26.6 °C.

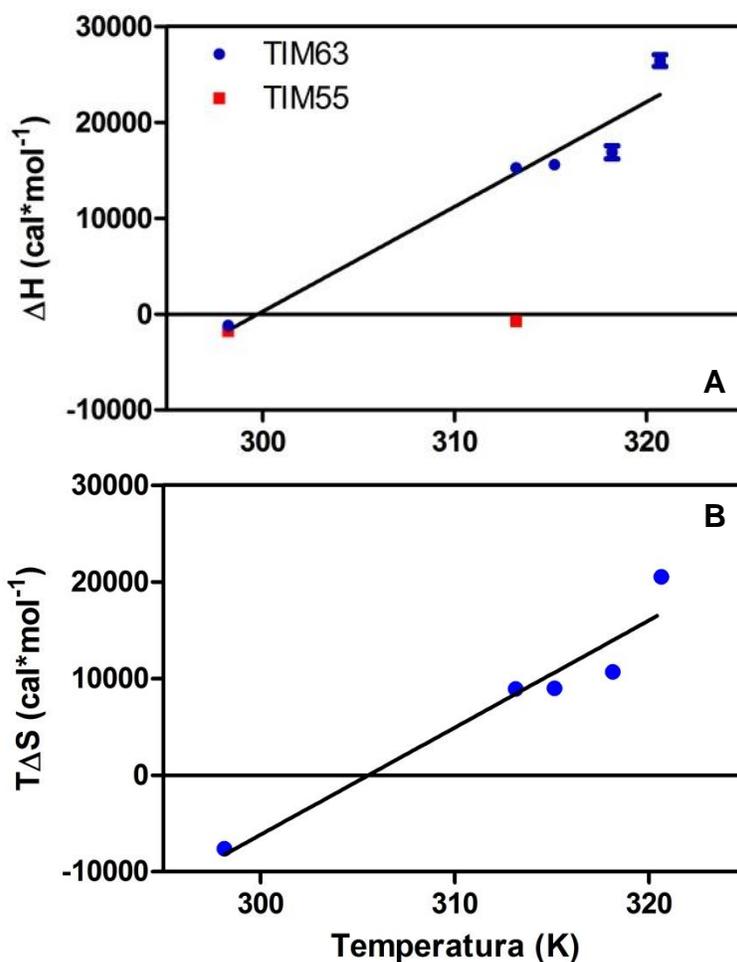


Figura 23. ΔC_p del proceso de disociación de TIM63. A) ΔH_{dis} vs T, $m = C_p = 1093 \pm 165.9$, B) $T\Delta S_{dis}$ vs T, $m = C_p = 1108 \pm 174.1$. En ambos casos la pendiente equivale a ΔC_p de disociación.

Tabla 10. ΔC_p de disociación para TIM63.

	$\text{Cal} \cdot \text{mol}^{-1} \cdot \text{K}^{-1}$
ΔC_{p_t}	480 – 840
$\Delta C_{p_{\text{exp}}}$	1093 ± 165.9

ΔC_{p_t} : teórico

$\Delta C_{p_{\text{exp}}}$: experimental

A partir de relaciones termodinámicas básicas (ver materiales y métodos) es posible obtener los valores para el ΔG y el $T\Delta S$. Un resumen de dichos valores se presenta en la tabla 11. La pendiente del gráfico de $T\Delta S$ contra T corresponde de nuevo al ΔC_p (fig. 23B), el valor obtenido de esta forma ($1108 \pm 174.1 \text{ cal} \cdot \text{mol}^{-1} \cdot \text{K}^{-1}$) concuerda con el anterior, lo que respalda los resultados reportados en la tabla 10.

Una de las características del proceso de disociación que resalta inmediatamente es la magnitud de la entalpía de disociación, ya que es muy pequeña, en especial complica el análisis de las isothermas porque muchas veces se enmascara con el ruido experimental. El ΔH obtenido para la disociación de la TIM de la amiba, en el que la disociación está conjugada con un desplegamiento parcial de los monómeros, es de 146.8 kcal/mol, sin embargo, cálculos teóricos a partir de simulaciones de cuerpo rígido para la disociación de un dímero en el que se exponen 74 residuos (37 por monómero) arroja valores entre 11 y 51 kcal/mol (basados en el ΔASA o en el número de residuos expuestos, respectivamente) (Téllez *et al.*, 2008). En todas las temperaturas empleadas ΔH_{dis} es menor a 30 kcal/mol. Esto sugiere que la disociación de la TIM63 no está acoplada a un desplegamiento generalizado de los monómeros. De manera alternativa, se ha demostrado que valores pequeños de ΔH pueden estar asociados a residuos específicos en la interfase del dímero que no estén correctamente orientados y por lo mismo no puedan formar interacciones estabilizantes (Robertson & Murphy, 1997). Esta interpretación concuerda con los valores de ΔC_p , es posible que el área enterrada en la interfase sea grande, pero que algún residuo en ella forma

una interacción desestabilizante y por lo mismo se necesita poca energía para la disociación.

Tabla 11. Parámetros termodinámicos del proceso de disociación para TIM63.

Temperatura (°C)	ΔG_{dis} (kcal/mol)	ΔH_{dis} (kcal/mol)	$T\Delta S_{\text{dis}}$ (kcal/mol)	K_{dis} (mM)
25	6.44	-1.20 ± 0.23	-7.56	0.02 ± 0.02
40	6.32	15.24 ± 0.55	-4.93	0.04 ± 0.01
42	6.59	15.57 ± 0.38	-5.18	0.03 ± 0.004
45	6.17	16.87 ± 0.69	-4.26	0.06 ± 0.01
47.5	5.93	26.43 ± 0.62	-3.49	0.09 ± 0.01
50	6.28	17.44 ± 0.46	-4.27	0.06 ± 0.01

Todos los datos obtenidos corresponden al proceso de disociación ya que es el que podemos manipular a partir de la dilución, sin embargo, para la claridad en la discusión de los datos se usará el proceso opuesto: la asociación. El único valor positivo para la entalpía de asociación es a 25°C, es decir, la asociación es entálpicamente desfavorable a esta temperatura. Para todas las otras temperaturas la asociación está favorecida entálpicamente. A pesar de esta diferencia, para todas las temperaturas el proceso está favorecido entrópicamente y la resultante global es que la asociación es favorable a todas las temperaturas. A partir de los datos en la tabla 11 y con los valores de ΔC_p obtenidos previamente en la tabla 10 se puede construir una curva de estabilidad de disociación para TIM63 (fig. 24). Los valores para ΔG_{dis} son muy pequeños en comparación a los obtenidos para TIMs extantes (15-20 Kcal/mol) (ver figura 5).

Por otra parte, al comparar los valores de la constante de disociación obtenidos a partir de los experimentos de inactivación por dilución y los obtenidos por ITC se observa que los primeros son 100 veces menores que los segundos, dicho de otra manera, TIM63 en el experimento de ITC presenta menor afinidad entre monómeros que en el experimento de inactivación por dilución. Una posible explicación para esta diferencia se puede deber a la presencia del sustrato. Como bien se sabe, la unión al sustrato estabiliza conformaciones funcionales, en el

caso de la TIM, dado que la enzima necesita la conformación dimérica para poder llevar a cabo la catálisis, se sugiere que la unión de G3P estabiliza al dímero y por lo mismo, la constante de disociación disminuye.

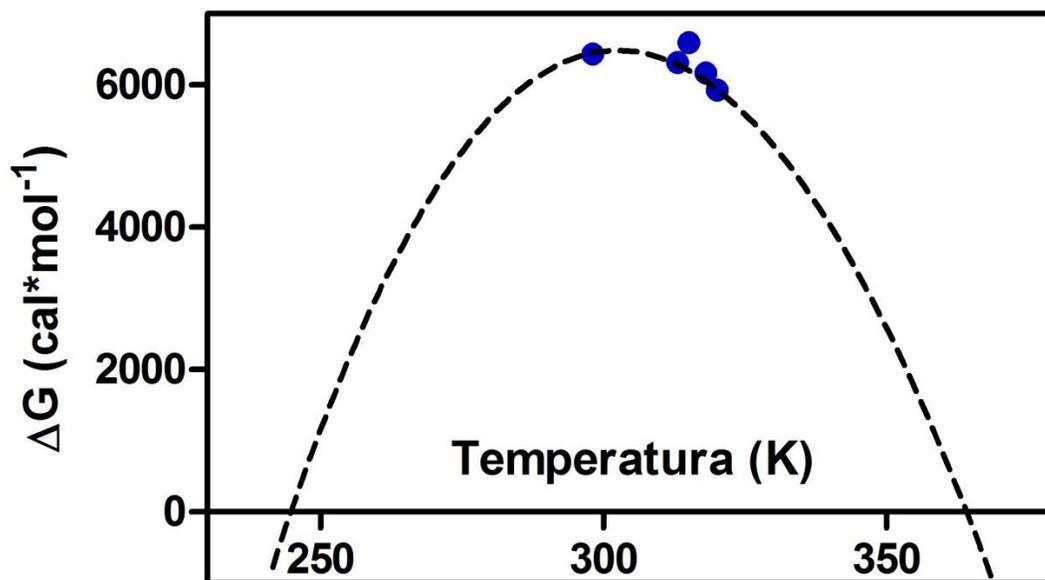


Figura 24. Curva de estabilidad a la disociación para TIM63. (●) Datos experimentales para ΔG_{dis} . Línea punteada: simulación de la curva energética de disociación a partir de la ecuación 16.

ITC para TIM55. Para TIM55 solamente fue posible realizar experimentos de ITC a dos temperaturas diferentes (25 y 40°C). En la figura 25 se presentan los termogramas y las integraciones de los mismos. Estos experimentos son aún más ruidosos que los de TIM63 y por lo mismo más difíciles de analizar y ajustar al modelo. Sin embargo, en la figura 23 se pueden comparar los valores de entalpía para ambos ancestros. Es importante notar que el proceso de disociación es exotérmico a 40°C para TIM55 a diferencia del proceso endotérmico para TIM63. Esto quiere decir que los enlaces de TIM55 de la interfase son energéticamente más débiles que los que se forman entre los residuos de la interfase en TIM63 o que es más favorable entálpicamente tener a los residuos de TIM55 solvatados.

Puesto de otra manera, a 40°C se necesita menos energía para disociar a TIM55 en comparación con TIM63.

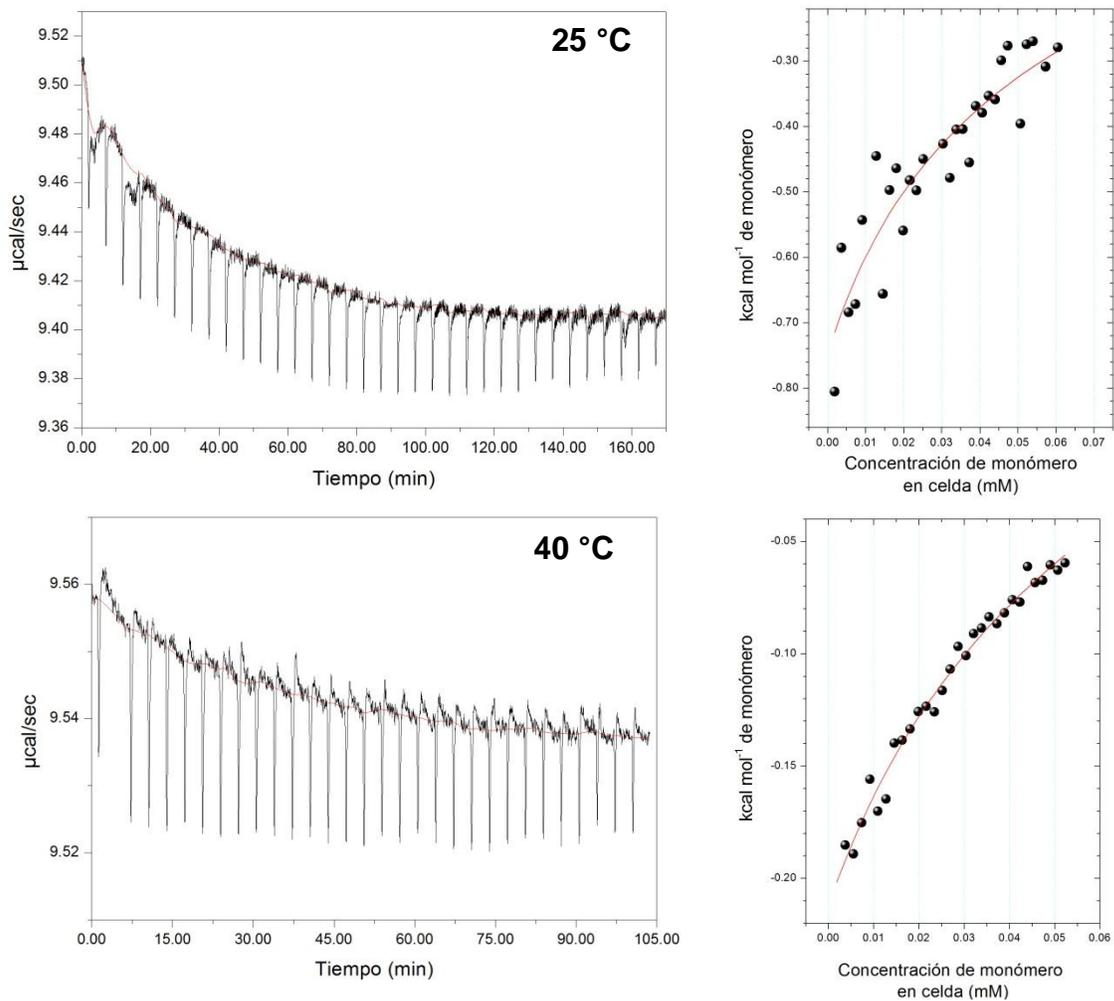


Figura 25. Experimentos de ITC para TIM55.

CONCLUSIONES Y PERSPECTIVAS

El plegamiento tipo barril TIM es una estructura muy versátil, esto se debe en gran parte a la capacidad de la estructura tridimensional para tolerar un considerable número de cambios en la secuencia; sin embargo, hay sitios que son altamente conservados y cualquier mutación en ellos lleva a la pérdida de actividad o de estabilidad (Banerjee *et al.*, 2009; Sampson & Knowles, 1992). A pesar de que la reconstrucción de los ancestros de la TIM se realizó manteniendo ciertas regiones conservadas, como la del sitio activo, una de las interrogantes principales era si las proteínas reconstruidas se plegarían. Nuestros resultados espectroscópicos sugieren que TIM63 y TIM55 cuentan con estructura secundaria y que el patrón de DC que presenta corresponde al de un barril $(\alpha/\beta)_8$ plegado. Además, de los espectros de fluorescencia se puede inferir que los triptófanos están ocultos en el núcleo hidrofóbico de la proteína, como debe de ser el caso en una estructura terciaria bien plegada. Las estructuras no parecen presentar parches hidrofóbicos extensos expuestos y los resultados de la cromatografía de exclusión molecular sugieren que se presentan principalmente como dímeros compactos.

Las altas temperaturas medias de desnaturalización para ambos ancestros indica una elevada termoestabilidad en comparación con TIMs extantes y TIMs estabilizadas por medio de diseño *in silico* y evolución dirigida. El proceso de desplegamiento está controlado cinéticamente y se describe por el modelo de dos estados irreversible de Lumry-Eyring con energías de activación altas y, por lo mismo, estados de transición energéticamente elevados. ¿A qué se debe esta alta energía de activación? Es algo que se tiene que investigar a fondo.

TIM63 es el único ancestro que tiene actividad catalítica apreciable, con una V_{max} comparable a la de TIMs extantes, sin embargo, la afinidad por el sustrato es baja, lo que resulta en una eficiencia catalítica baja. Esto nos podría decir que la actividad de la TIM no siempre ha sido “perfecta” como pasa en las enzimas extantes, sino que el tiempo evolutivo se ha encargado de refinarla. Dado que el valor para la actividad de TIM55 está en un rango en el que no se puede descartar

contaminación, no podemos confiar en bajo valor observado. Existen dos posibles experimentos que darían una respuesta clara y eliminarían la duda. El primero consiste en expresar y purificar a los ancestros en una cepa mutante en el gen endógeno de TIM (Δtim); la segunda opción es correr un gel de isoelectroenfoque y revelar la actividad directamente del gel. Es interesante destacar que los dos ancestros tienen una identidad de secuencia de casi 96 %, es decir, 11 aminoácidos diferentes de los cuales 8 son similares (fig. 8). Ninguno de estos aminoácidos corresponde a los componentes del sitio activo o a aminoácidos que participen en la estabilización del dímero, sin embargo, las diferencias que se presentan con respecto a la catálisis son evidentes y muy significativas.

Ambos ancestros presentan inestabilidad a la dilución que genera monómeros inactivos, aparentemente compactos. Los experimentos por ITC para TIM63 demuestran que la entalpía de disociación es desfavorable a temperaturas mayores a 40°C y presenta un valor pequeño aunque dentro del rango teórico, lo que sugiere que el dímero se disocia en monómeros compactos. El $\Delta C_{p_{dis}}$ experimental es un poco mayor que el teórico, hay que tomar en cuenta que el teórico se basa en el promedio del área oculta en la interfase de dímeros extantes, por lo que la diferencia de los valores puede tener dos implicaciones importantes: 1) si los monómeros se comportan como cuerpos rígidos en el proceso de disociación, entonces el área en la interfase de los dímeros ancestrales debería ser mayor que la de los dímeros extantes; esta opción es poco probable, o 2) el área expuesta en la disociación de los dímeros corresponde no solamente al área de contacto entre monómeros, sino que el proceso está acoplado a cierto desplegamiento de los mismos. La comparación entre $\Delta C_{p_{dis}}$ de TIM63 y $\Delta C_{p_{tot}}$ de las TIMs bacterianas (Romero-Romero, 2013) sugiere que si bien hay cierto grado de desplegamiento en la disociación, éste es pequeño en comparación con el desplegamiento de la proteína completa. Por un lado, el hecho de que $\Delta C_{p_{dis}}$ experimental sea cercano al teórico implica que el área no polar de TIM63 en la interfase es de un tamaño parecido al área no polar de las TIMs extantes; pero por otro lado, ΔH_{dis} experimental es menor al teórico, este comportamiento puede ser reflejo de un menor número de interacciones favorables en la interfase de TIM63

en comparación a las TIMs extantes o un mayor número de interacciones desfavorables en la TIM ancestral.

El cambio en entalpía favorece la disociación de TIM55 a temperaturas menores a 40°C. Una diferencia evidente entre TIM63 y TIM55 es que el proceso de asociación a 40°C está entálpicamente favorecido para la primera y desfavorecido para la segunda. Esto sugiere que las interacciones de la interfase en TIM55 son menos energéticas que para TIM63, lo que se refleja en una constante de disociación mayor para TIM55.

Normalmente las constantes de afinidad entre monómeros son tan altas en las TIMs extantes que las constantes de unión no se pueden medir, sin embargo, el estudio por ITC de la TIMs ancestrales abre la posibilidad de caracterizar el ciclo termodinámico completo de la enzima no solamente a partir de las constantes termodinámicas de unión (monómero-monómero o monómero/dímero-ligando), sino también de las constantes cinéticas (fig. 26). Estas últimas se podrían obtener a partir del análisis de las curvas de inactivación.

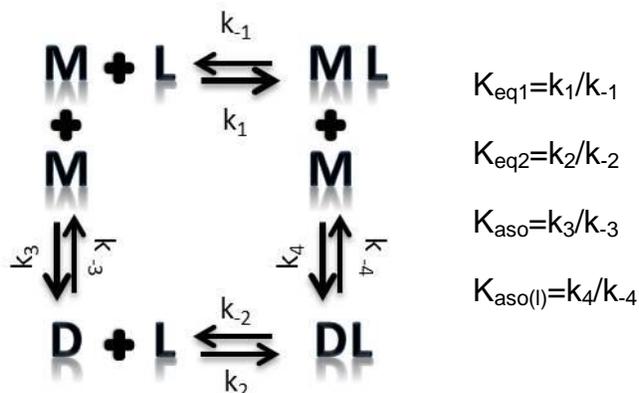


Figura 26. Ciclo termodinámico para la TIM. M: monómero, D: dímero, L: ligando, k: constantes cinéticas, K: constantes termodinámicas. k_1 y k_{-1} : constantes de unión entre monómero y ligando, k_2 y k_{-2} : constantes de unión entre dímero y ligando, k_3 y k_{-3} : constantes disociación del dímero sin y con ligando, respectivamente, k_4 y k_{-4} : constantes asociación del dímero sin y con ligando, respectivamente. K_{aso} y $K_{aso(l)}$: constantes de asociación sin y con ligando, respectivamente.

Todas las incógnitas presentadas anteriormente tienen un trasfondo estructural que es fundamental resolver para poder entender las determinantes fisicoquímicas que llevan a los comportamientos observados. Por ello, está en proceso la obtención de la estructura cristalográfica de ambos ancestros. Ya se han puesto pruebas de cristalización en el Laboratorio Nacional de Estructura de Macromoléculas (LANEM) y se han encontrado condiciones de cristalización que necesitan ser optimizadas. De esta manera será posible realizar una comparación estructural entre los ancestros así como entre ellos y las TIMs extantes para determinar cuáles son las diferencias e identificar a los residuos que forman las interacciones intra e intermonómeros.

Por último, falta llevar a cabo el trabajo de analizar las vías de plegamiento y desplegamiento de las proteínas ancestrales y compararlas con las de alguna proteína extante. Así podremos empezar a desenmarañar la historia de la evolución de las vías de plegamiento de la TIM.

Antes de aventurar cualquier conclusión con respecto a cómo eran las TIMs en el pasado, es necesario hacer un paréntesis reflexivo acerca del método de reconstrucción ancestral. Para empezar se debe tener en cuenta que todo el tiempo se está trabajando con una hipótesis evolutiva, es decir, las secuencias ancestrales obtenidas son las secuencias con mayor probabilidad de reconstrucción en un paisaje de muchas secuencias, esto no quiere decir que sean las únicas ni que se trate exactamente de la secuencia de la proteína que existió hace miles de años. Durante el proceso de reconstrucción se eligen un árbol filogenético y una matriz de sustitución (en el caso de los aminoácidos) que van a modelar el proceso evolutivo a partir de sus parámetros internos. Es un hecho que los modelos de sustitución no son perfectos y falta mucho para poder describir de manera exacta los diversos procesos evolutivos (Gaucher *et al.*, 2010). La verosimilitud de las secuencias se refiere a la probabilidad de observar dichas secuencias ancestrales dado el modelo (el árbol y la matriz). A pesar de que los valores de verosimilitud para las secuencias reconstruidas sean muy altos, la reconstrucción puede ser “errónea” si el modelo es demasiado simplista.

En la reconstrucción de los ancestros estudiados en esta tesis, se obtuvo una gran proporción (>60%) de sitios que se reconstruyen con una alta probabilidad posterior, sin embargo, éste valor es el producto de las probabilidades para cada uno de los sitios en la secuencia. Por ejemplo, para una proteína de 250 aminoácidos, si cada sitio se reconstruye con una probabilidad de 0.95, la probabilidad posterior de haber inferido la secuencia correcta es 0.95^{250} (2.7×10^{-6}). A pesar de lo dicho anteriormente suponemos que nuestras secuencias ancestrales son la mejor aproximación y serán un buen acercamiento al fenotipo ancestral (Thornton, 2004).

Por el momento, a partir de los resultados obtenidos en la presente tesis, se puede concluir que la catálisis de la enzima no siempre ha sido perfecta, pero siempre ha dependido de la dimerización, al menos en eucariontes, y a pesar de la inestabilidad a la dilución, la termoestabilidad de la estructura correlaciona con las altas temperaturas que pudieran existir en el tiempo del que son los ancestros (~1.37 mil millones de años, tiempo estimado para la divergencia entre animales y hongos). Un argumento que refuerza la idea de la inestabilidad del dímero es que los sitios con baja probabilidad de reconstrucción se encuentran principalmente en la parte externa del barril, lejos de la interfase entre monómeros o del sitio catalítico. La implicación evolutiva que tiene el haber encontrado una TIM activa pero inestable a la dilución sugiere que la evolución trabaja sobre andamios flexibles en cuestión de estabilidad para poder así ir dando forma a una enzima con mayor eficiencia catalítica y a su vez, mayor estabilidad estructural.

El trabajo aquí presentado tiene muchos caminos por donde seguir. El primero e inmediato es la continuación de la caracterización de los ancestros. Esto implica la determinación de las constantes termodinámicas y cinéticas presentadas en la figura 26 realizando ITC en presencia de ligandos; además queda por estudiar la inactivación por dilución de las proteínas a diferentes temperaturas para poder hacer una comparación mejor de las constantes de disociación; y, como ya se dijo, determinar la estructura tridimensional de los ancestros con y sin ligandos.

Otra continuación de este trabajo es llevar a cabo un segundo ciclo de reconstrucción, porque como ya se ha mencionado en repetidas ocasiones, las secuencias reconstruidas son las secuencias con mayor probabilidad de reconstrucción y esto no quiere decir que sean las únicas ni que se trate exactamente de la secuencia de la proteína que existió hace miles de años. Para hacer el análisis más robusto se podría hacer un segundo ciclo de reconstrucción con un mayor número de secuencias extantes. Por otra parte, hay posiciones en las secuencias de los ancestros en las que la probabilidad de que haya estado presente otro aminoácido también es muy alta; es necesario identificar esas posiciones y explorar qué tanto cambia la probabilidad del segundo aminoácido en cuestión. Esto cobra una mayor relevancia si se localizan esas posiciones en las estructuras de los ancestros y se determina si son las responsables de las “interacciones desestabilizantes”. En caso de ser así, una última alternativa es realizar un ciclo de diseño computacional con los ancestros presentes para estabilizarlos al permitir mutaciones específicas en los aminoácidos antes mencionados. Este último enfoque usa los datos termodinámicos, estructurales y evolutivos que se han obtenido y refuerza la idea de que una aproximación integrativa al entendimiento del comportamiento de las proteínas arroja mucha más información de la que se obtiene individualmente a partir de cada uno de los aspectos.

BIBLIOGRAFÍA

- Albery, W.J. & Knowles, J.R., 1976. Evolution of enzyme function and the development of catalytic efficiency. *Biochemistry*, 15(25), pp.5631–40.
- Alvarez, M., J. P. Zeelen, V. Mainfroid, F. Rentier-Delrue, J.A. Martial, L. Wyns, T. K. Wierenga and D. Maes, 1998. Triose-phosphate isomerase (TIM) of the psychrophilic bacterium *Vibrio marinus*. Kinetic and structural properties. *The Journal of biological chemistry* 273 (4), pp. 2199-206.
- Anfinsen, C. B., E. Haber, M. Sela, and F. H. White. 1961. The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain. *Proceedings of the National Academy of Sciences of the United States of America* 47, pp. 1309-14.
- Banerjee, M., H. Balaram, and P. Balaram, 2009. Structural effects of a dimer interface mutation on catalytic activity of triosephosphate isomerase. The role of conserved residues and complementary mutations. *The FEBS journal* 276 (15), pp. 4169-83.
- Ciccarelli, F.D., T. Doerks, C. Von Mering, C.J. Creevey, B. Snel, and P. Bork. 2006. Toward automatic reconstruction of a highly resolved tree of life. *Science* 311 (5765), p. 1283.
- Cisneros Armas, D.A., 2003. Termodinámica de la disociación de la triosafosfato isomerasa de *Entamoeba histolytica*. Tesis de licenciatura en Investigación Biomédica Básica.
- Costas, M., D. Rodríguez-Larrea, L. De Maria, T. V. Borchert, A. Gómez-Puyou, and J. M. Sanchez-Ruiz. 2009. Between-species variation in the kinetic stability of TIM proteins linked to solvation-barrier free energies. *Journal of molecular biology* 385 (3), pp. 924-37.
- Enríquez-Flores, S., 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, 2011. Determining the molecular mechanism of inactivation by chemical modification of triosephosphate isomerase from the human parasite *Giardia lamblia*: a study for antiparasitic drug design. *Proteins* 79 (9), pp. 2711-24.
- Farber, G.K. and Petsko, G.A., 1990. The evolution of alpha/beta barrel enzymes. *Trends in biochemical sciences*, 15(6), pp. 228–34.

- Gaucher, E. A., S. Govindarajan, and O. K. Ganesh, 2008. Palaeotemperature trend for Precambrian life inferred from resurrected proteins. *Nature*, 451(7179), pp. 704–7.
- Gaucher, E. A., J. T. Kratzer, and R. N. Randall, 2010. Deep phylogeny--how a tree can help characterize early life on Earth. *Cold Spring Harbor perspectives in biology* 2 (1), pp. a002238.
- Gokhale, R. S., S. S. Ray, H. Balaram, and P. Balaram. 1999. Unfolding of Plasmodium falciparum triosephosphate isomerase in urea and guanidinium chloride: evidence for a novel disulfide exchange reaction in a covalently cross-linked mutant. *Biochemistry* 38 (1), pp. 423-31.
- Guzmán-Luna, V., 2006. *Participación de las cisteínas en la estabilidad y estructura de la triosafosfato isomerasa de Trypanosoma Brucei*. Universidad Nacional Autónoma de México.
- Knobeloch, D., A. Schmidt, P. Scheerer, N. Krauss, H. Wessner, Ch. Scholz, G. Küttner, T. von Rintelen, A. Wessel and W. Höhne. 2010. A coleopteran triosephosphate isomerase: X-ray structure and phylogenetic impact of insect sequences. *Insect molecular biology* 19 (1), pp. 35-48.
- Knowles, J. R., and W. J. Albery. 1977. Perfection in enzyme catalysis: the energetics of triosephosphate isomerase. *Accounts of Chemical Research* 10(4), pp. 105-111.
- Liberles, D., 2007. Ancestral Sequence Reconstruction. *Annals of Physics*. First Edit. Vol. 54. Oxford University Press.
- López-Velázquez, G., D. Molina-Ortiz, N. Cabrera, G. Hernández-Alcántara, J. Peon-Peralta, L. Yépez-Mulia, R. Pérez-Montfort, and Horacio Reyes-Vivas, 2004. An unusual triosephosphate isomerase from the early divergent eukaryote Giardia lamblia. *Proteins* 55 (4), pp. 824-34.
- Mainfroid, V., P. Terpstra, M. Beauregard, J. M. Frère, S. C. Mande, W. G. Hol, J. a Martial, and K. Goraj, 1996. Three hTIM mutants that provide new insights on why TIM is a dimer. *Journal of molecular biology* 257 (2), pp. 441-56.
- Maithal, K., G. Ravindra, G. Nagaraj, S. Kumar Singh, H. Balaram, and P. Balaram, 2002. Subunit interface mutation disrupting an aromatic cluster in Plasmodium falciparum triosephosphate isomerase: effect on dimer stability. *Protein engineering* 15 (7), pp. 575-84.
- Morgan, C. J., D. K. Wilkins, L. J. Smith, Y. Kawata, and C. M. Dobson, 2000. A compact monomeric intermediate identified by NMR in the denaturation of dimeric triose phosphate isomerase. *Journal of molecular biology* 300 (1), pp. 11-6.

- Nagano, N., C. A. Orengo and J. M. Thornton, 2002. One Fold with Many Functions: The Evolutionary Relationships between TIM Barrel Families Based on their Sequences, Structures and Functions. *Journal of Molecular Biology*, 321(5), pp.741–765.
- Nájera, H., M. Costas, and D. A. Fernández-Velasco, 2003. Thermodynamic characterization of yeast triosephosphate isomerase refolding: insights into the interplay between function and stability as reasons for the oligomeric nature of the enzyme. *The Biochemical journal*, 370(Pt 3), pp.785–92.
- Ortlund, E. a, J. T. Bridgham, M. R. Redinbo, and J. W. Thornton, 2007. Crystal structure of an ancient protein: evolution by conformational epistasis. *Science (New York, N.Y.)* 317 (5844), pp. 1544-8.
- Page, R.D. and E. Holmes, 1998. *Molecular Evolution. A Phylogenetic Approach*, Blackwell Science.
- Pardo-Ávila, F.A., 2011. *Evolución de la Triosafosfato Isomerasa*. Tesis de licenciatura en Investigación Biomédica Básica. Universidad Nacional Autónoma de México.
- Peimbert, M., L. Domínguez-Ramírez and D. A. Fernández-Velasco, 2008. Hydrophobic repacking of the dimer interface of triosephosphate isomerase by in silico design and directed evolution. *Biochemistry*, 47(20), pp. 5556–64.
- Perez-Jimenez, R., A. Inglés-Prieto, Zi-Ming Zhao, I. Sanchez-Romero, J. Alegre-Cebollada, P. Kosuri, S. Garcia-Manyes, T. J. Kappock, M. Tanokura, A. Holmgren, J. M. Sanchez-Ruiz, E. a Gaucher and J. M. Fernandez, 2011. Single-molecule paleoenzymology probes the chemistry of resurrected enzymes. *Nature Structural & Molecular Biology* (January) (April): 1-6.
- Pompliano, D.L., A. Peyman and J. R. Knowles, 1990. Stabilization of a reaction intermediate as a catalytic device: definition of the functional role of the flexible loop in triosephosphate isomerase. *Biochemistry*, 29(13), pp.3186–94.
- Reyes-Vivas, H., A. Diaz, J. Peon, G. Mendoza-Hernandez. G. Hernandez-Alcantara, I. De la Mora-De la Mora, S. Enriquez-Flores, L. Dominguez-Ramirez, and G. Lopez-Velazquez, 2007. Disulfide bridges in the mesophilic triosephosphate isomerase from *Giardia lamblia* are related to oligomerization and activity. *Journal of molecular biology* 365 (3), pp. 752-63.
- Robertson, A.D. and K. P. Murphy, 1997. Protein Structure and the Energetics of Protein Stability. *Chemical reviews*, 97(5), pp.1251–1268.
- Romero-Romero, S., 2013. Control cinético y termodinámico en el desplegamiento térmico de la triosafosfato isomerasa proveniente de especies bacterianas.

Tesis de maestría en Ciencias Bioquímicas. Universidad Nacional Autónoma de México.

- Sampson, N. S., and J. R. Knowles. 1992. Segmental movement: definition of the structural requirements for loop closure in catalysis by triosephosphate isomerase. *Biochemistry* 31 (36), pp. 8482-7.
- Sampson, N. S., and J. R. Knowles. 1992. Segmental motion in catalysis: investigation of a hydrogen bond critical for loop closure in the reaction of triosephosphate isomerase. *Biochemistry* 31 (36), pp. 8488-94.
- Sanchez-Ruiz, J. M., 1992. Theoretical analysis of Lumry-Eyring models in differential scanning calorimetry. *Biophysical journal*, 61(4), pp.921–35.
- Schliebs, W., N. Thanki, R. Eritja, and R. Wierenga. 1996. Active site properties of monomeric triosephosphate isomerase (monoTIM) as deduced from mutational and structural studies. *Protein science : a publication of the Protein Society* 5 (2), pp. 229-39.
- Schliebs, W., N. Thanki, R. Jaenicke, and R. K. Wierenga, 1997. A double mutation at the tip of the dimer interface loop of triosephosphate isomerase generates active monomers with reduced stability. *Biochemistry* 36 (32), pp. 9655-62.
- Silverman, J.A., R. Balakrishnan, and P. B. Harbury,, 2001. Reverse engineering the (beta/alpha)8 barrel fold. *Proceedings of the National Academy of Sciences of the United States of America*, 98(6), pp.3092–7.
- Téllez, L. A., L. M. Blancas-Mejia, E. Carrillo-Nava, G. Mendoza-Hernández, D. A. Cisneros, and D. A. Fernández-Velasco, 2008. Thermal unfolding of triosephosphate isomerase from *Entamoeba histolytica*: dimer dissociation leads to extensive unfolding. *Biochemistry* 47 (44), pp. 11665-73.
- Téllez, L. A., L. M. Blancas, A. D. Fernández-Velasco, M. Peimbert, and L. Domínguez-Ramírez. 2008. Triosephosphate Isomerase. *Advances in Protein Physical Chemistry*, 661, pp. 193-217.
- Thornton, J. W., 2004. Resurrecting ancient genes: experimental analysis of extinct molecules. *Nature reviews. Genetics* 5(5), pp. 366-75.
- Ugalde, J. A., B. S. W. Chang, and M. V. Matz, 2004. Evolution of coral pigments recreated. *Science (New York, N.Y.)*, 305(5689), p.1433.
- Uversky, V.N., 1993. Use of fast protein size-exclusion liquid chromatography to study the unfolding of proteins which denature through the molten globule. *Biochemistry*, 32(48), pp.13288–98.
- Voet, D. & Voet, J.G., 1995. *Biochemistry* 2nd ed., John Wiley & Sons, Inc.

- Wierenga, R., 2001. The TIM-barrel fold: a versatile framework for efficient enzymes. *FEBS Letters*, 492(3), pp.193–198.
- Zuckerklund, E. and L. Pauling, 1965. Molecules as documents of evolutionary history. *Journal of theoretical biology*, 8(2), pp.357–66.

APÉNDICE I

CATH

www.cathdb.info

La base de datos del CATH se basa en una clasificación jerárquica de los dominios de las estructura de proteínas depositadas en el PDB (Protein Data Bank). Las estructuras se clasifican semi automáticamente. Los cuatro niveles de jerarquía son los siguientes:

- **Clase** – clasificación de acuerdo a la composición de estructura secundaria:
 - mayoritariamente alfa
 - mayoritariamente beta
 - alfa/beta mezclada o
 - poca estructura secundaria.
- **Arquitectura** – clasificación conforme a la silueta determinada por la orientación de los elementos de estructura secundaria en el espacio 3D (ignora la conectividad).
- **Topología (familia de plegamiento)** – la estructuras se agrupan según su plegamiento dependiente tanto de la forma en 3D como la conectividad de la estructura secundaria.
- **Superfamilia Homóloga** – en este nivel se agrupan los dominios proteicos que probablemente comparten un ancestro común y, por lo mismo, se clasifican como homólogos.

-

Para los barriles TIM la clasificación es la siguiente:

- Clase: alfa/beta
- Arquitectura: barril alfa/beta
- Topología: barril TIM

- 29 superfamilias homólogas (6 no se han nombrado y no aparecen en la lista):

1. Alanine racemase
2. *Dihydropteroate (DHP) synthetase*
3. *FMN dependent fluorescent proteins*
4. *7-stranded glycosidases (cellulases)*
5. *Phosphoenolpyruvate-binding domains*
6. *Aldolase class I*
7. Glycosidases
8. *NADP-dependent oxidoreductase*
9. *tRNA-guanine (tRNA-G) transglycosylase*
10. *Rubisco*
11. *Enolase superfamily*
12. *Metal-dependent hydrolases*
13. *Divalent-metal-dependent TIM barrel enzymes*
14. *Phosphatidylinositol (PI) phosphodiesterase*
15. *Malate synthase, domain 3*
16. *Glycoside hydrolase/deacetylase*
17. *CutC-like (Pfam 03932)*
18. *FMN-linked oxidoreductases*
19. *GlpP-like*
20. *Hypothetical protein tm1631.*
21. *D-lysine 5,6-aminomutase alpha subunit. Chain A*
22. *EAL domain-like*
23. *Monomethylamine methyltransferase MtmB*

SCOP

<http://scop.mrc-lmb.cam.ac.uk/scop/index.html>

La base de datos SCOP fue creada a partir de inspección manual y curada por una batería de métodos automatizados. Su fin es proveer de una descripción detallada y comprensiva acerca de las relaciones estructurales y evolutivas entre todas las proteínas de las que se conoce la estructura.

La clasificación jerárquica intenta reflejar tanto la relación estructural entre proteínas como la evolutiva. Los principales niveles de clasificación son familia, superfamilia y plegamiento. Las fronteras exactas entre los niveles son hasta cierto punto subjetivas. La clasificación evolutiva es en general conservativa: si existe duda acerca de la existencia de relación, se generan nuevas divisiones en los niveles de familia y superfamilia.

- **Familia:** *Relación evolutiva clara*

Generalmente esto significa que la identidad de secuencia entre proteínas es del 30% o mayor. Sin embargo, en algunos casos, a pesar de un bajo valor para la identidad de secuencia, funciones y estructuras similares proveen evidencia definitiva de un linaje evolutivo común.

- **Superfamilia:** *probable origen evolutivo común*

Aquellas proteínas que tienen bajas identidades de secuencias, pero cuyas características funcionales y estructurales sugieren un origen evolutivo común se agrupan en superfamilias.

- **Plegamiento:** *semejanza estructural*

Se dice que las proteínas tienen el mismo plegamiento cuando los elementos de estructura secundaria están en el mismo arreglo espacial y con las mismas conexiones topológicas. Proteínas diferentes con el mismo plegamiento muy a menudo tienen elementos de estructura secundaria periféricos y regiones de giros que se diferencian en tamaño y conformación. Proteínas en la misma categoría de plegamiento pueden no tener un origen evolutivo común.