



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Maestría y Doctorado en Ciencias Bioquímicas

Estudio de la estabilidad cinética y la flexibilidad térmica de proteínas usando como modelo las triosafosfato isomerasas de dos protozoarios cercanos en la evolución.

T E S I S

QUE PARA OPTAR POR EL GRADO DE:

DOCTORA EN CIENCIAS

P R E S E N T A

Q. ANDREA ELIZABETH GUTIÉRREZ QUEZADA

Tutor principal:

Dr. Miguel Antonio Costas Basín
Facultad de Química - UNAM

Comité Tutor:

Dr. Daniel Alejandro Fernández Velasco
Facultad de Medicina - UNAM

Dr. Ruy Pérez Montfort
Instituto de Fisiología Celular - UNAM

CIUDAD DE MÉXICO. FEBRERO, 2017.



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.

Reconocimientos

Deseo expresar mi reconocimiento y gratitud al Dr. Miguel Costas bajo cuya dirección se llevó a cabo este trabajo en el laboratorio de Biofísicoquímica de la Facultad de Química de la Universidad Nacional Autónoma de México.

Al Dr. Ángel Piñeiro Guillén por su ayuda en la planeación, lanzamiento y análisis de las simulaciones de dinámica molecular durante mi estancia de investigación en la Facultad de Ciencias Físicas de la Universidad de Santiago de Compostela, así como por proporcionarme varios de los códigos que sirvieron tanto para el análisis de desplazamientos de residuos como para el ajuste de las trazas calorimétricas.

Al Dr. Ruy Pérez Montfort en cuyo laboratorio se llevó a cabo la parte de biología molecular y purificación de las proteínas y que formó parte de mi comité tutorial.

Al Dr. Alejandro Fernández Velasco por sus valiosos comentarios en los tutoriales.

A la Biol. Elizabeth Nallely Cabrera González por su asistencia técnica en diversos aspectos concernientes a la clonación, purificación y determinación de la actividad específica de las enzimas.

A la M. en C. Alma Jessica Díaz Salazar quien determinó las constantes catalíticas presentadas en este trabajo.

Al Dr. José Manuel Sánchez Ruiz quien me abrió las puertas de su laboratorio en la Facultad de Ciencias de la Universidad de Granada durante la etapa inicial de este proyecto.

El Centro de Supercomputo de Galicia (CESGA) otorgó tiempo de computo para la realización de las simulaciones de dinámica molecular.

El Programa de Apoyo a la Investigación y el Posgrado de la Facultad de Química de la UNAM (PAIP 500-9018) otorgó recursos para el desarrollo de este proyecto.

Obtuve ayuda financiera por parte del Programa de Apoyo a los Estudios del Posgrado (PAEP) de la UNAM para realizar una estancia de investigación en la Facultad de Ciencias de la Universidad de Granada y otra en la Facultad de Ciencias Físicas de la Universidad de Santiago de Compostela, España. En ambos casos también obtuve un respaldo complementario por parte del programa de Becas Mixtas otorgado por el CONACyT.

Durante mis estudios de maestría/doctorado conté con una beca del CONACyT (No. de becario 226014), con una beca de ayudante de investigador por parte del SNI/CONACyT y con el apoyo de los proyectos DGAPA-UNAM-PAPIIT IN112813 e IN204616.

Jurado asignado:

Presidente	Dr. Enrique Rudiño Piñera
Vocal	Dra. Marietta Tuena Sangri
Vocal	Dr. Enrique García Hernández
Vocal	Dr. Rafael Arturo Zubillaga Luna
Secretario	Dr. Alfredo Torres Larios

Parte de este trabajo fue presentado en los siguientes congresos:

- “Determinación de la estabilidad cinética de algunas enzimas quiméricas construidas a partir de las triosafosfato isomerasas de *Trypanosoma cruzi* y *Trypanosoma brucei*” (Cartel). Andrea E. Gutiérrez Quezada y Miguel Costas. XXVIII Congreso de la Sociedad Mexicana de Bioquímica, Tuxtla Gutiérrez, Chiapas. Noviembre del 2010.
- “Kinetic stability determination of two closely related Trypanosomatidae’s TIM chimeric enzymes” (Cartel). Andrea E. Gutiérrez Quezada y Miguel Costas. 3er Congreso de la Rama de Fisicoquímica de Proteínas, México DF. Noviembre del 2011.
- “Simulaciones de dinámica molecular de desplegamiento inducido por temperatura de las triosafosta isomerasas de *Trypanosoma brucei* y *Trypanosoma cruzi*” (Cartel). Andrea E. Gutiérrez Quezada, Ángel Piñero Guillén y Miguel Costas. XXIX Congreso de la Sociedad Mexicana de Bioquímica; Oaxaca, Oaxaca. Noviembre del 2012.
- “The Structural Basis of Protein Kinetic Stability: the Trypanosomatidae’s TIMs case” (Ponencia). Andrea E. Gutiérrez Quezada y Miguel Costas. 4to. Congreso de la Rama de Fisicoquímica de Proteínas, Guanajuato, Guanajuato. Noviembre del 2013.

Agradecimientos

Quiero agradecer con respeto y cariño a todos aquellos que de muy diversas maneras me ayudaron a concluir este mandala/laberinto que fue el doctorado:

A Miguel Costas.

A los compañeros de laboratorio.

A mis alumnos de la Facultad de Química.

A los compañeros de running y natación.

A los compañeros de bicicleta.

A la banda del náhuatl (gracias por negarse a dejar que el mundo siga encogiéndose).

A los compañeros de la montaña.

A mis maestros y compañeros de danza.

A Syama y a Mahendra.

A mis ex-roomies.

A mis amigos.

A mis tios y primos.

A Pilar y a Jesús.

A Rosa María, a Rafael y a Esteban.

A Magos y a Luna.

A Mónica.

A la Universidad, a toda la gente que la ha hecho posible.

Y a los dioses que sí existen.

*Cuando emprendas tu viaje hacia Ítaca
debes rogar que el viaje sea largo,
lleno de peripecias, lleno de experiencias.*

*No has de temer ni a los lestrigones, ni a los cíclopes,
ni la cólera del airado Poseidón.*

*Nunca tales monstruos hallarás en tu ruta
si tu pensamiento es elevado, si selecta
es la emoción que toca tu espíritu y tu cuerpo.*

*Los lestrigones y los cíclopes
y el feroz Poseidón no podrán encontrarte
si tú no los llevas ya dentro, en tu alma,
si tu alma no los conjura ante ti.*

*Debes rogar que el viaje sea largo,
que sean muchos los días de verano;
que te vean arribar con gozo, alegremente
a puertos que tú antes ignorabas.*

*Que puedas detenerte en los mercados de Fenicia,
y comprar unas bellas mercancías.*

*Acude a muchas ciudades de Egipto
para aprender, y aprender de quienes saben.*

*Conserva siempre en tu alma la idea de Ítaca,
llegar allí, he aquí tu destino.*

*Mas no hagas con prisas tu camino;
mejor será que dure muchos años,
y que llegues, ya viejo, a la pequeña isla,
rico de cuanto habrás ganado en el camino.*

*No has de esperar que Ítaca te enriquezca,
Ítaca te ha concedido ya un hermoso viaje.*

*Sin ella, jamás habrías partido;
mas no tiene otra cosa que ofrecerte.*

*Y si la encuentras pobre, Ítaca no te ha engañado.
Y así, sabio como te has vuelto, con tanta experiencia,
entenderás ya qué significan las Ítacas.*

Konstantinos Kavafis

Índice General

1	Introducción	1
1.1	Estabilidad cinética de proteínas	1
1.2	La triosafosfato isomerasa: un barril TIM	6
1.3	Las triosafosfato isomerasas de <i>Trypanosoma cruzi</i> (TcTIM) y <i>Trypanosoma brucei</i> (TbTIM)	8
1.4	Flexibilidad de proteínas	9
2	Métodos	14
2.1	Diseño, clonación, expresión y purificación de las quimeras de TbTIM/TcTIM	14
2.2	Actividad enzimática	15
2.3	Calorimetría diferencial de barrido	16
2.4	Simulaciones de dinámica molecular	17
2.4.1	Cajas de simulación y parametrización	17
2.4.2	Análisis estándar de las trayectorias	18
2.4.3	Desplazamientos de residuos y su dependencia con temperatura (TGD)	18
2.4.4	Estimación de TGD para las quimeras y contribución de cada residuo al TGD de las silvestres	22
3	Resultados y discusión	24
3.1	Caracterización de las quimeras de TbTIM/TcTIM	24
3.1.1	Determinación de la actividad catalítica	26
3.1.2	Desnaturalización térmica caracterizada por DSC	26

3.2	Simulaciones de dinámica molecular de desplegamiento por temperatura . . .	38
3.2.1	Agrupación conformacional de las estructuras en clústers	45
3.2.2	Desplazamiento de residuos y regiones	47
3.2.3	Dependencia con temperatura de los desplazamientos (TGD)	49
3.3	Flexibilidad térmica y estabilidad cinética	54
3.3.1	Definición de flexibilidad térmica	54
3.3.2	Correlación entre la estabilidad cinética y la flexibilidad térmica . . .	55
3.4	Contribución de cada residuo a la diferencia en flexibilidad térmica entre TbTIM y TcTIM	61
4	Conclusiones	65
5	Perspectivas	66
	Bibliografía	68
	Apéndices	83
	Apéndice A Artículo publicado	84
	Apéndice B DSC de otras quimeras TbTIM/TcTIM	99
	Apéndice C Análisis general de las simulaciones a 298 K	103
	Glosario	108

Índice de Figuras

1.1	Calorimetría diferencial de barrido	6
1.2	TbTIM y TcTIM: alineamiento estructural y trazas calorimétricas	9
1.3	Flexibilidad estática y Flexibilidad dinámica	11
1.4	Caja de simulación	13
2.1	Análisis dimérico y análisis monomérico de los desplazamientos	20
2.2	Procedimiento del análisis de desplazamiento de residuos	22
3.1	Separación de la estructura del barril TIM en ocho regiones	25
3.2	Comparación entre las eficiencias catalíticas de varias TIMs silvestres y las quimeras	26
3.3	Trazas de DSC y gráficos de Arrhenius: TbTIM, TcTIM, Tc(5)Tb(1-4,6-8), Tc(1-4,6-8)Tb(5)	30
3.4	Trazas de DSC y gráficos de Arrhenius: Tc(1-6)Tb(7-8), Tc(1-5)Tb(6-8), Tc(1-4)Tb(5-8), Tc(1-3)Tb(4-8)	31
3.5	Trazas de DSC y gráficos de Arrhenius: Tc(1-2)Tb(3-8), Tc(1)Tb(2-8), Tc(2)Tb(2-8), Tc(4)Tb(1-3,5-8)	32
3.6	Trazas de DSC y gráficos de Arrhenius: Tc(1-3,5-8)Tb(4), Tc(2,5)Tb(1,3-4,6-8), Tc(3)Tb(1-2,4-8), Tc(2-8)Tb(1)	33
3.7	Trazas de DSC y gráficos de Arrhenius: Tc(5-6)Tb(1-4,7-8), Tc(6)Tb(1-5,7-8), Tc(1,3,5-8)Tb(2,4), Tc(5-8)Tb(1-4)	34
3.8	Trazas de DSC y gráficos de Arrhenius: Tc(6-8)Tb(1-5), Tc(2,4)Tb(1,3,5-8), Tc(4-5)Tb(1-3,6-8), Tc(1,7)Tb(2-6,8)	35

3.9	Trazas de DSC y gráficos de Arrhenius: Tc(2-6,8)Tb(1,7), Tc(1,4)Tb(2,3,5-8), Tc(1-3,5-8)Tb(4,5), Tc(2,4,5)Tb(1,2,6-8)	36
3.10	Trazas de DSC y gráficos de Arrhenius: Tc(2,3,5-8)Tb(1,4), Tc(2,3,5)Tb(1,4,6- 8), Tc(1,3,6-8)Tb(2,4,5)	37
3.11	Imágenes de las estructuras de TbTIM y TcTIM a diferentes tiempos de simulación	40
3.12	Radio de giro	41
3.13	Puentes de hidrógeno entre subunidades	42
3.14	Puentes de hidrógeno totales	42
3.15	Área accesible al solvente	43
3.16	Desviación cuadrática media (RMSD) de los $C\alpha$	44
3.17	Cambios en la estructura secundaria	46
3.18	Proyección de la matriz de $C\alpha$ -RMSD en 3D	47
3.19	Dispersión progresiva de movilidad (PMD)	48
3.20	Desplazamiento de residuos ($\Delta d_{j,c}$)	50
3.21	Desplazamiento por regiones ($D_{r,T}$)	51
3.22	Desplazamiento total por proteína ($D_{wt,T}$) para TbTIM y TcTIM	51
3.23	Desplazamiento total por región ($D_{r,T}$)	52
3.24	Dependencia con la temperatura de los desplazamientos por región (TGD_r) .	52
3.25	Efecto del intercambio de las regiones 4 y 5 en la E_{act} de TbTIM y TcTIM .	53
3.26	Unidades de desplegamiento de TbTIM y TcTIM	54
3.27	Correlación entre la estabilidad cinética y la flexibilidad térmica	57
3.28	Correlación entre la cooperatividad de desplegamiento y la flexibilidad térmica	58
3.29	T_m de las quimeras en función de TGD	60
3.30	Representación esquemática del efecto en de la flexibilidad estática y la flexibi- lidad térmica en la estabilidad cinética de TbTIM y TcTIM	61
3.31	Comparación de los valores de TGD_{res} entre TbTIM y TcTIM	63
B.1	Trazas de DSC y gráficos de Arrhenius: Tc(1,3-8)Tb(2), Tc(1,3,4,6-8)Tb(2,5), TbTIM+V61Q/M122H, TbTIM+S43P/V46L/A49P	100

B.2	Trazas de DSC y gráficos de Arrhenius: TbTIM+9mut, TbTIM+5mut, TbTIM+5mut*, TbTIM+7mut	101
B.3	Trazas de DSC y gráficos de Arrhenius: TbTIM+A69T/P81Q, TbTIM+12mut	102
C.1	Estructuras de TbTIM y TcTIM después de 400 ns de simulación a 298 K	103
C.2	RMSD de TbTIM y TcTIM a 298 K	104
C.3	RMSF de TbTIM y TcTIM a 298 K	104
C.4	Estructura secundaria de TbTIM y TcTIM a 298 K	105
C.5	Puentes de hidrógeno de TbTIM y TcTIM a 298 K	106
C.6	Radio de giro de TbTIM y TcTIM a 298 K	107
C.7	Área accesible al solvente de TbTIM y TcTIM a 298 K	107

Índice de Tablas

3.1	Residuos que comprenden las ocho regiones del barril TIM	25
3.2	Caracterización de las quimeras de TbTIM/TcTIM	27
3.3	Eficiencia catalítica de TIMs silvestres	28
3.4	Residuos de la intercara de TbTIM y TcTIM	38
3.5	Residuos de TcTIM con valores de TGD_{res} significativamente mayores que TbTIM	64
B.1	DSC de otras quimeras de TbTIM/TcTIM	99

Resumen

La estabilidad cinética es un concepto que está ligado al tiempo promedio en que una proteína mantiene su estructura y función bajo ciertas condiciones. Como tal, su estudio es esencial para esbozar un panorama completo del comportamiento de las proteínas y es importante para comprender cómo la evolución ha alcanzado el sutil balance entre estructura y función en tiempos relevantes para la célula. Utilizando una estrategia que incluye simulaciones, ingeniería de proteínas y calorimetría, en este trabajo se demostró que existe una correlación entre la estabilidad cinética, determinada a través de calorimetría diferencial de barrido y un parámetro que hemos llamado flexibilidad térmica, obtenido con una metodología nueva basada en simulaciones de dinámica molecular de alta temperatura. La flexibilidad térmica mide cuantitativamente el incremento del espacio conformacional accesible a la proteína cuando se le suministra energía. El barril $(\beta/\alpha)_8$ de las triosafosfato isomerasas de los parásitos evolutivamente cercanos *Trypanosoma brucei* y *Trypanosoma cruzi* fue utilizado como modelo de estudio. La correlación entre la estabilidad cinética y la flexibilidad térmica tiene poder predictivo para las proteínas estudiadas, lo que abre la posibilidad de aplicar esta estrategia y la metodología discutida aquí a otras proteínas de interés biotecnológico, terapéutico y evolutivo.

Abstract

Kinetic stability is a key parameter to comprehend protein behaviour and it plays a central role to understand how evolution has reached the balance between function and stability in cell-relevant time scales. Using an approach that includes molecular dynamics simulations, protein engineering and calorimetry, we show that there is a clear correlation between kinetic stability determined by differential scanning calorimetry, and protein thermal flexibility obtained from a novel method based on temperature-induced unfolding molecular dynamic simulations. Thermal flexibility quantitatively measures the increment of the conformational space available to the protein when energy is provided. The $(\beta/\alpha)_8$ barrel fold of two closely related by evolution triosephosphate isomerases from two trypanosomes are used as model systems. The kinetic stability-thermal flexibility correlation has predictive power for the studied proteins, suggesting that the strategy and methodology discussed here might be applied to other proteins in biotechnological developments, evolutionary studies and the design of protein based therapeutics.

Introducción

1.1 Estabilidad cinética de proteínas

Las proteínas son macromoléculas de origen biológico que están formadas por aminoácidos. Es la naturaleza química de estos aminoácidos lo que le confiere sus propiedades fisicoquímicas a una proteína y por lo tanto su vasta versatilidad de formas y funciones. Los experimentos de Anfinsen [5] demostraron que toda la información que se requiere para que una proteína adopte su estado nativo y funcional se encuentra en la secuencia específica de aminoácidos que la conforman.

Las proteínas tienen diversas funciones vitales en las células, como la síntesis de diversos compuestos, transporte, captación de nutrientes, regulación, de carácter estructural, etc. Para llevar a cabo estas funciones es necesario que se encuentren en su estado nativo, esto es, que la cadena de aminoácidos se encuentre plegada en una estructura tridimensional específica (la excepción son las proteínas intrínsecamente desordenadas, IDP). El proceso de plegamiento sucede dentro de la célula en el momento mismo que termina su síntesis en el ribosoma (o cuando ésta se encuentra ya avanzada).

Las interacciones que se establecen entre los residuos en las proteínas plegadas y entre estos y el solvente, son principalmente de naturaleza electrostática e hidrofóbicas. Desde el punto de vista de la termodinámica, estas interacciones forman una estructura conformacionalmente estable (N) porque el proceso de desplegamiento no está favorecido por la energía libre. Sin embargo, la oposición de los factores entálpicos y entrópicos que intervienen en la formación

de la estructura de la proteína y de su continua interacción con el solvente, tiene como consecuencia que la estabilidad sea pequeña. Por lo tanto es común que se establezca un equilibrio entre el estado nativo y el estado desplegado (U).



Para una proteína termodinámicamente estable, se cumple que $k_N > k_U$. Sin embargo, la estabilidad termodinámica de las proteínas no es suficiente para garantizar que la proteína se mantendrá en su estado nativo y funcional indefinidamente. El equilibrio 1.1 es proclive a establecer reacciones secundarias que lo lleven irreversiblemente a un estado final (F), en un proceso dependiente del tiempo.



La susceptibilidad de la proteína de sufrir estas interacciones está determinada por su estabilidad cinética. Es decir, que la estabilidad cinética de una proteína determinará su tiempo de vida media, o sea, el tiempo en el que la proteína permanecerá plegada (N) y funcional. Si la proteína es propensa a pasar al estado irreversible (F) ($k_F > k_N$), el equilibrio se desplazará al estado desplegado (U), vaciando eventualmente el estado nativo. Se dice entonces que el desplegamiento de la proteína se encuentra bajo control cinético.

Por lo tanto, existen dos conceptos diferentes y no excluyentes para estudiar la estabilidad de las proteínas: la termodinámica y la cinética. La estabilidad termodinámica hace referencia al signo de la energía libre asociada al proceso de desplegamiento, mientras que la estabilidad cinética está asociada a la dependencia del tiempo de dicho proceso [104]. Las proteínas que son consideradas cinéticamente estables, son aquellas que cuentan con una barrera energética inusualmente alta que separa el estado nativo (N) del estado de transición (TS), a partir del cual se favorece la aparición de estados no funcionales desplegados o parcialmente desplegados (U), que son los más propensos a padecer alteraciones irreversibles, tales como la agregación y la proteólisis.

El concepto de estabilidad cinética viene a contribuir a un panorama global de la estabilidad de las proteínas *in vitro*, lo cual es de interés en muchas ramas de la biotecnología [50,75,88,119]. En este contexto cabe enfatizar que muchas de las acciones que realizamos en el laboratorio para preservar las proteínas, como por ejemplo buscar el amortiguador a un pH adecuado, mantenerlas en refrigeración, precipitarlas con sulfato de amonio, etc, tienen el propósito de encontrar las condiciones en las que su estabilidad cinética sea mayor y de esta manera tener el tiempo suficiente de realizar experimentos con ellas.

Por otro lado, el estudio de la estabilidad cinética de las proteínas también contribuye al discernimiento de como la evolución ha alcanzado el sutil balance entre función y estabilidad en escalas de tiempo relevantes para la célula, en donde el ambiente puede llegar a ser altamente hostil [91]. De hecho se piensa que el ambiente extracelular adverso puede ser el principal promotor de la aparición de proteínas con una gran estabilidad cinética [20]. La perturbación de este balance en un medio tan poblado como el medio celular, puede llevar a situaciones como la degradación temprana de la proteína y la agregación amorfa o amiloidogénica que traen como resultado condiciones patológicas tales como la fibrosis cística, la deficiencia de α 1-antitripsina, la deficiencia de SCAD (short-chain acyl-CoA deshidrogenase), la enfermedad de Alzheimer, la enfermedad de Parkinson, la enfermedad de Huntington, la anemia falciforme, la diabetes tipo II, etc. La lista de enfermedades relacionadas con la agregación de proteínas ha ido creciendo los últimos años [57]. Diversos esfuerzos se han hecho en la dirección de aumentar la estabilidad cinética de proteínas que tienden a formar fibras amiloides [18,37,89,118]. Se ha sugerido que la prevención de la agregación y la degradación temprana en la célula es una importante fuerza directora de la evolución de las proteínas [24,79].

Ha habido mucha especulación acerca de las bases moleculares de la estabilidad cinética, sin que hasta el día de hoy se conozca a fondo el mecanismo por el cual la selección natural modela grandes barreras energéticas para prevenir el desplegamiento de algunas proteínas. Entre los estudios que se han llevado a cabo sobre este tema (ver la introducción de la referencia [75]) están el de mutaciones que buscan incrementar la hidrofobicidad de la superficie [71], la incorporación de puentes disulfuro [76,102] y la introducción de sitios de unión a metales [70,92].

Se ha sugerido también que la estructura cuaternaria de homo-oligómeros surgió como un mecanismo para aumentar la estabilidad cinética de las proteínas [26, 31, 32, 81, 96].

Este trabajo pretende contribuir a contestar la pregunta sobre cuáles son las bases moleculares que hacen a una proteína más estable cinéticamente que otra. Para ello hemos tomado como modelo de estudio a las triosafosfato isomerasas de *Trypanosoma cruzi* (TcTIM) y *Trypanosoma brucei* (TbTIM). Estas dos proteínas poseen elementos estructurales de alta identidad, pese a lo cual TcTIM tiene una estabilidad cinética dos veces más grande que la de TbTIM (ver más abajo) [19]. Por tal motivo constituyen un excelente modelo de estudio para investigar cómo la selección natural ha moldeado a nivel molecular las barreras energéticas que le confieren estabilidades cinéticas tan distintas a dos proteínas tan cercanas en la evolución.

Una de las maneras de cuantificar la estabilidad cinética es determinando los valores de las constantes que rigen la cinética del equilibrio 1.2 y del proceso irreversible que lleva a la proteína a un estado desnaturalizado. Este proceso se puede simplificar si se considera que los únicos estados poblados son el nativo y el final.

$$N \xrightleftharpoons[k_f]{k} F \quad (1.3)$$

Esto es lo que se conoce como el modelo de dos estados irreversible. La constante cinética de primer orden k determinará la rapidez del proceso de desnaturalización a una determinada temperatura. Se asume que la dependencia de k con la temperatura se puede describir mediante la ecuación de Arrhenius:

$$k = \exp \left[-\frac{E_{act}}{R} \left(\frac{1}{T} - \frac{1}{T^*} \right) \right] \quad (1.4)$$

Donde T^* es una temperatura de referencia y E_{act} es la energía de activación, es decir la barrera energética que separa el estado nativo del estado de transición (TS). Este parámetro es también una medida de la estabilidad cinética de una proteína, ya que entre más grande sea esta barrera, la probabilidad de cruzarla será menor y el tiempo de vida media del estado

nativo será mayor. Comúnmente el valor de E_{act} es constante en el intervalo de temperatura en el que ocurre la transición, por tal motivo resulta un parámetro más confiable para cuantificar la estabilidad cinética de las proteínas, que el valor de k cuya dependencia con la temperatura lleva en muchas ocasiones a hacer grandes extrapolaciones a temperaturas en donde no hay garantía de que el modelo de Arrhenius permanezca lineal.

Tanto los valores de k como de E_{act} se pueden obtener del gráfico de Arrhenius, el cual se puede construir con los datos obtenidos a partir de experimentos de calorimetría diferencial de barrido (DSC por las siglas de differential scanning calorimetry) para proteínas cuyo proceso de desplegamiento se encuentre bajo control cinético. La calorimetría diferencial de barrido es una técnica que nos permite caracterizar indirectamente los cambios conformacionales en proteínas como función de la temperatura. Dicha caracterización consiste principalmente en determinar la energía involucrada en el proceso de plegamiento/desplegamiento.

En los equipos de DSC se utilizan dos celdas aisladas a través de una chaqueta adiabática (recuadro de la figura 1.1). En la celda de referencia se coloca el disolvente “puro” (normalmente un amortiguador) y en la otra, una muestra de proteína disuelta en el mismo amortiguador. La técnica consiste en medir la diferencia en capacidad calorífica entre ambas celdas. Para ello se aplica una variación de temperatura en ambas celdas simultáneamente, a una velocidad constante y se mide la potencia eléctrica necesaria para mantener entre ellas una diferencia de temperatura muy cercana a cero. Esta potencia eléctrica es directamente proporcional a la diferencia en capacidad calorífica entre los líquidos contenidos en ambas celdas.

La capacidad calorífica es una medida de la capacidad del sistema de almacenar energía. Ya que la desnaturalización térmica es un proceso que requiere energía, cuando una proteína cambia del estado nativo al estado desnaturalizado, está almacenando energía en las moléculas desplegadas. Esto es lo que se observa en la traza calorimétrica como la diferencia de capacidad calorífica dadas por la línea pre-pico y post-pico (línea base química). Esta diferencia usualmente se interpreta en términos de las contribuciones de hidratación de los residuos, que están o no expuestos al solvente en determinados estados.

Dado que en realidad ambas celdas no son idénticas, se necesita de una línea base

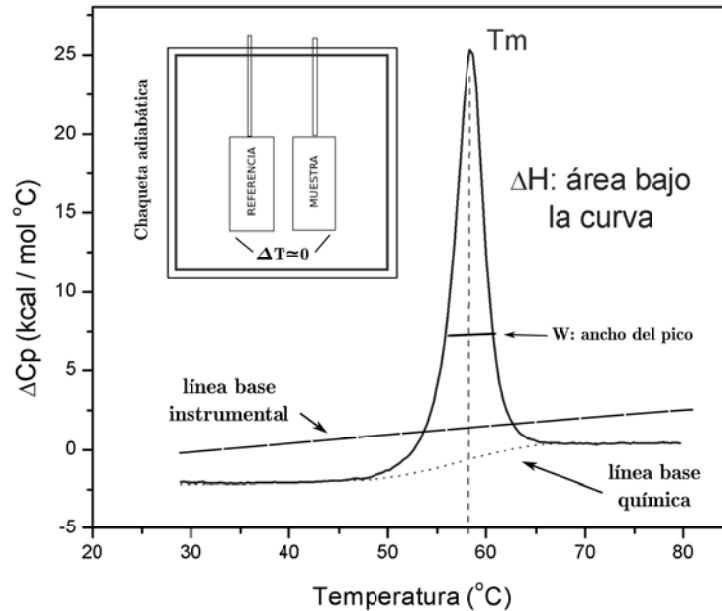


Figura 1.1: Ejemplo de una endoterma obtenida con un equipo de DSC. El área bajo la curva corresponde al cambio de entalpía del proceso y la temperatura máxima de la transición se conoce como T_m (línea discontinua). El valor de W o ancho de pico, corresponde al ancho de la traza a la mitad del área total de la transición. La diferencia de capacidad calorífica entre el estado nativo y el desplegado está dada por las líneas pre-pico y post-pico (línea base química). En el recuadro se representan las dos celdas del calorímetro.

instrumental. Ésta se obtiene cuando ambas celdas contienen únicamente el amortiguador. Posteriormente, esta línea se resta de la endoterma obtenida del desplegamiento de la proteína. El área bajo la curva corresponde al ΔH calorimétrico del proceso y la temperatura máxima de la transición se conoce como T_m . El ancho de la traza calorimétrica a la mitad del área total del pico se conoce como $T_{1/2}$ o W y es un parámetro que está ligado a la cooperatividad del proceso de desplegamiento [19]. Las proteínas cuyo proceso de desplegamiento se encuentre bajo control cinético tendrán trazas calorimétricas dependientes de la velocidad de barrido e irreversibles, esto significa que no se observará una segunda señal tras enfriar el sistema y volver a calentar.

1.2 La triosafosfato isomerasa: un barril TIM

El motivo estructural conocido como barril TIM o barril $(\beta/\alpha)_8$ fue denominado así por la estructura de la proteína triosafosfato isomerasa (TIM). Este motivo estructural se encuen-

tra presente en enzimas que catalizan cinco de los seis tipos de reacciones conocidas [41]. Aproximadamente el 10% de las proteínas cuya estructura se encuentra depositada en el Protein Data Bank (PDB) es un barril TIM [44]. Como resultado de su versatilidad, ha sido ampliamente usado en investigaciones de ingeniería de proteínas [43, 47, 100, 106].

La estructura del barril consiste en una parte central formada por ocho hebras β paralelas, cada una de ellas está unida a través de un giro $\alpha\beta$ con una hélice α . Estas hélices forman una especie de envoltorio para el centro hidrofóbico de hojas beta (ver figura 3.1 en la sección 3.1). Pese a la diversidad de reacciones que catalizan, todos los barriles TIM conocidos tienen ubicado el sitio activo en los giros $\beta\alpha$, localizados en el extremo C-terminal de las hebras β , por lo que a este lado del barril se le conoce como la cara catalítica [44].

Hay diversos estudios que argumentan que la evolución del barril TIM consistió en distintos eventos convergentes, esto es que una gran variedad de enzimas llegaron independientemente a este motivo estructural. Sin embargo, la hipótesis más aceptada es que hubo un ancestro común de donde provienen todos los barriles TIM [28, 45, 105]. Se ha especulado con la posibilidad de que este ancestro común haya evolucionado a su vez gracias a un evento de duplicación genética y fusión de medios barriles [42, 43, 108].

La triosafosato isomerasa (TIM) es una enzima dimérica de la ruta metabólica de la glucólisis. La TIM cataliza la interconversión de dihidroxiacetona fosfato a D-gliceraldehído 3-fosfato a una velocidad impuesta por la difusión, lo que ejemplifica su excelencia como catalizador [116]. Para llevar a cabo su función no requiere de algún cofactor, ni de alguna coenzima. Su forma activa es un homodímero, cada uno con un peso molecular de 27 kDa. Es interesante que a pesar de que cada monómero cuenta con un sitio catalítico independiente, esta enzima solo es funcional en su forma dimérica [2]. El sitio activo está formado por cuatro residuos: Asn11, Lys13, His95 and Glu167 (la numeración depende de la especie, aquí se muestra la de la TIM de *Trypanosoma brucei*) [117].

La construcción de quimeras de barriles TIM ha ayudado a contestar diversas preguntas en el amplio campo de investigación que tiene como objetivo comprender las propiedades de las proteínas [41, 107]. De interés especial para este trabajo es la construcción de quimeras de las TIMs de *Trypanosoma cruzi* (TcTIM) y *Trypanosoma brucei* (TbTIM) que han sido

utilizadas como herramienta para estudiar la reactividad de estas dos enzimas ante agentes sulfidrilos [35].

1.3 Las triosafosfato isomerasas de *Trypanosoma cruzi* (TcTIM) y *Trypanosoma brucei* (TbTIM)

Entre los protozoarios del género *Trypanosoma*, existen dos que son especialmente dañinos para la humanidad: *Trypanosoma cruzi* y *Trypanosoma brucei*. El primero es el causante de la enfermedad de Chagas en el continente Americano y el segundo produce la llamada enfermedad del sueño en el África subshariana. Ambas patologías constituyen un serio problema de salud pública en las zonas rurales de ambos continentes. *T. cruzi* y *T. brucei* comparten diversas características. Ambos son organismos unicelulares flagelados, cuyos vectores de transmisión son insectos hematófagos. *T. brucei* se transmite a través de la saliva de la mosca tse-tse y *T. cruzi* se trasmite a través de las heces de una variedad de chinches nativas del continente Americano [7].

Cuando se encuentra en el torrente sanguíneo de un mamífero, *T. brucei* es totalmente dependiente de la energía producida en la glucólisis. De hecho, el 90% de su contenido proteico consiste en las enzimas glucolíticas. Esta vía metabólica se encuentra en su mayoría confinada a un tipo especial de peroxisoma, llamado glicosoma. La confinación de la glucólisis en este organelo es exclusiva de los protozoarios pertenecientes al grupo Kinetoplastida, ya que esta vía es principalmente citosólica en la mayoría de los organismos. *T. cruzi* también cuenta con un glicosoma, sin embargo su dependencia por la vía glucolítica es menor [4]. Se ha especulado que sin la confinación de esta vía metabólica al glicosoma, la concentración de los carbohidratos fosforados se elevaría a niveles patológicos dentro de la célula y también que ésta sería incapaz de recuperarse de una carencia de glucosa [6].

Es interesante notar que aunque la mayoría de las enzimas glucolíticas se encuentran confinadas en este organelo, su síntesis es citosólica, por ello estos protozoarios han debido desarrollar un método de translocación post-traducciona para llevar a cabo la glucólisis [86], que aún no se conoce con detalle [85].

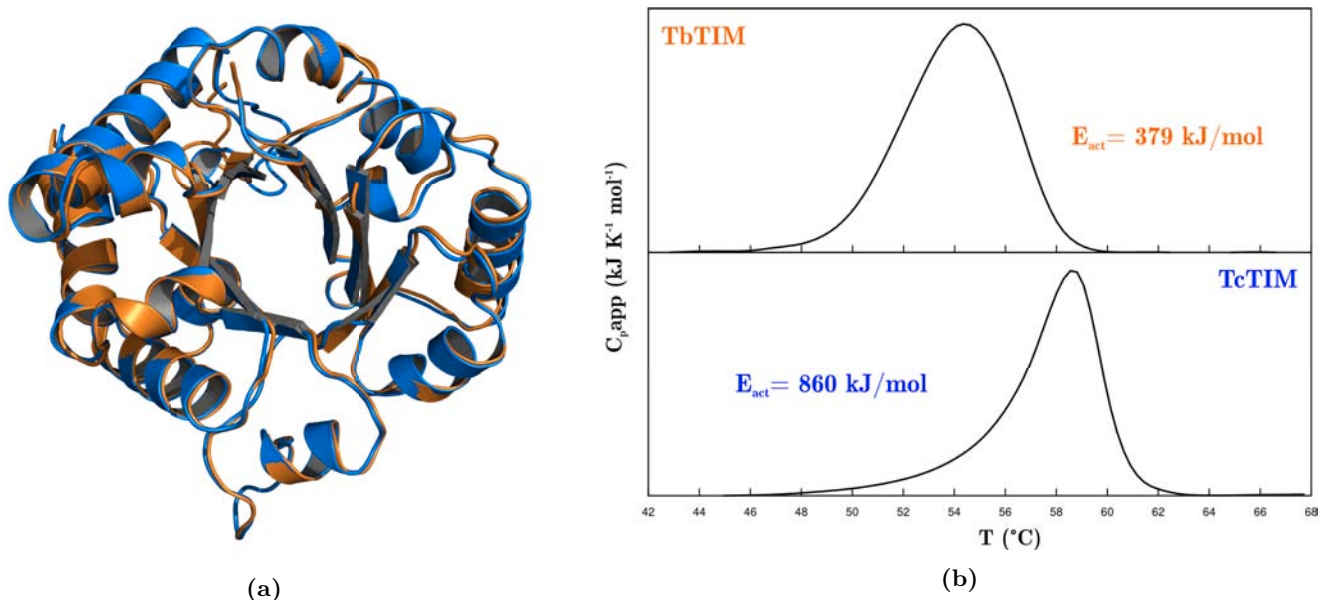


Figura 1.2: (a) Alineamiento de la estructura de los monómeros de TbTIM (naranja, PDB:5TIM) y de TcTIM (azul, PDB:1TCD). (b) Trazas calorimétricas de TbTIM y TcTIM a una velocidad de barrido de 1.5 K/min. Datos tomados de la referencia [19].

Las triosafofato isomerasas de *Trypanosoca cruzi* (TcTIM) y *Trypanosoca brucei* (TbTIM) son por lo tanto, dos enzimas muy cercanas en la evolución. Tienen un 74% de identidad en su secuencia de aminoácidos y sus estructuras se sobreponen con una desviación cuadrática media (RMSD) menor a 1 Å (figura 1.2a). A pesar de ello, presentan un comportamiento marcadamente diferente en varios aspectos [15, 35, 36, 95], incluida su estabilidad cinética. De hecho, la E_{act} de TcTIM es el doble que la de TbTIM [19]. La selección natural ha moldeado a estas dos enzimas para tener barreras energéticas muy diferentes frente al proceso de desnaturalización, sin perder eficiencia catalítica y únicamente con pocos cambios en su secuencia, lo que hace a estas dos enzimas un excelente sistema modelo para incrementar nuestro conocimiento de las bases moleculares de la estabilidad cinética.

1.4 Flexibilidad de proteínas

Las proteínas poseen una estructura que continuamente sufre fluctuaciones debido a diversos factores, tales como la interacción con el solvente, las interacciones no específicas con otras proteínas y en general, el movimiento térmico al que están sometidas. Los estados

conformacionales que visita una proteína durante estas fluctuaciones definen su flexibilidad. Esta flexibilidad es esencial para que tengan lugar diversos eventos que definen la función de las proteínas, como la actividad catalítica en el caso de las enzimas o la unión a diversos ligandos, el control alostérico, la interacción con la membrana celular, etc.

La flexibilidad de las proteínas es un parámetro que no sólo se se liga con su función, sino también con su estabilidad [114] y con su evolución [82]. De tal manera que en los últimos años ha habido un incremento en el interés de la comunidad científica por estudiar la flexibilidad de las proteínas en solución de manera cualitativa y cuantitativa [9, 49, 112].

La relación entre la flexibilidad y la estabilidad de una proteína es una interrogante que hasta la fecha no se ha resuelto satisfactoriamente. Una de las razones principales es el hecho de que los conceptos de flexibilidad y estabilidad pueden tomar diferentes significados, de tal manera que quizá lo más correcto es hablar sobre las relaciones que existen entre las flexibilidades y las estabilidades de una proteína.

Tang y Dill [111] dividen el concepto de flexibilidad de una proteína en estática y dinámica. La flexibilidad estática se refiere a la diversidad estructural de conformaciones que se encuentran en equilibrio en el estado nativo de la proteína, mientras que la flexibilidad dinámica se refiere a la rapidez con la que una proteína puede transitar de una conformación a otra (figura 1.3).

La mayoría de las investigaciones enfocadas en estudiar la relación flexibilidad/estabilidad son en realidad estudios de la relación entre la flexibilidad estática y la estabilidad térmica, medida como el valor de T_m de una proteína a través de diversas técnicas [54, 111]. La conclusión general de estos estudios es que hay una relación inversa entre ambas propiedades: entre más flexible es una proteína su estabilidad térmica es menor. Sin embargo, ha habido algunos resultados que no apoyan esta hipótesis general [30, 38], sobre todo cuando la flexibilidad que se mide tiene como variable el tiempo, es decir, cuando se estudia una propiedad que es más similar al concepto de flexibilidad dinámica (ver las revisiones [54] y [55]).

En cuanto a las investigaciones de la relación entre la estabilidad cinética con la flexibilidad, los resultados también parecen ser contradictorios. Hay dos estudios que afirman que entre

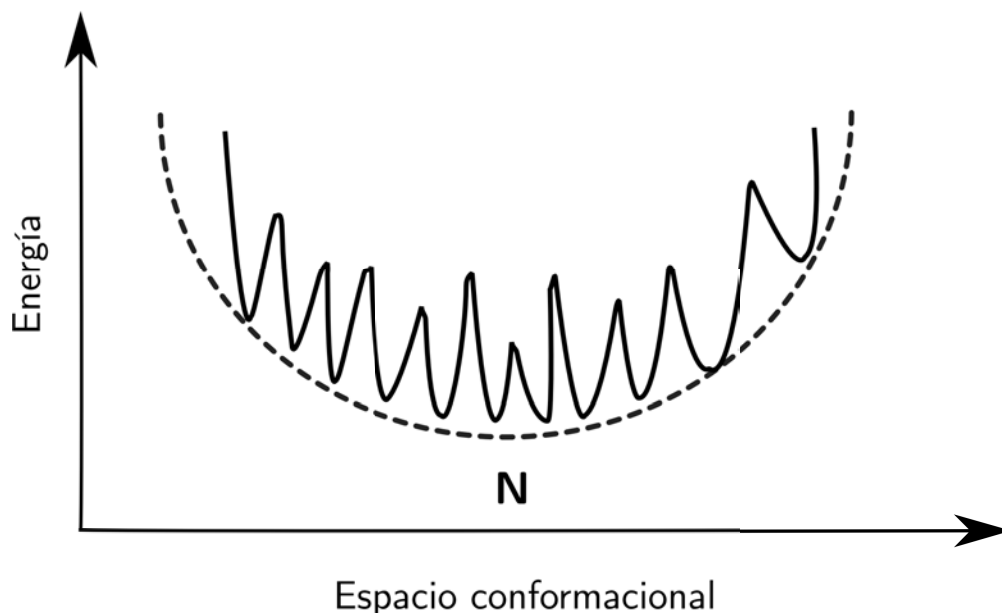


Figura 1.3: El estado nativo de una proteína sufre continuamente fluctuaciones conformacionales. La flexibilidad estática determina el espacio conformacional accesible para la proteína (micro-pozos que determinan micro-estados), mientras que la flexibilidad dinámica determina que tan rápido ocurre el tránsito de la estructura por estos micro-pozos. Imagen modificada de [109].

más rígida sea una proteína su estabilidad cinética será mayor [75, 120]. Los parámetros de flexibilidad que se utilizan en estos estudios son la susceptibilidad a una proteasa a temperatura ambiente, en el primer caso y en el segundo los factores B de la estructura determinada por rayos X. En ambos casos se puede pensar que la flexibilidad a la que hacen mención coincide más con el concepto de flexibilidad estática. Por otro lado, hay otras investigaciones que concluyen que una proteína cinéticamente estable tiene una flexibilidad alta [12, 23]. Sin embargo, en estos dos estudios no queda muy claro a qué tipo de flexibilidad se refieren. El primero cuantifica la flexibilidad como la cantidad de Newton/metro que se aplica en una serie de experimentos de desplegamiento usando microscopía de fuerza dinámica de una molécula dentro de proteoliposomas depositados en mica y el segundo utiliza dispersión cuasielástica de neutrones para caracterizar la dinámica del sistema en el orden de picosegundos.

En este trabajo utilizaremos las simulaciones de dinámica molecular (MDS) de alta temperatura [21] para cuantificar el desplazamiento de los residuos de TbTIM y TcTIM (que tienen estabilidades cinéticas muy diferentes) en condiciones desnaturalizantes. Este

análisis proporcionará una medida de la flexibilidad de las enzimas como una función de la temperatura, por lo que la hemos llamado flexibilidad térmica en analogía al concepto de flexibilidad dinámica, con la diferencia de que el parámetro que varía es la temperatura y no el tiempo.

Las MDS de la triosafosfato isomerasa ya han contribuido al estudio de diferentes cuestiones tales como la dinámica del giro 6 (o giro catalítico) [48, 61–63], el efecto de la dimerización en la dinámica de la proteína [14], el efecto de la unión de sustratos a la intercara [64], y la dinámica de esta enzima en presencia de solventes orgánicos [25].

Las MDS son una herramienta computacional que nos permite estudiar la dinámica de un sistema confinado a una caja virtual de forma y tamaño definidos. De manera muy general, esta técnica utiliza las ecuaciones de movimiento de Newton para estudiar la evolución de un sistema como función del tiempo. Antes de empezar la simulación se realiza una minimización de la conformación inicial de la proteína dentro de la caja simulación para optimizar las interacciones interatómicas y la energía potencial total del sistema. Posteriormente se solvata la proteína con las moléculas de agua y se realiza una pequeña simulación para equilibrar el sistema, manteniendo los átomos pesados restringidos a sus posiciones iniciales (figura 1.4). Finalmente se realiza la simulación de producción.

Para alcanzar el objetivo planteado en este trabajo de incrementar el conocimiento de las bases moleculares de la estabilidad cinética, se recurrirá entonces a una aproximación que consiste en el empleo de tres diferentes técnicas: ingeniería de proteínas, calorimetría diferencial de barrido y simulaciones de dinámica molecular. La construcción de quimeras de TbTIM/TcTIM [35] se retomó en este estudio para reconocer las regiones estructurales responsables de la diferencia en estabilidad cinética entre ambas proteínas. La determinación de la estabilidad cinética de estas quimeras se realizó a través de experimentos de DSC (ver sección 1.1). Por último, las MDS se utilizaron para estudiar el desplazamiento de los residuos de las enzimas silvestres a diferentes temperaturas. Estos desplazamientos serán confrontados con los resultados de la calorimetría diferencial de barrido obtenidos para las enzimas silvestres y las quimeras de TbTIM/TcTIM.

Este trabajo está organizado de la siguiente manera: en el capítulo 2 se describen los

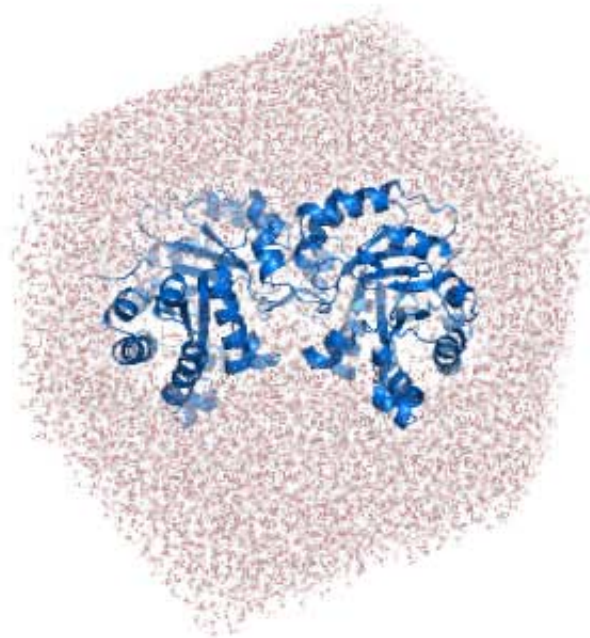


Figura 1.4: Caja de simulación con forma de dodecaedro truncado que contiene la conformación inicial de la proteína (TbTIM) y 23287 moléculas de agua del modelo SPC (ver la sección 2.4).

métodos que se utilizaron para el diseño, clonación y purificación de las quimeras, las condiciones en las que se corrieron los experimentos de DSC, los detalles del modelo de dos estados irreversible al que se ajustaron las trazas, las condiciones en las que se lanzaron las simulaciones de dinámica molecular de alta temperatura, así como el método que se desarrolló para cuantificar la flexibilidad térmica. En el capítulo 3 se presentan los resultados de la eficiencia catalítica y de DSC de las veintinueve quimeras que se utilizaron en este trabajo, el análisis general de las MDS, así como la cuantificación de la flexibilidad térmica de las enzimas silvestres y de las diferentes regiones en las que se descompone. Posteriormente se establece una correlación entre la estabilidad cinética de las proteínas estudiadas mediada a través de DSC y la flexibilidad térmica obtenida de las MDS. Por último se hace un análisis de la contribución de cada residuo a los valores de flexibilidad térmica de las enzimas silvestres.

Métodos

2.1 Diseño, clonación, expresión y purificación de las quimeras de TbTIM/TcTIM

Para el diseño de las quimeras se utilizaron las secuencias de ADN depositadas en la base de datos del National Center for Biotechnology Information (NCBI) con las etiquetas X03921 para TbTIM y U53867 para TcTIM. Los genes de las quimeras Tc(4)Tb(1,3,5-8), Tc(1,2)Tc(3-8), Tc(1)Tb(2-8) y Tc(2,3,5-8)Tb(1,4) fueron construidos utilizando diferentes reacciones en cadena de la polimerasa (PCR), utilizando la Accuzyme DNA polimerasa (Bioline, Taunton MA) y los oligonucleótidos promotor y terminator T7 externos. Los detalles sobre las enzimas y los oligonucleótidos empleados en estas síntesis pueden encontrarse en las referencias [34] y [35]. Los demás genes fueron sintetizados por la compañía GenScript con sede en Piscataway, New Jersey, Estados Unidos.

Los genes fueron insertados en el plásmido de expresión con resistencia a ampicilina pET-3a. Las enzimas de restricción usadas fueron Nde-I y BAMHI. Tras comprobar las secuencias, se transformaron células de BL21 (DE3)pLysS. Las bacterias fueron sembradas en un medio de Luria Bertani con 100 $\mu\text{g}/\text{mL}$ de ampicilina e incubadas a 310 K hasta alcanzar una absorbancia de aproximadamente 0.6 a una longitud de onda de 600 nm. Para inducir la sobreexpresión, se añadió una concentración final de 0.4 mM de isopropil- β -D-1-tiogalactopiranosido (IPTG) y se incubaron durante 12 horas a 303 K. Posteriormente, las células fueron centrifugadas y resuspendidas en 20 mL de un amortiguador que contenía 100

mM de ácido 2-(N-morfolino)etanosulfónico (MES), 1 mM de ditioneitol (DTT), 0.5 mM de ácido etilendiaminotetracético (EDTA), 0.2 mM de fluoruro de fenilmetilsulfonil (PMSF) y 300 mM de NaCl a pH 6.3. Las células fueron sonicadas y centrifugadas a $1.44 \times 10^5 \times g$ durante una hora. El sobrenadante fue diluido hasta una concentración aproximada de 20 mM de NaCl y pasado por la columna de intercambio iónico SP-Sepharosa Fast Flow. Esta columna se equilibró previamente con un amortiguador de 50mM MES a pH 6.3 y la muestra se eluyó con un gradiente lineal de NaCl de 0 a 500 mM, en el mismo amortiguador. Las fracciones fueron precipitadas con 70 % de saturación de $(\text{NH}_4)_2\text{SO}_4$ con agitación continua durante 3 horas. Posteriormente fueron centrifugadas a $2.3 \times 10^4 \times g$ durante 15 minutos y disueltas en 3 mL de una solución que contenía 100 mM de trietanolamina (TEA), 1 mM de EDTA y 2.2 M de $(\text{NH}_4)_2\text{SO}_4$ a pH 7.4. El último paso de purificación consistió en pasar esta muestra por una columna hidrofóbica Toyopearl, previamente equilibrada con la misma solución. La muestra se eluyó con un gradiente lineal de 2.2 a 0 M de sulfato de amonio. Los pasos de purificación fueron verificados utilizando electroforesis en gel de poliacrilamida con dodecilsulfato de sodio (SDS-PAGE). Las proteínas se almacenaron a 277 K en una solución de 70 % de saturación de $(\text{NH}_4)_2\text{SO}_4$.

2.2 Actividad enzimática

La actividad enzimática de las quimeras se determinó en la dirección de gliceraldehído 3-fosfato (GAP) a dihidroxiacetona fosfato (DHAP) a través de un ensayo acoplado con la enzima α -glicerolfosfato deshidrogenasa (α -GDH). La α -GDH toma el producto de la actividad enzimática de la TIM, el DHAP y cataliza la conversión redox de éste a sn-glicerol 3-fosfato, utilizando como co-enzima a el β -nicotinamida adenina dinucleótido (NADH). La reacción comienza al añadir 5 ng/mL de la enzima a la celda que contiene la mezcla: 100 mM de trietanolamina (TEA), 10 mM de ácido etilendiaminotetraacético (EDTA) a pH 7.4, 1 mM de GAP, 0.2 mM de NADH y 5 $\mu\text{g}/\text{mL}$ de α -GDH. La oxidación del NADH se siguió espectrofotométricamente a una longitud de onda de 340 nm.

Para calcular los parámetros catalíticos, la concentración de GAP que se utilizó fue entre 0.02 y 3.2 mM. Los resultados fueron ajustados al modelo de Michaelis-Menten y los valores

de K_M y V_{max} fueron calculados mediante una regresión no-lineal. Para calcular la eficiencia catalítica, el valor de K_M fue corregido de tal forma que se considera que sólo el 4% de GAP está libre en la solución [113]. Todos los ensayos se llevaron a cabo en un espectrofotómetro Cary-50 (Varian, Australia Pty, Ltd) con un arreglo multiceldas. La temperatura de la celda se mantuvo a 298 K con un baño de recirculación.

2.3 Calorimetría diferencial de barrido

Los experimentos de calorimetría de diferencial de barrido se realizaron en un calorímetro capilar VP-DSC (Malvern, Reino Unido). Las proteínas fueron preparadas mediante una diálisis contra un amortiguador de 100 mM de TEA y 10 mM de EDTA a pH 7.4. La concentración de proteína fue de 0.4 mg/mL (a menos que otra cosa se especifique en las figuras que muestran las trazas correspondientes) y fue determinada midiendo su absorbancia a una longitud de onda de 280 nm con un espectrofotómetro Cary-50.

El intervalo de velocidades de barrido que se exploró fue de 0.5 a 3 K/min. Las trazas fueron ajustadas al modelo de dos estados irreversible (ecuación 1.3) [1, 19, 97], usando un software diseñado para tal propósito. En este modelo se asume que, al proveer de energía al sistema, la proteína sufre una transición conformacional que la lleva del estado nativo a un estado final del cual no puede regresar. La conversión cinética entre estos dos estados se describe a través de una constante de primer orden (k). Se asume que k varía con la temperatura de acuerdo a la ecuación de Arrhenius (ecuación 1.4).

La capacidad térmica aparente en este modelo está dado por:

$$C_p^{app} = C_p^{pre} + (C_p^{post} - C_p^{pre})(1 - X_N) - \Delta H \left(\frac{dX_N}{dT} \right) \quad (2.1)$$

donde

$$X_N = \exp \left[-\exp \left(\frac{E_{act} \Delta T}{RT_m^2} \right) \right] \quad (2.2)$$

y

$$\frac{dX_N}{dT} = -\frac{E_{act}}{RT_m^2} \exp\left(\frac{E_{act}\Delta T}{RT_m^2}\right) \exp\left[-\exp\left(\frac{E_{act}\Delta T}{RT_m^2}\right)\right] \quad (2.3)$$

X_N es la fracción mol del estado nativo, ΔH es la entalpía de la transición de N a F, C_p^{pre} y C_p^{post} son las líneas base pre y post-transición, que se toman como funciones lineales de la temperatura, T_m es la temperatura correspondiente al máximo de la transición, $\Delta T = T - T_m$ y E_{act} es la energía de activación que separa el estado nativo del estado de transición. Todos los experimentos de DSC se ajustaron con las ecuaciones 2.1, 2.2 y 2.3.

2.4 Simulaciones de dinámica molecular

2.4.1 Cajas de simulación y parametrización

Las simulaciones de dinámica molecular de las proteínas silvestres TbTIM y TcTIM se llevaron a cabo a partir de las estructuras depositadas en el PDB como 5TIM y 1TCD, respectivamente. Se utilizó el paquete GROMACS [11, 40, 67], versión 4.5.1 y el campo de fuerza GROMOS53a6 [84]. En todos los casos se utilizaron cajas en forma de dodecaedros truncados con condiciones periódicas en las fronteras del espacio tridimensional. Para evitar la interacción de imágenes periódicas, el tamaño de la celda unitaria de simulación en la configuración inicial se eligió de tal forma que hubiera una distancia mínima de 0.9 nm entre cada átomo de la proteína y las paredes de la caja. Las cajas se solvataron con el modelo explícito de agua Simple Point Charge (SPC) [10]. Se añadieron 23287 moléculas de agua en el caso de TbTIM y 19729 en el caso de TcTIM. La energía de cada sistema se minimizó usando el método de gradiente descendiente (steepest descent) y la configuración resultante se equilibró a 298 K con una simulación de 5 ns usando el conjunto NPT. Los átomos pesados (que son todos menos el hidrógeno) fueron restringidos a su posición inicial a través de un potencial armónico con una constante de fuerza de $1000 \text{ kJ mol}^{-1}\text{nm}^{-2}$ durante el proceso de equilibración. Las simulaciones se corrieron a volumen y temperatura constante (conjunto NVT) durante 24 ns a 400, 450, 490 y 510 K y durante 7 ns a 700 K. La temperatura

se controló mediante un termostato de rescalado de velocidades (v-rescale) [83] con una constante de acoplamiento de 0.1 ps. Para las simulaciones a presión constante (únicamente los 5 ns de equilibración a 298 K) se utilizó un barostato de Berendsen [87] con una constante de acoplamiento de 0.5 ps y considerando una compresibilidad isotérmica de 4.5×10^{-5} . Las interacciones electrostáticas de largo alcance se calcularon usando el método de particle mesh Ewald (PME) [22, 27] con un “cutoff” de 1.2 nm, una malla (spaced grid) de 0.15 y una interpolación B-spline de cuarto orden. Las interacciones de van der Waals se truncaron a 1.2 nm. Una corrección de dispersión de largo alcance se aplicó a la energía y a la presión. La lista de pares (neighbor list) fue actualizada cada 5 pasos. Se utilizó el método de “leapfrog” [46] para integrar la ecuación de movimiento cada 2 fs. Para restringir la longitud de los enlaces de la proteína se utilizó el algoritmo LINCS [39] y el algoritmo de SETTLE [78] para restringir la longitud de los enlaces y los ángulos de las moléculas de agua. Las coordenadas y las energías fueron guardadas cada picosegundo para su análisis.

2.4.2 Análisis estándar de las trayectorias

Se espera que las estructuras de las proteínas se desplieguen a diferente velocidad en el rango de temperatura entre 400 y 700 K. Las herramientas del paquete de GROMACS se utilizaron para caracterizar la evolución del área accesible al solvente (ASA), la estructura secundaria, los puentes de hidrógeno, el radio de giro y la desviación cuadrática media (RMSD) de las estructuras de TbTIM y TcTIM.

2.4.3 Desplazamientos de residuos y su dependencia con temperatura (TGD)

Los efectos que el desplegamiento inducido por temperatura tuvo en cada residuo, en cada subunidad y en las proteínas completas, así como diferencias específicas entre TbTIM y TcTIM, fueron evaluados por un procedimiento que se esquematiza en la figura 2.2 y que consiste en los siguientes pasos:

- (i) Para cada trayectoria se tomó la estructura de la proteína cada 4 ps y se construyó una matriz con los valores de RMSD de los carbonos α (C_α) obtenidos de la alineación

de cada conformación. Este paso se llevó a cabo con dos diferentes alineamientos. El primero fue utilizando todo el dímero y el segundo con cada subunidad por separado.

- (ii) Se realizó un análisis de escalamiento multidimensional (MDS) a partir de la diagonalización de cada matriz de C_α -RMSD. En todos los casos, sólo las tres dimensiones con los eigenvalores más grandes fueron considerados. La proyección de la trayectoria en este espacio reducido de RMSD nos arroja una representación donde la semejanza entre las estructuras es equivalente a su distancia espacial.
- (iii) Las estructuras en el espacio reducido de RMSD fueron agrupadas en clústers usando un cálculo de dispersión progresiva de movilidad (PMD) sobre un grupo de 200 estructuras secuenciales, lo que corresponde a un intervalo de simulación de 800 ps. El valor de PMD se determinó como la distancia promedio entre cada estructura en el grupo y su posición media. El grupo de 200 estructuras fue redefinido progresivamente como $[i, \dots, i + 200]$ donde i va desde 1 a $n - 200$, siendo n la última estructura de la trayectoria. Hay que hacer notar que el alineamiento hecho para calcular las matrices C_α -RMSD puede ser significativamente distinto si se lleva a cabo con toda la proteína o con las subunidades por separado. Por tal motivo, en las gráficas de PMD se escogieron los mínimos simultáneos del dímero y las subunidades, como las estructuras representativas de cada clúster a lo largo de las trayectorias. Los mínimos en las gráficas de PMD representan estructuras que permanecen relativamente estables, mientras que los máximos representan segmentos de la trayectoria donde la proteína sufre cambios conformacionales rápidos.
- (iv) Para cada una de estas estructuras representativas la distancia interna promedio entre cada C_α y todos los demás C_α de la estructura fue determinada como:

$$d_{j,c} = \frac{1}{N} \sum_{i=1}^N \vec{r}_j^c - \vec{r}_i^c \quad (2.4)$$

donde j y c se refiere al número de residuo y al número de clúster, respectivamente. N es el número total de residuos que se toman en cuenta en el análisis, que puede ser el total del dímero para el análisis dimérico o el de cada subunidad para el análisis monomérico. En la figura 2.1 se esquematiza la diferencia entre ambos tipos de análisis. Es importante

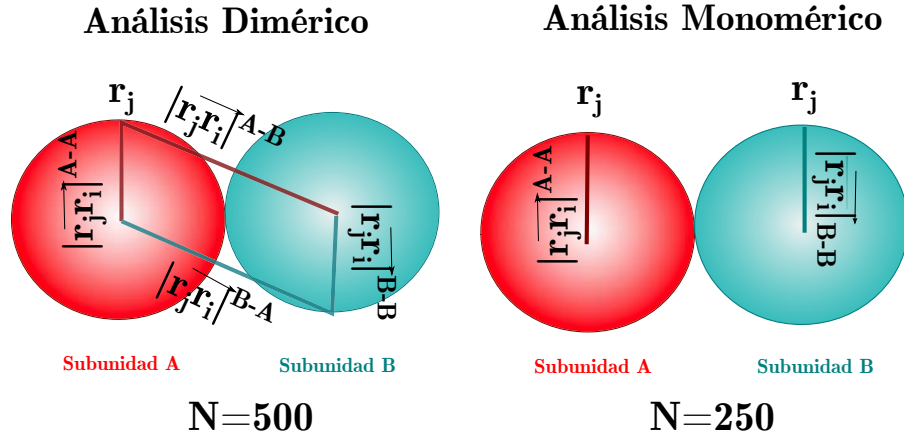


Figura 2.1: Análisis dimérico y análisis monomérico de los desplazamientos d_j . El análisis dimérico mide las distancias promedio entre cada C_α y todos los demás C_α del dímero, mientras que en el análisis monomérico se miden las distancias entre los C_α de la misma subunidad

resaltar que en los dos casos, no existe dependencia de ningún alineamiento estructural, dado que las distancias son internas para cada estructura.

- (v) Se determinó la forma en la que cada residuo se movía en cada estructura representativa de los clústers obteniendo la diferencia entre cada distancia interna, $d_{j,c}$ y el valor correspondiente en la estructura cristalina, $d_{j,0}$:

$$\Delta d_{j,c} = d_{j,c} - d_{j,0} \quad (2.5)$$

Los valores positivos de $\Delta d_{j,c}$ indican que, en promedio, el residuo j se aleja de toda la estructura en dicha conformación, mientras que los valores negativos indican que el residuo se acerca en promedio al resto de la estructura.

- (vi) El desplazamiento total de la proteína $D_{X,T}$ a cada temperatura T , donde X corresponde a las enzimas silvestres TbTIM o TcTIM, se obtuvo como el promedio de la suma de los valores absolutos de $d_{j,c}$ para las subunidades A y B, promediadas para todos los clústers (N_c), excepto los tres primeros (identificados entre los 4 y 7 ns iniciales de las simulaciones), porque en ese intervalo la evolución conformacional de la proteína es muy

rápida:

$$D_{X,T} = \frac{1}{N_c - 3} \sum_{c=4}^{N_c} \frac{1}{2} \sum_{j=1}^N (\Delta d_{j,c} A + \Delta d_{j,c} B) \quad (2.6)$$

(vii) La dependencia con la temperatura de los desplazamientos ($TGD = dD_{X,T}/dT$) para cada proteína fue calculado a través de un ajuste lineal:

$$D_{X,T} = a + TGD \cdot T \quad (2.7)$$

donde la variable independiente fue la temperatura de simulación (400, 450, 490 y 510 K). El desplazamiento de cada región, r , a cada temperatura ($D_{r,T}$) para ambas enzimas silvestres se calculó asumiendo que las regiones son independientes, usando la ecuación 2.6 pero restringiendo la suma a los residuos presentes en cada región (ver figura 3.1 en la sección 3.1). Por lo tanto, el TGD de cada región está dado también por la ecuación 2.7, pero usando $D_{r,T}$ en lugar de $D_{X,T}$. Es importante resaltar que los valores de $D_{r,T}$ y $D_{X,T}$ usados en este cálculo son de las contribuciones independientes de los residuos que constituyen cada proteína o región. De esta manera los valores de $D_{X,T}$ se pueden obtener de la suma de los valores de $D_{r,T}$ de cada una de las ocho regiones que constituyen TbTIM o TcTIM:

$$TGD_X = \sum_{r=1}^8 TGD_r \quad (2.8)$$

Los desplazamientos obtenidos de la ecuaciones 2.6 y los valores de TGD obtenidos con la ecuación 2.7 se pueden normalizar usando el número total de residuos en la proteína, con la finalidad de permitir comparaciones directas con otras proteínas.

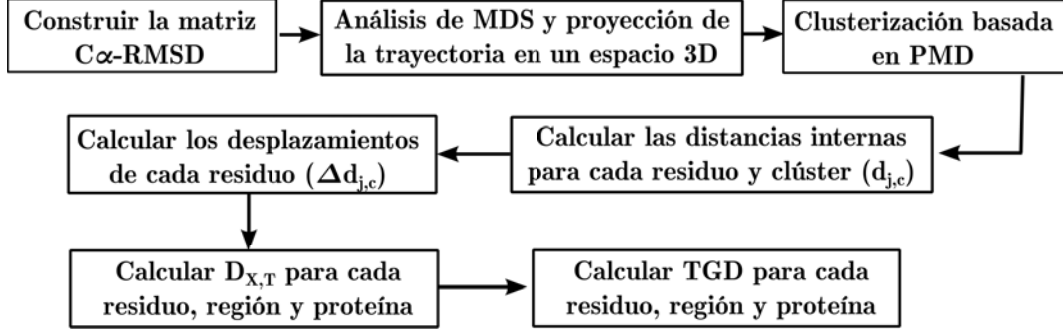


Figura 2.2: Procedimiento que se siguió para evaluar los efectos que el desplegamiento inducido por temperatura tuvo en cada residuo, en cada subunidad y en las proteínas completas

2.4.4 Estimación de TGD para las quimeras y contribución de cada residuo al TGD de las silvestres

Para cada una de las proteínas quiméricas (Q) el desplazamiento $D_{Q,T}$ fue estimado asumiendo aditividad de las regiones:

$$D_{Q,T} = \sum_{rTc} D_{rTc} + \sum_{rTb} D_{rTb} \quad (2.9)$$

en donde rTc y rTb indican respectivamente las regiones de TcTIM y TbTIM que forman la quimera. Una vez estimado el desplazamiento, se calcula el TGD de la quimera (TGD_Q) usando la ecuación 2.7.

Para cada residuo, el desplazamiento ($D_{res,T}$) está dado por:

$$D_{res,T} = \frac{1}{N_c - 3} \sum_{c=4}^{N_c} \frac{1}{2} (\Delta d_{j,c}^A + \Delta d_{j,c}^B) \quad (2.10)$$

y el valor de TGD_{res} se estima a su vez con la ecuación 2.7.

Es importante hacer notar que la selección de los clústers esta basada en los alineamientos entre los C_α de las diferentes conformaciones a lo largo de la trayectoria y que cada valor de RMSD es un solo número que representa la diferencia entre dos estructuras alineadas que están formadas por miles de átomos. En contraste, la determinación del desplazamiento de los residuos y de los valores de TGD no depende de ningún alineamiento, sino únicamente de la estructura que se está analizando. En el análisis de desplazamiento de los residuos descrito aquí se hace una diferencia entre un análisis utilizando el dímero completo (análisis dimérico)

y un análisis de las subunidades por separado (análisis monomérico) (ver figura 2.1). Esto se debe a que la distancia promedio entre un residuo y la subunidad a la que pertenece no será la misma a la distancia promedio entre ese mismo residuo y la proteína completa, aún cuando ambas distancias sean internas y no dependan de ningún alineamiento. En el análisis dimérico se espera obtener información tanto del movimiento relativo de las subunidades (acercamiento, separación o rotación), así como del desplegamiento de la proteína. Esta información sobre la estructura cuaternaria de la proteína, si bien es útil en otro contexto, puede ocultar información sobre cambios en la estructura terciaria y secundaria. Debido a la aproximación que hacemos para calcular los desplazamientos y los valores de *TGD* de las quimeras, en los resultados que se presentan a continuación se ha utilizado únicamente el análisis monomérico, dado que está libre de la posible “contaminación” debida a los cambios de la estructura cuaternaria de las enzimas silvestres.

Resultados y discusión

Este capítulo está dividido en cuatro secciones. En la primera se presentarán los resultados obtenidos del diseño, clonación y purificación de quimeras TbTIM/TcTIM y los resultados de su estabilidad cinética obtenidos con calorimetría diferencial de barrido (DSC). La segunda sección compila y discute los resultados de las simulaciones de dinámica molecular (MDS). En la tercera sección se establece una correlación entre los resultados de DSC y los de las MDS. En la cuarta y última sección se discuten las diferencias entre ambas proteínas a nivel residuo.

3.1 Caracterización de las quimeras de TbTIM/TcTIM

Las características de la estructura del barril $(\beta/\alpha)_8$ se utilizaron para dividir la estructura de la triosafosfato isomerasas en ocho regiones diferenciadas. Cada región consta de una lámina beta, de una hélice alfa y de un giro que las une (Figura 3.1). Esta división se ha utilizado con anterioridad para diseñar y construir quimeras de TbTIM/TcTIM, en las cuales una o varias regiones de una proteína son sustituidas por la o las regiones correspondientes de la otra proteína [34, 35]. La nomenclatura que empleamos aquí para denotar a cada quimera es la siguiente: Tc(R1)Tb(R2) donde Tc y Tb hacen referencia a TcTIM y TbTIM respectivamente y R1 y R2 indican las regiones de cada proteína que forman la quimera. De esta manera, la quimera Tc(2,3,5)Tb(1,4,6-8) es una quimera que tiene las regiones 2, 3 y 5 de TcTIM y las regiones 1,4,6-8 de Tb.

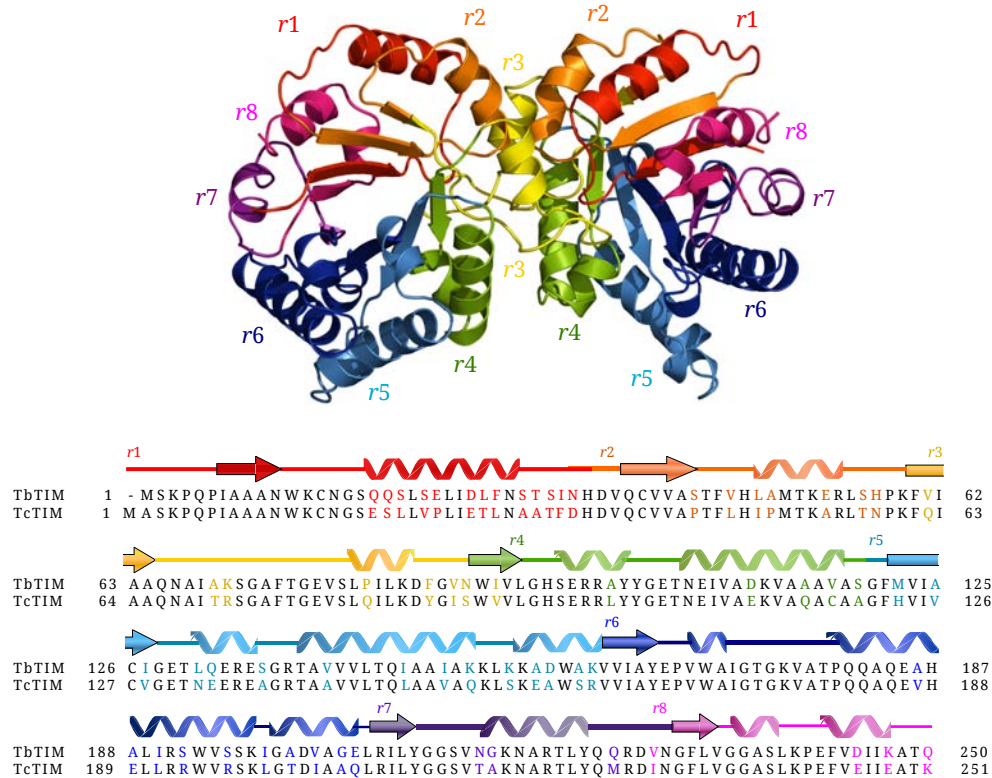


Figura 3.1: El barril $(\beta/\alpha)_8$ fue dividido en ocho regiones diferentes ($r1$ a $r8$). Cada región comprende una lámina β , una hélice α y un giro. En esta figura se indica la posición de las ocho regiones en el dímero de TbTIM y en las secuencias alineadas de ambas enzimas silvestres. Los residuos idénticos entre ambas se muestran en negro mientras que los residuos en colores indican las diferencias. El tipo de diferencias en cada región se encuentra en la tabla 3.1

Región	aminoácidos	sustituciones	conservativas	semiconservativas	s/homología
1	1-35	13	8	1	4
2	36-60	7	4	2	1
3	61-91	8	3	–	2
4	92-119	5	4	1	3
5	120-161	15	8	3	4
6	162-206	10	5	2	3
7	207-227	3	–	2	1
8	228-250	4	4	–	–
Total		65	36	11	18

Tabla 3.1: Diferencias conservativas, semiconservativas y sin homología entre las secuencias de TcTIM y TbTIM. Tabla modificada de [34]

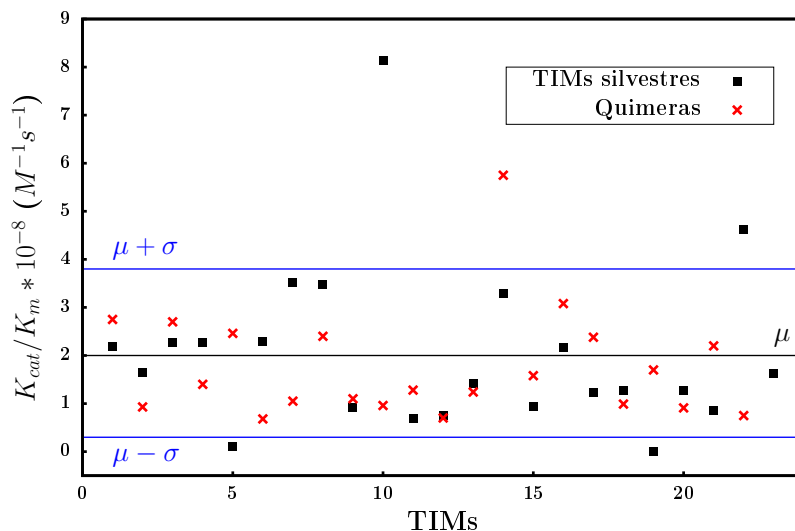


Figura 3.2: Comparación de las eficiencias catalíticas de las TIMs silvestres contenidas en la tabla 3.3 (cuadros negros) y las eficiencias catalíticas de las quimeras obtenidas en este trabajo y en la referencia [35] (cruces rojas). La línea negra es el valor promedio de las eficiencias catalíticas (μ) y las líneas azules corresponden a $\mu \pm \sigma$, en donde σ es la desviación estándar

3.1.1 Determinación de la actividad catalítica

Todas las quimeras presentaron actividad catalítica en la dirección de gliceraldehído 3-fosfato (GAP) a dihidroxiacetona fosfato (DHAP). La mayoría de los valores de k_{cat}/K_M (tabla 3.2) se encuentran dentro del intervalo de $\mu \pm \sigma$ (μ = promedio y σ = desviación estándar) que fue estimado usando los datos de veintitrés triosafosfato isomerasas silvestres obtenidos en la literatura (tabla 3.3 y Figura 3.2). Pese a que no se ha determinado la estructura de la mayoría de las quimeras de la tabla 3.2 (la excepción es la quimera Tc(2,3,5-8)Tb(1,4) con un PDB-ID 3Q37 [35]), el hecho de que todas sean catalíticamente competentes es un indicio de que las quimeras de TbTIM/TcTIM son proteínas correctamente plegadas.

3.1.2 Desnaturalización térmica caracterizada por DSC

La calorimetría diferencial de barrido fue utilizada para caracterizar el proceso de desnaturalización con temperatura de las quimeras de TbTIM/TcTIM. Las enzimas silvestres TbTIM y TcTIM fueron caracterizadas de nueva cuenta, aunque sus parámetros de DSC ya se habían reportado previamente [19]. Las figuras 3.3 a 3.10 muestran los perfiles de

Tabla 3.2: Parámetros derivados de los experimentos DSC y actividades catalíticas de las enzimas silvestres y las quimeras ^a

Proteína	E_{act} (kJ/mol)	W (K)	T_m (K) ^b	$k_{cat}/K_M \times 10^{-8}$ ($M^{-1}s^{-1}$) ^c
Enzimas silvestres				
TbTIM	434 ± 26	5.2	328.2	1.24
TcTIM	822 ± 30	2.8	331.8	1.28
Una región de TcTIM en TbTIM				
Tc(1)Tb(2-8)	361 ± 4 ^d	5.7	323.9	2.75 ^g
Tc(2)Tb(1,3-8)	510 ± 21	4.8	342.6	0.93
Tc(3)Tb(1,2,4-8)	419 ± 22	5.2	325.7	ND ^h
Tc(4)Tb(1-3,5-8)	609 ± 35	3.6	324.4	2.7 ^g
Tc(5)Tb(1-4,6-8)	600 ± 52	3.6	328.6	0.75
Tc(6)Tb(1-5,7-8)	470 ± 47 ^e	6.4	325.6	1.4
Una región de TbTIM en TcTIM				
Tc(2-8)Tb(1)	838 ± 51	2.4	337.5	ND
Tc(1-3,5-8)Tb(4)	789 ± 31	3.0	333.8	2.46 ^g
Tc(1-4,6-8)Tb(5)	639 ± 37	3.2	332.9	0.68
Dos regiones de TcTIM en TbTIM				
Tc(2,4)Tb(1,3,5-8)	493 ± 15	4.8	337.6	1.05
Tc(1-2)Tb(3-8)	493 ± 21	4.5	333.0	2.40 ^g
Tc(4,5)Tb(1-3,6-8)	626 ± 22	3.6	325.7	1.1
Tc(2,5)Tb(1-3,6-8)	459 ± 15	4.8	341.2	0.96
Tc(5,6)Tb(1-4,7-8)	526 ± 23	4.0	324.9	1.28
Tc(1,7)Tb(2-6,8)	373 ± 5	6.0	324.7	0.7
Tc(1,4)Tb(2,3,5-8)	377 ± 0 ^f	4.8	320.5	ND
Dos regiones de TbTIM en TcTIM				
Tc(1,3,5-8)Tb(2,4)	620 ± 62 ^e	6.8	320.3	ND
Tc(2-6,8)Tb(1,7)	822 ± 53	2.8	336.5	1.24
Tc(2,3,5-8)Tb(1,4)	472 ± 13	4.9	340.3	5.75 ^g
Tc(1-3,6-8)Tb(4,5)	456 ± 7	5.2	341.7	1.58
Tc(1-6)Tb(7,8)	846 ± 29	2.6	330.4	3.08 ^g
Tres regiones de TcTIM en TbTIM				
Tc(1-3)Tb(4-8)	526 ± 23	4.7	336.2	2.38 ^g
Tc(2,4,5)Tb(1,3,6-8)	659 ± 24	3.6	336.9	0.99
Tc(2,3,5)Tb(1,4,6-8)	433 ± 1	5.2	342.2	ND
Tc(6-8)Tb(1-5)	558 ± 22 ^f	4.0	327.5	ND
Tres regiones de TbTIM en TcTIM				
Tc(1-5)Tb(6-8)	953 ± 27	2.2	332.8	1.70 ^g
Tc(1,3,6-8)Tb(2,4,5)	426 ± 17	5.2	317.0	0.91
Cuatro regiones de TcTIM en TbTIM				
Tc(1-4)Tb(5-8)	639 ± 16	3.2	332.4	2.20 ^g
Tc(5-8)Tb(1-4)	706 ± 38 ^f	3.2	328.0	ND

^a Velocidades de barrido de 0.5 a 3 K/min, a menos que se indique otra cosa; ^b Los valores corresponden a 1.5 K/min; ^c Valor de K_M considerando 4% de GAP libre en solución [113];

^d De 0.5 a 2.5 K/min; ^e Único experimento a 1.5 K/min; ^f De 0.5 a 1.5 K/min;

^g Valor tomado de [35]; ^h ND: No determinado

Organismo	Eficiencia catalítica x 10 ⁻⁸ (M ⁻¹ s ⁻¹)	Referencia
<i>Escherichia coli</i>	2.18	[3]
<i>Entamoeba histolytica</i>	1.64	[66]
<i>Gallus gallus</i>	2.27	[90, 93]
<i>Giardia lamblia</i>	2.28	[69, 94]
<i>Helicobacter pylori</i>	0.11	[17]
<i>Homo sapiens</i>	2.30	[72–74]
<i>Leishmania donovani</i>	3.51	[60]
<i>Leishmania mexicana</i>	3.47	[58]
<i>Moritella marina</i>	0.92	[3]
<i>Mycobacterium tuberculosis</i>	8.13	[77]
<i>Oryctolagus cuniculus</i>	0.68	[56, 59]
<i>Plasmodium falciparum</i>	0.75	[53]
<i>Rhipicephalus microplus</i>	1.41	[80]
<i>Saccharomyces cerevisiae</i>	3.28	[59, 65]
<i>Sus scrofa</i>	0.93	[51]
<i>Schistosoma japonicum</i>	2.17	[110]
<i>Trypanosoma brucei</i>	1.24	Este trabajo
<i>Trypanosoma cruzi</i>	1.28	Este trabajo
<i>Thermotoga maritima</i>	0.0018	[8]
<i>Tenebrio molitor</i>	1.28	[56]
<i>Taenia solium</i>	0.85	[51, 52]
<i>Thermoproteus tenax</i>	4.62	[115]
<i>Trichomonas vaginalis</i>	1.63	[29]

Tabla 3.3: Eficiencia catalítica de veintitrés triosafosfato isomerasas obtenidas de la literatura. Tabla tomada de [98]

DSC a las velocidades de barrido indicadas y el gráfico de Arrhenius de las veintinueve quimeras caracterizadas y de las enzimas silvestres. Tras enfriar la muestra en el calorímetro y volver a calentarla, no se observó ninguna señal. Por lo tanto, para todas las proteínas estudiadas la desnaturalización es un proceso irreversible. También se observa que las trazas son dependientes de la velocidad de barrido, lo que indica que el proceso está cinéticamente controlado.

En las figuras 3.3 a 3.10, el ajuste al modelo de dos estados irreversible (ecuación 1.3) se muestra como una línea continua. El modelo propuesto se ajusta razonablemente bien a los datos experimentales, lo que indica que durante el proceso de desnaturalización los únicos estados significativamente poblados son el nativo (N) y el final (F). En el modelo no se contempla la disociación de las subunidades, lo que implica que la transición cinéticamente relevante ocurre desde el dímero al estado final, como es el caso de las enzimas silvestres [19],

de las mutantes E104D de varias TIMs [1] y de algunas TIMs bacterianas [99].

Los valores de la energía de activación (E_{act}), que separa el estado nativo del estado de transición, se obtuvieron al ajustar las trazas obtenidas a diferentes velocidades de barrido a la ecuación de Arrhenius (ecuación 1.4). La temperatura correspondiente al máximo de cada transición (T_m) y los valores del ancho de pico a la mitad de la altura máxima (W) se obtuvieron directamente de las trazas, por lo tanto son independientes de cualquier modelo. Todos estos parámetros se reportan en la tabla 3.2.

Los valores de E_{act} en la tabla 3.2 proporcionan información sobre la importancia relativa de ciertas regiones en la estabilidad cinética de las enzimas silvestres. La quimera Tc(5)Tb(1-4,6-8) tiene una $E_{act} = 600 \pm 52$ kJ/mol, que es 40% más grande que el valor para la TbTIM (434 ± 26 kJ/mol), lo que indica que la región 5 de TcTIM puede estar jugando un papel importante en la alta estabilidad cinética de la TcTIM (822 ± 30 kJ/mol). La E_{act} de la quimera Tc(1-3,6-8)Tb(4,5) es muy cercana (456 ± 7 kJ/mol) al de la enzima silvestre TbTIM, lo que apunta a que las regiones 4 y 5 pueden ser cruciales para la baja estabilidad cinética de esta enzima.

La intercara del barril TIM esta compuesta por residuos que se encuentran en las regiones 1, 2, 3 y 4 (tabla 3.4). Las quimeras Tc(1-4)Tb(5-8) ($E_{act}=639 \pm 16$ kJ/mol) y Tc(5-8)Tb(1-4) ($E_{act}=706 \pm 38$ kJ/mol) tienen la intercara completa de TcTIM y TbTIM, respectivamente. Las energías de activación de estas proteínas son muy cercanas y se encuentran a “medio camino” entre los valores de las silvestres TbTIM y TcTIM y además, como se verá en la sección 3.3.2, tienen un valor de flexibilidad térmica muy similar. Este resultado apoya la hipótesis de que el barril TIM evolucionó a partir de dos medios barriles (cada uno formado por cuatro “regiones”) que se fusionaron para dar nueva funcionalidad a esta versátil enzima [42, 43, 108].

Sin embargo, los resultados en la tabla 3.2 nos muestran también que independientemente del número de regiones que se intercambien entre cada proteína, o de su identidad, no se observa un cambio sistemático en los valores de E_{act} de las quimeras. De hecho, hay algunas quimeras con una estabilidad cinética más grande o más pequeña que las enzimas silvestres. De tal manera que la variación no-sistemática de E_{act} no parece ser el resultado del hecho de

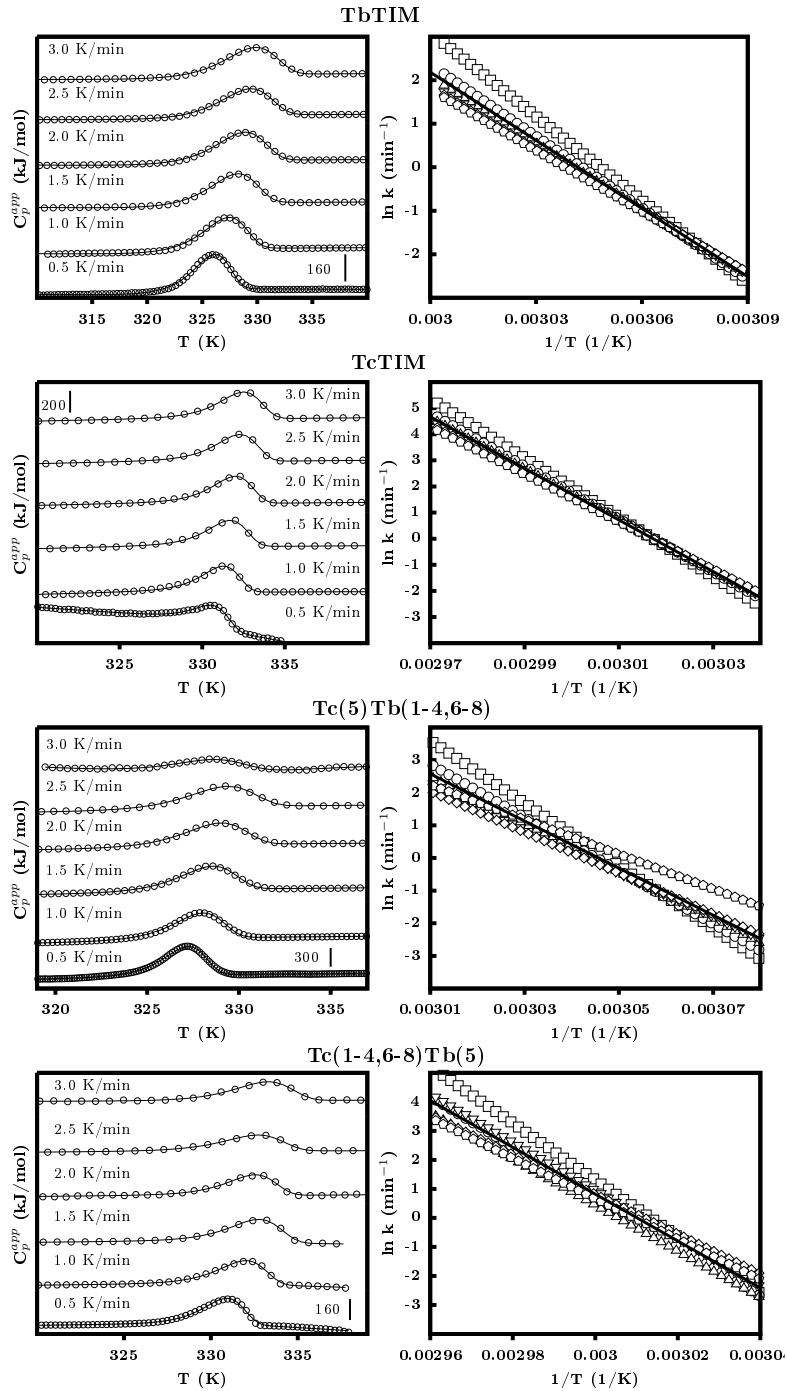


Figura 3.3: Se muestran los termogramas de DSC de las enzimas silvestres y de dos quimeras a las velocidades de barrido indicadas (izquierda). La concentración de proteína fue de 0.4 mg/mL, excepto para los experimentos a 0.5, 1 y 1.5 K/min de Tc(1-4,6-8)Tb(5) donde se usó 1 mg/mL. Los símbolos representan la capacidad calorífica aparente y las líneas continuas representan el ajuste al modelo de dos estados irreversible. Por claridad, los perfiles fueron desplazados en el eje de las ordenadas. Se muestran los gráficos de Arrhenius (derecha) para las mismas proteínas. Los símbolos corresponden a las diferentes velocidades de barrido empleadas: 0.5, 1.0, 1.5, 2.0, 2.5 y 3.0 K/min. Las líneas continuas representan el mejor ajuste lineal de todos los puntos.

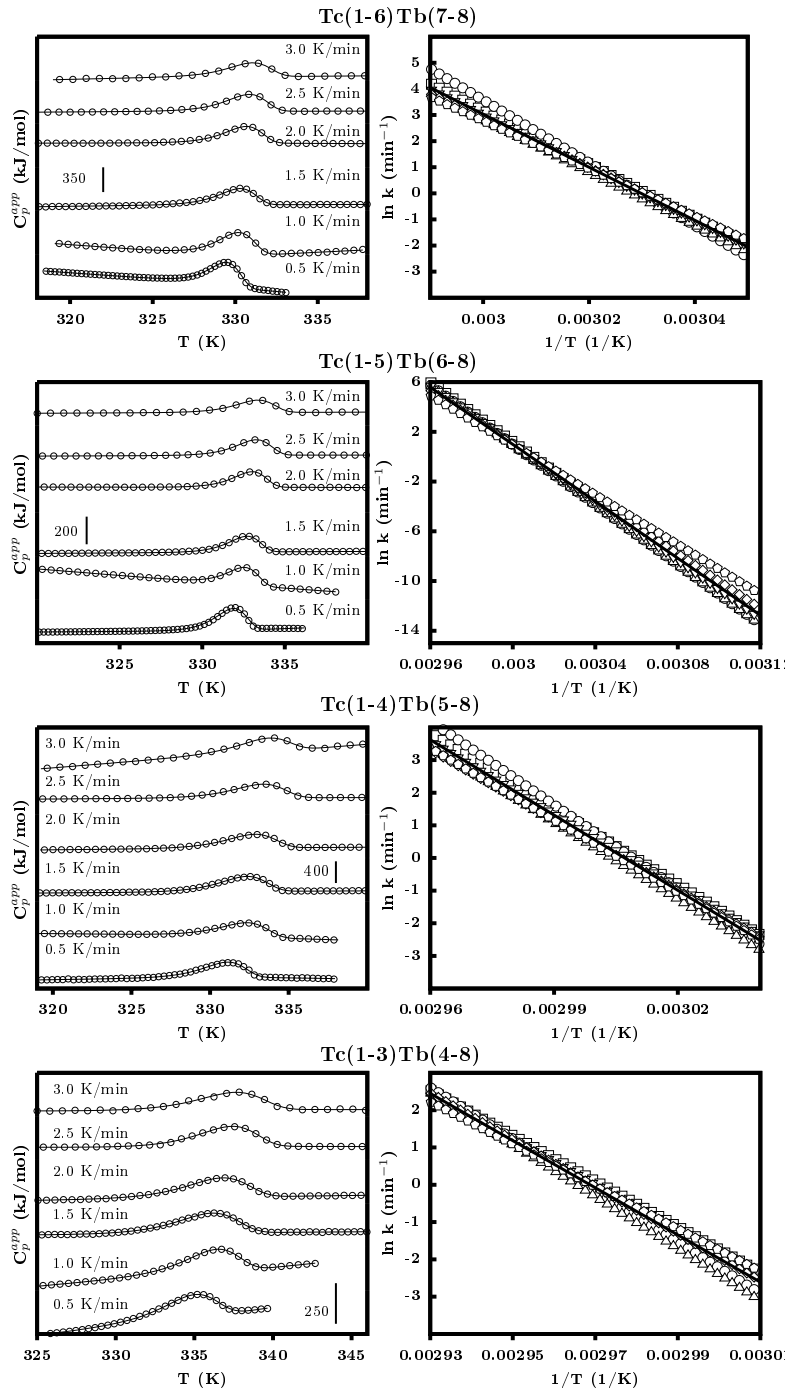


Figura 3.4: Se muestran los termogramas de DSC de cuatro quimeras a las velocidades de barrido indicadas (izquierda). La concentración de proteína fue de 0.4 mg/mL, excepto para los experimentos a 0.5 y 1 K/min en todos los casos, donde se usó 1 mg/mL. Los símbolos representan la capacidad calorífica aparente y las líneas continuas representan el ajuste al modelo de dos estados irreversible. Por claridad, los perfiles fueron desplazados en el eje de las ordenadas. Se muestran los gráficos de Arrhenius (derecha) para las mismas proteínas. Los símbolos corresponden a las diferentes velocidades de barrido empleadas: 0.5, 1.0, 1.5, 2.0, 2.5 y 3.0 K/min. Las líneas continuas representan el mejor ajuste lineal de todos los puntos.

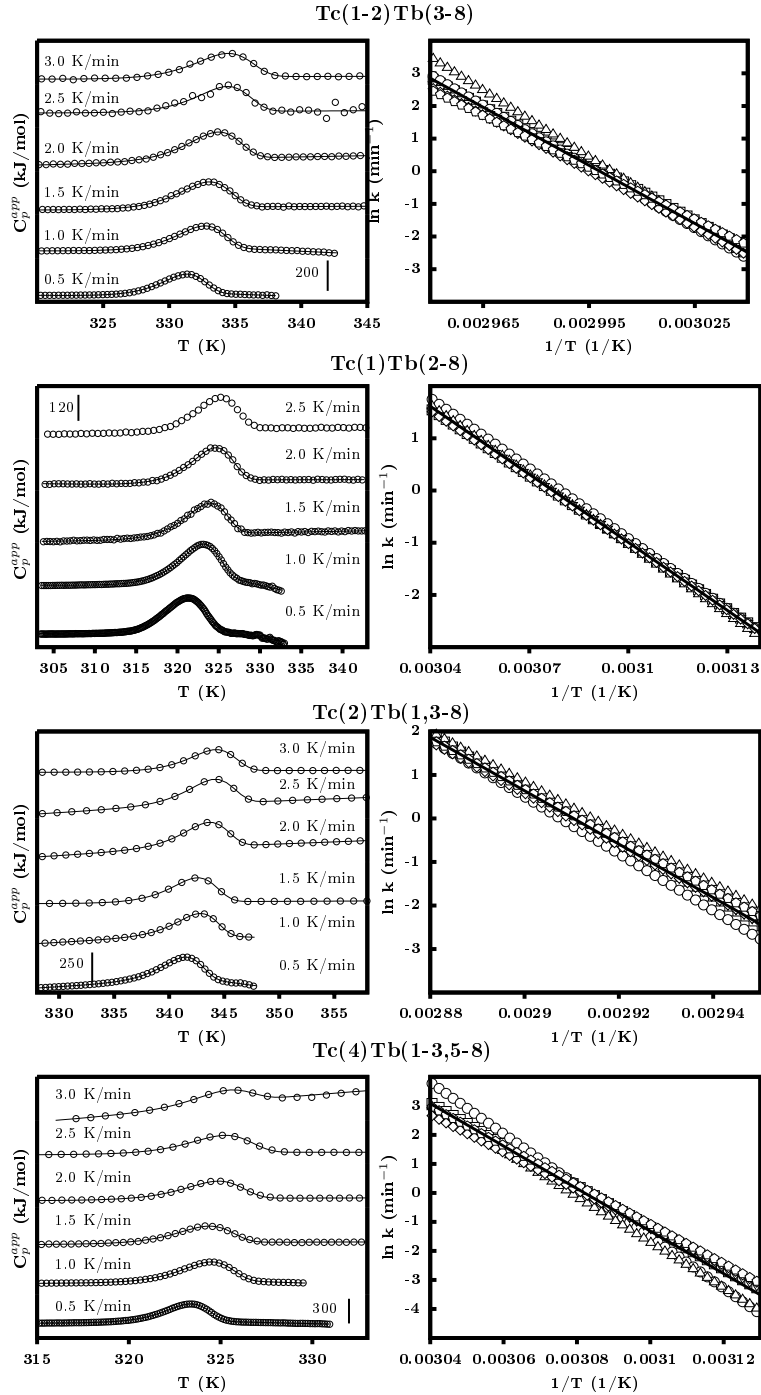


Figura 3.5: Se muestran los termogramas de DSC de cuatro quimeras a las velocidades de barrido indicadas (izquierda). La concentración de proteína fue de 0.4 mg/mL, excepto para los experimentos a 0.5 y 1 K/min de Tc(1-2)Tb(3-8), Tc(1)Tb(2-8) y Tc(4)Tb(1-3,5-8), donde se usó 1 mg/mL. Los símbolos representan la capacidad calorífica aparente y las líneas continuas representan el ajuste al modelo de dos estados irreversible. Por claridad, los perfiles fueron desplazados en el eje de las ordenadas. Se muestran los gráficos de Arrhenius (derecha) para las mismas proteínas. Los símbolos corresponden a las diferentes velocidades de barrido empleadas: 0.5, 1.0, 1.5, 2.0, 2.5 y 3.0 K/min. Las líneas continuas representan el mejor ajuste lineal de todos los puntos.

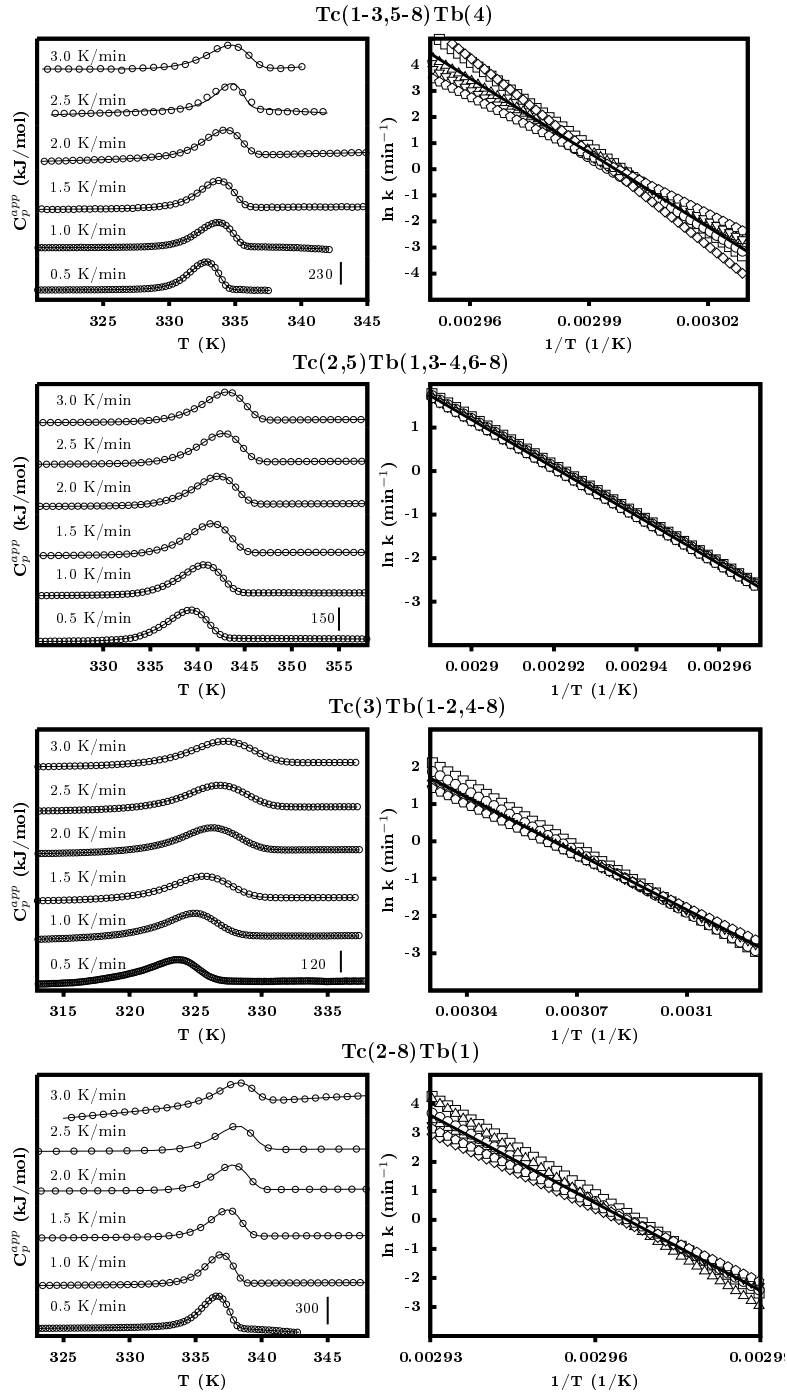


Figura 3.6: Se muestran los termogramas de DSC de cuatro quimeras a las velocidades de barrido indicadas (izquierda). La concentración de proteína fue de 0.4 mg/mL, excepto para los experimentos a 0.5 y 1 K/min de Tc(1-3,5-8)Tb(4) y 0.5 K/min de Tc(2-8)Tb(1), donde se usó 1 mg/mL. Los símbolos representan la capacidad calorífica aparente y las líneas continuas representan el ajuste al modelo de dos estados irreversible. Por claridad, los perfiles fueron desplazados en el eje de las ordenadas. Se muestran los gráficos de Arrhenius (derecha) para las mismas proteínas. Los símbolos corresponden a las diferentes velocidades de barrido empleadas: 0.5, 1.0, 1.5, 2.0, 2.5 y 3.0 K/min. Las líneas continuas representan el mejor ajuste lineal de todos los puntos.

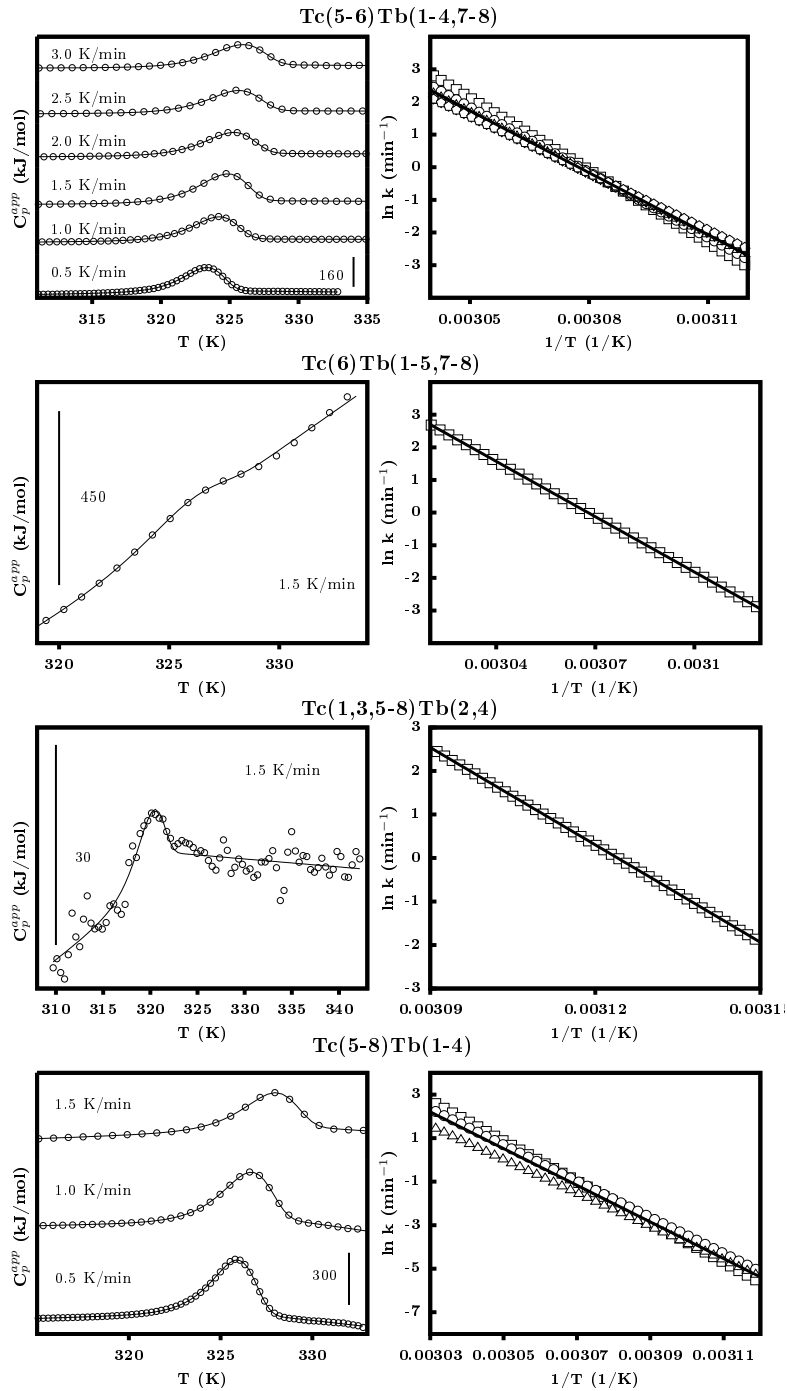


Figura 3.7: Se muestran los termogramas de DSC de cuatro quimeras a las velocidades de barrido indicadas (izquierda). La concentración de proteína fue de 0.4 mg/mL, excepto para los experimentos a 0.5, 1 y 1.5 K/min de Tc(5-8)Tb(1-4), donde se usó 1 mg/mL. Los símbolos representan la capacidad calorífica aparente y las líneas continuas representan el ajuste al modelo de dos estados irreversible. Por claridad, los perfiles fueron desplazados en el eje de las ordenadas. Se muestran los gráficos de Arrhenius (derecha) para las mismas proteínas. Los símbolos corresponden a las diferentes velocidades de barrido empleadas: 0.5, 1.0, 1.5, 2.0, 2.5 y 3.0 K/min. Las líneas continuas representan el mejor ajuste lineal de todos los puntos.

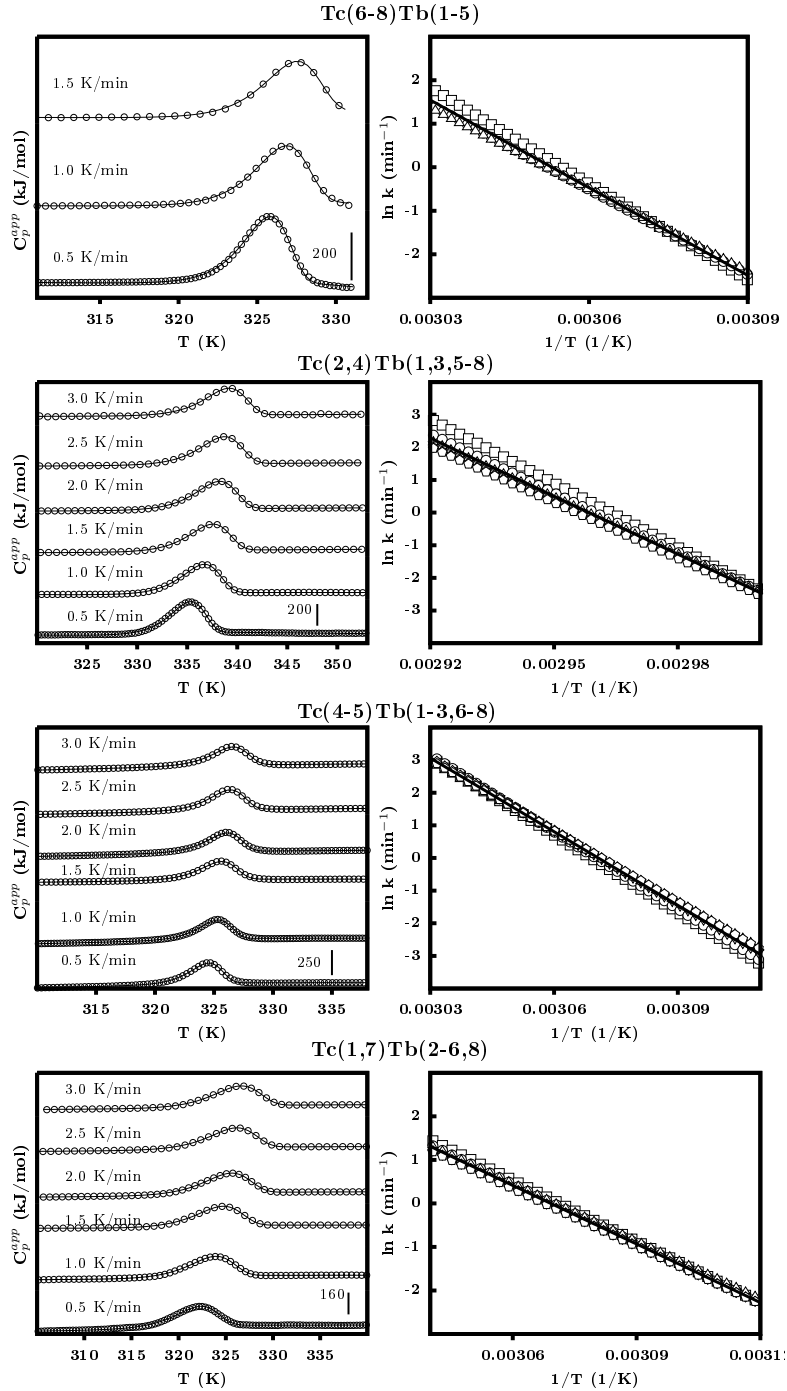


Figura 3.8: Se muestran los termogramas de DSC de cuatro quimeras a las velocidades de barrido indicadas (izquierda). La concentración de proteína fue de 0.4 mg/mL, excepto para los experimentos a 0.5, 1 y 1.5 K/min de Tc(6-8)Tb(1-5), donde se usó 1 mg/mL. Los símbolos representan la capacidad calorífica aparente y las líneas continuas representan el ajuste al modelo de dos estados irreversible. Por claridad, los perfiles fueron desplazados en el eje de las ordenadas. Se muestran los gráficos de Arrhenius (derecha) para las mismas proteínas. Los símbolos corresponden a las diferentes velocidades de barrido empleadas: 0.5, 1.0, 1.5, 2.0, 2.5 y 3.0 K/min. Las líneas continuas representan el mejor ajuste lineal de todos los puntos.

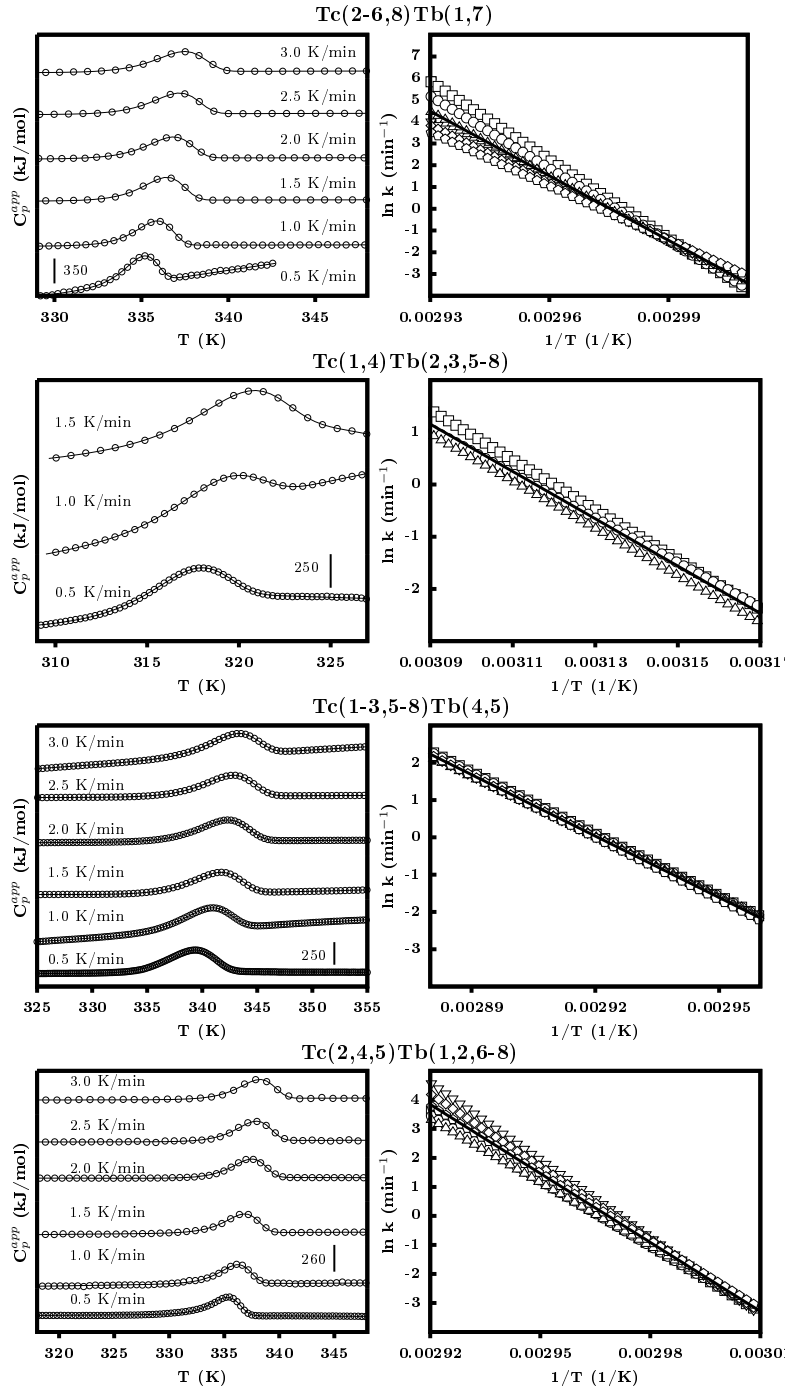


Figura 3.9: Se muestran los termogramas de DSC de cuatro quimeras a las velocidades de barrido indicadas (izquierda). La concentración de proteína fue de 0.4 mg/mL, excepto para los experimentos a 0.5, 1 y 1.5 K/min de Tc(1,4)Tb(2,3,5-8), donde se usó 1 mg/mL. Los símbolos representan la capacidad calorífica aparente y las líneas continuas representan el ajuste al modelo de dos estados irreversible. Por claridad, los perfiles fueron desplazados en el eje de las ordenadas. Se muestran los gráficos de Arrhenius (derecha) para las mismas proteínas. Los símbolos corresponden a las diferentes velocidades de barrido empleadas: 0.5, 1.0, 1.5, 2.0, 2.5 y 3.0 K/min. Las líneas continuas representan el mejor ajuste lineal de todos los puntos.

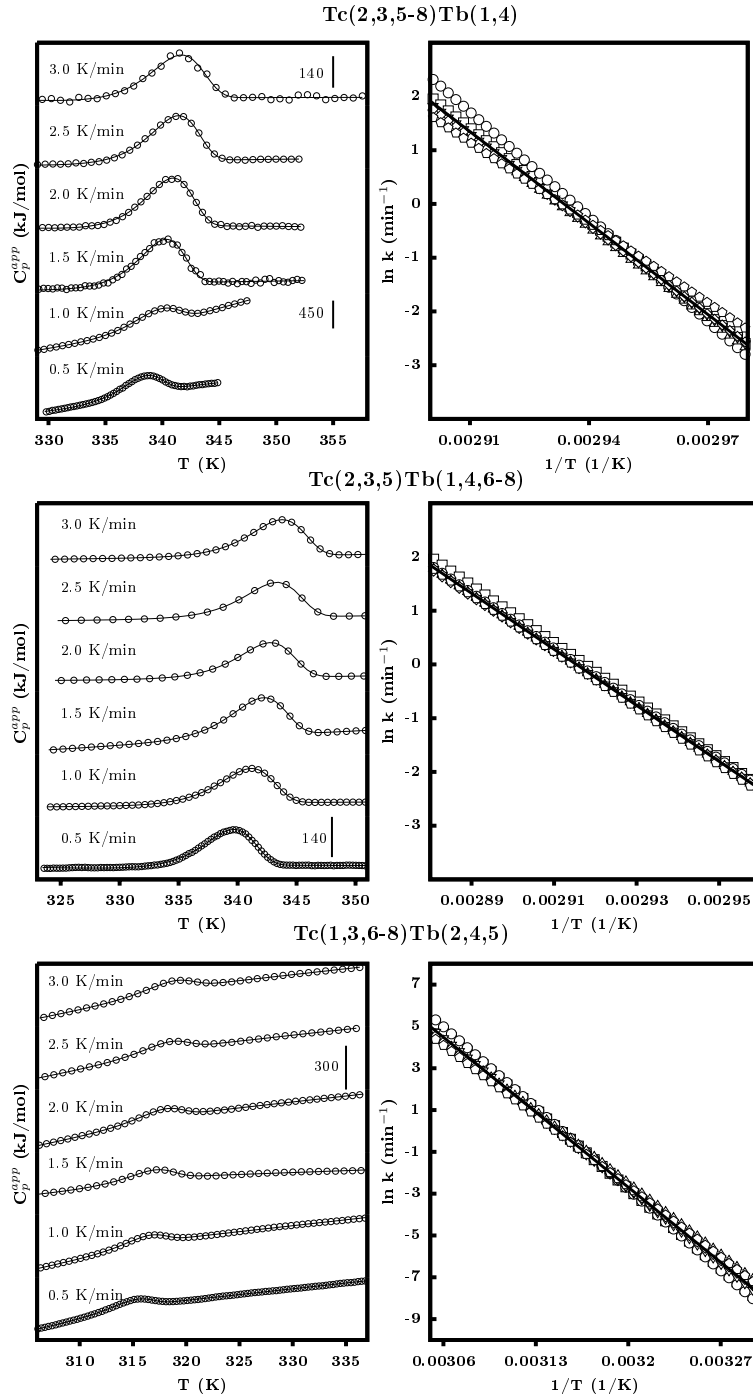


Figura 3.10: Se muestran los termogramas de DSC de tres quimeras a las velocidades de barrido indicadas (izquierda). La concentración de proteína fue de 0.4 mg/mL, excepto para los experimentos a 0.5 y 1 K/min de Tc(2,3,5-8)Tb(1,4), donde se usó 1 mg/mL. Los símbolos representan la capacidad calorífica aparente y las líneas continuas representan el ajuste al modelo de dos estados irreversible. Por claridad, los perfiles fueron desplazados en el eje de las ordenadas. Se muestran los gráficos de Arrhenius (derecha) para las mismas proteínas. Los símbolos corresponden a las diferentes velocidades de barrido empleadas: 0.5, 1.0, 1.5, 2.0, 2.5 y 3.0 K/min. Las líneas continuas representan el mejor ajuste lineal de todos los puntos.

región 1		región 2		región 3		región 4	
TbTIM	TcTIM	TbTIM	TcTIM	TbTIM	TcTIM	TbTIM	TcTIM
11 Asn	12 Asn	44 Thr	45 Thr	65 Gln	66 Gln	92 Val	93 Val
13 Lys	14 Lys	45 Phe	46 Phe	66Asn	67 Asn	95 His	96 His
14 Cys	15 Cys	46 Val	47 Leu	68 Ile	69 Ile	97 Glu	98 Glu
15 Asn	16 Asn	47 His	48 His	71 Ser	72 Ser	98 Arg	99 Arg
16 Gly	17 Gly	48 Leu	49 Ile	72 Gly	73 Gly	101 Tyr	102 Tyr
17 Ser	18 Ser	49 Ala	50 Pro	73 Ala	74 Ala	102 Tyr	103 Tyr
18 Gln	19 Glu	50 Met	51 Met	74 Phe	75 Phe		
19 Gln	20 Ser			75 Thr	76 Thr		
				76 Gly	77 Gly		
				77 Glu	78 Glu		
				78 Val	79 Val		
				79 Ser	80 Ser		
				82 Ile	83 Ile		
				83 Leu	84 Leu		
				85 Asp	86 Asp		
				86 Phe	87 Tyr		
				88 Val	89 Ile		

Tabla 3.4: Residuos de la intercara de TbTIM y TcTIM. El criterio para considerar a un residuo como parte de la intercara, es que dicho residuo tenga al menos un átomo a una distancia no mayor a 4 Å de la otra subunidad. Tabla modificada de [34].

que únicamente fueron estudiadas una pequeña fracción de quimeras que es posible construir siguiendo esta metodología (254 en total). En general, de los resultados experimentales de la tabla 3.2 se concluye que la estabilidad cinética de los barriles TIM es el resultado de una compleja red de interacciones concertadas entre regiones, interacciones locales de corto alcance y globales de largo alcance.

Con el propósito de darle continuidad al texto se han omitido de esta sección los resultados de DSC de algunas quimeras y mutantes puntuales sobre TbTIM y TcTIM que no se usaron en la correlación que se estableció entre los resultados de DSC y MDS (ver sección 3.3.2). Estos resultados se encuentran en el Apéndice B.

3.2 Simulaciones de dinámica molecular de desplegamiento por temperatura

Con el objetivo de buscar una explicación a nivel molecular de los determinantes de la estabilidad cinética y de la cooperatividad del proceso de desplegamiento, se llevaron a cabo

simulaciones de dinámica molecular a alta temperatura. A continuación se presentan los resultados del análisis general de estas simulaciones.

El desplegamiento térmico de las enzimas silvestres TbTIM y TcTIM fue estudiado en simulaciones con duración de 24 ns, a cuatro diferentes temperaturas: 400, 450, 490 y 510 K. También se llevó a cabo una simulación a 700 K de 7 ns de duración. Las temperaturas usadas en las simulaciones son artificialmente altas para acelerar el proceso de desnaturalización. Por lo tanto, se asume que las estructuras parcialmente desplegadas observadas entre 400 y 510 K son equivalentes a estructuras de desplegamiento a temperaturas menores.

A medida que las simulaciones avanzan, el desplegamiento se incrementa y se acelera (figura 3.11). A 700 K la perturbación es más significativa y rápida. La estructura permanece con las dos subunidades juntas durante todas las simulaciones, (incluso a 700 K) y el radio de giro no presenta cambios importantes durante las trayectorias (figura 3.12), lo cual se esperaría si hubiera una disociación del dímero. La unión de ambas subunidades durante las simulaciones también se puede observar en el número de puentes de hidrógeno que las unen durante las trayectorias (figura 3.13). Si bien se observa una fluctuación, estos se mantienen oscilando entre 10 y 20 durante el tiempo de simulación. Únicamente a 700 K se observa que tienden a cero, pero no se pierden por completo. Aunque el objetivo de realizar estas simulaciones no fue determinar o reproducir la ruta de desplegamiento de las proteínas silvestres, cabe destacar que el hecho de que observemos el desplegamiento sin separación previa de las subunidades, es consistente con la conclusión alcanzada con el modelo al que se han ajustado las trazas calorimétricas (ver sección 2.3).

Los puentes de hidrógeno totales en la proteína y aquellos formados entre la proteína y el solvente se muestran en la figura 3.14. Observamos lo que se espera para las simulaciones a alta temperatura: a medida que ésta aumenta, los puentes de hidrógeno dentro de la proteína decrecen, mientras que los puentes formados entre la proteína y el solvente se incrementan. No se observa ninguna diferencia significativa entre los patrones de formación y ruptura de los puentes de hidrógeno entre TbTIM y TcTIM. Algo similar se observa en el área expuesta al solvente de ambas proteínas (figura 3.15).

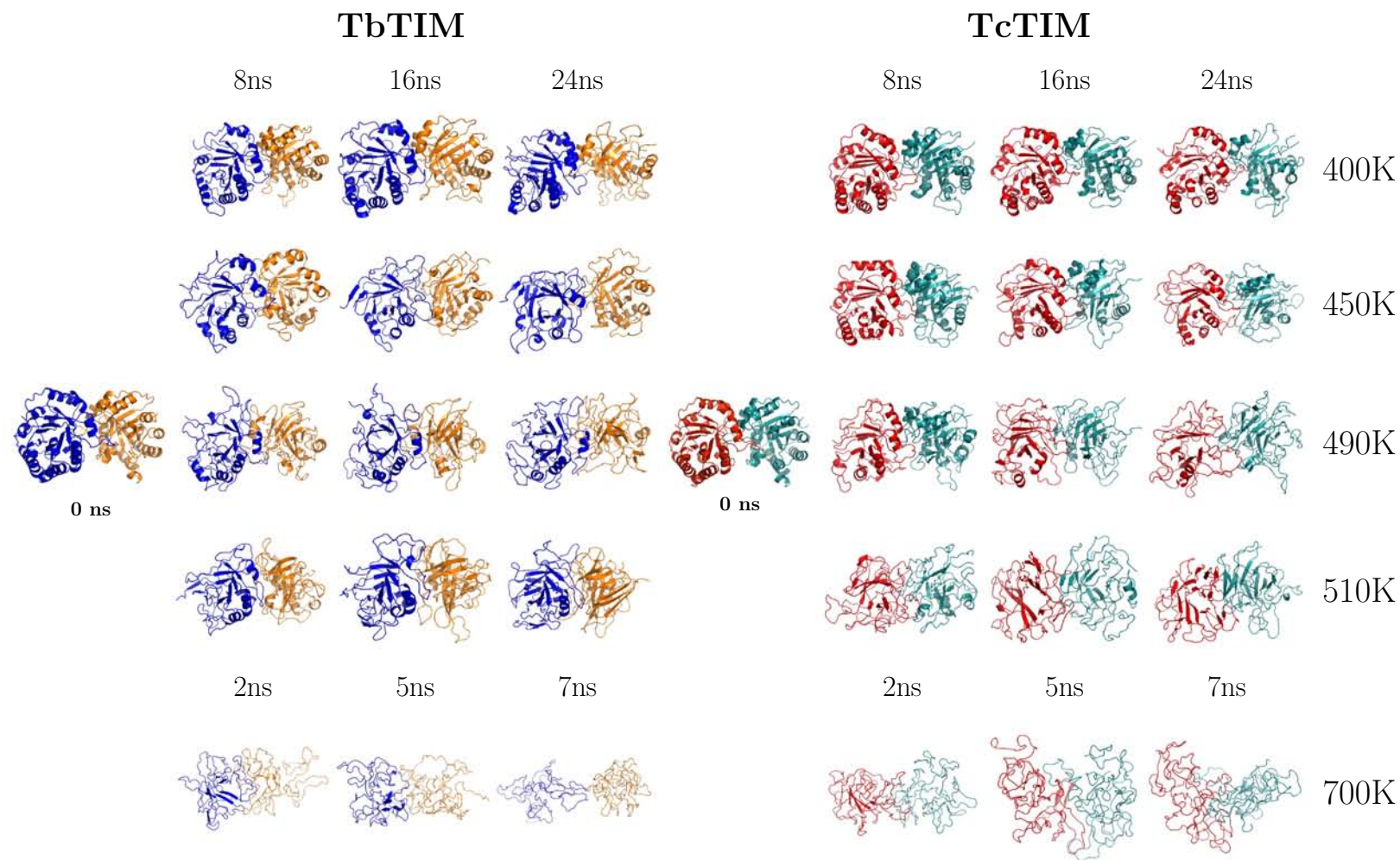


Figura 3.11: Imágenes de las trayectorias a 8, 16 y 24 ns de las simulaciones a 400, 450, 490 y 510 K. De las simulaciones a 700 K se muestran imágenes a 2, 5 y 7 ns. Las subunidades A y B de TbTIM se muestran en azul y naranja, mientras que las de TcTIM se muestran en rojo y cian, respectivamente.

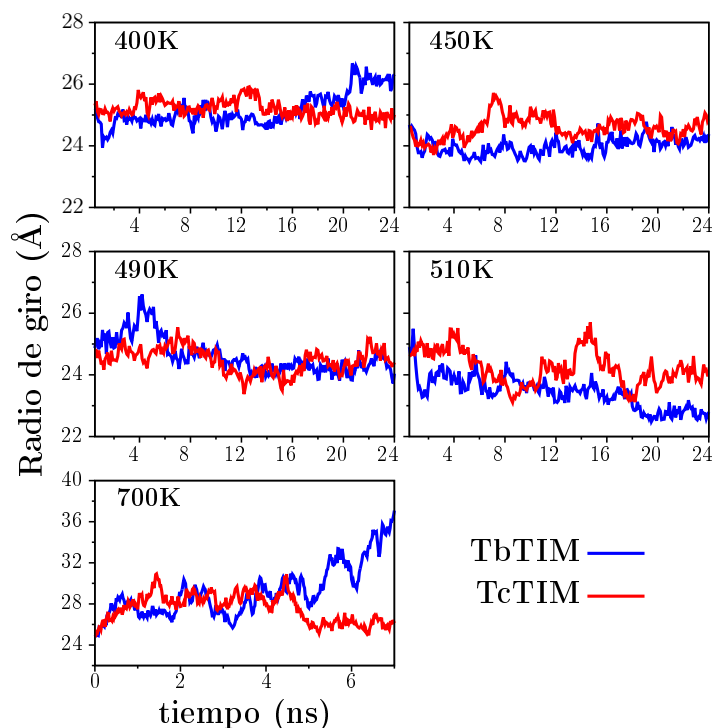


Figura 3.12: Radio de giro de TbTIM y TcTIM a las temperaturas de simulación.

La desviación cuadrática media (RMSD) de los C- α a lo largo de las trayectorias se muestra en la figura 3.16. El valor de C α -RMSD de la estructura completa de TbTIM es independiente y de mayor magnitud que los valores correspondientes a las subunidades. Esto es indicativo de un desacoplamiento entre ellas, es decir, un cambio en la estructura cuaternaria. Para TcTIM, este desacoplamiento solo es evidente en las simulaciones de temperatura más alta, 510 y 700 K. Este desacoplamiento de las subunidades puede deberse a un movimiento relativo entre ellas, como por ejemplo la contra-rotación observada en la TIM de pollo y en TcTIM [14, 63] en simulaciones de dinámica molecular a 300 K.

Las ocho hélices α de la estructura se despliegan fácilmente con la temperatura (figura 3.17). En contraste, la parte central del barril, formada por ocho hebras β paralelas, es mucho más resistente en las dos subunidades de ambas proteínas. La excepción se observa en las trayectorias a 700 K en donde se pierde por completo la estructura secundaria en menos

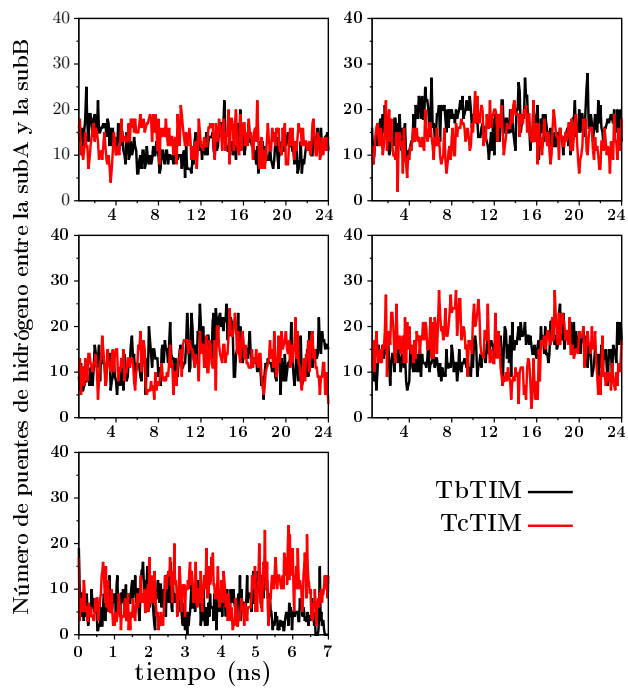


Figura 3.13: Puentes de hidrógeno que forman las subunidades de TbTIM y TcTIM

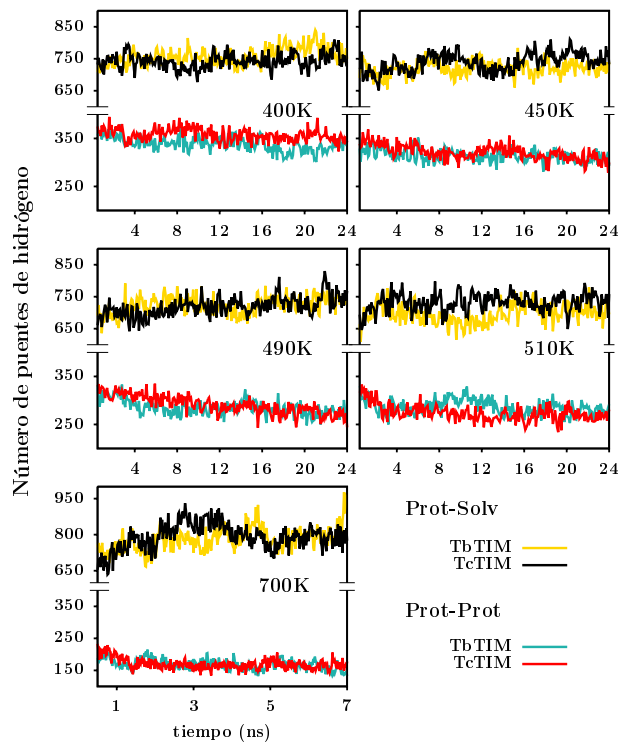


Figura 3.14: Número de puentes de hidrógeno totales de la proteína y formados entre la proteína y el solvente

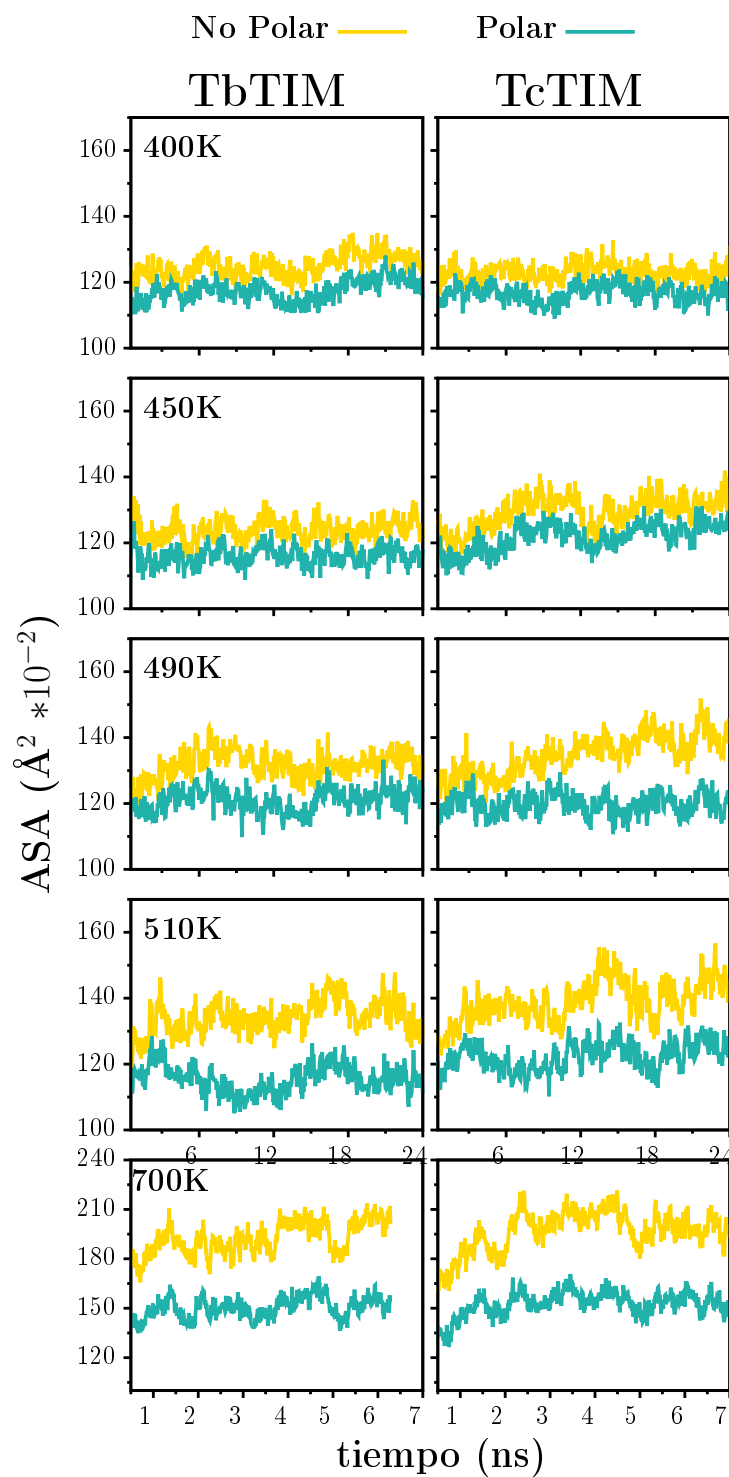


Figura 3.15: Área accesible al solvente durante las simulaciones a alta temperatura. Se muestran por separado las contribuciones polares y no polares

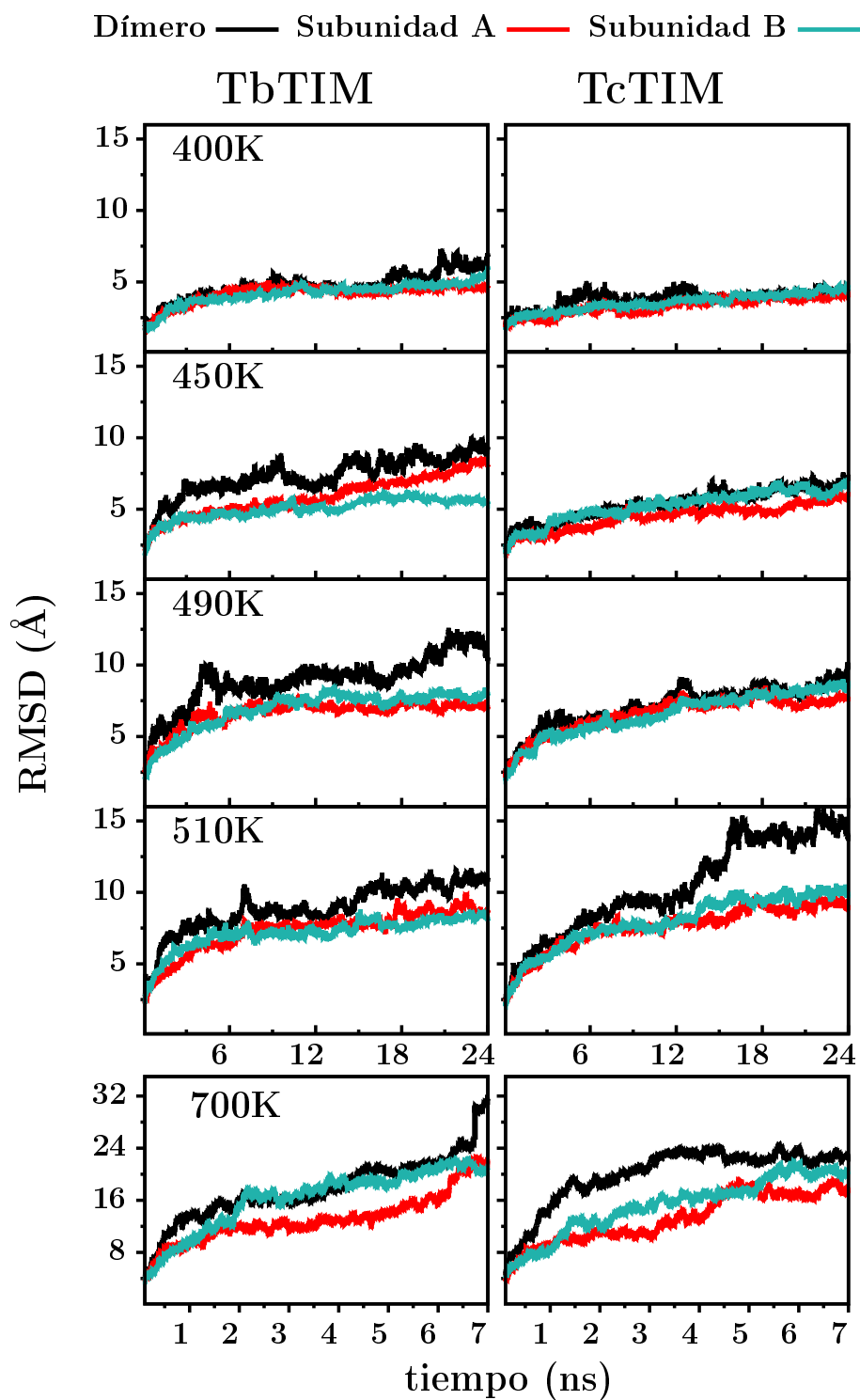


Figura 3.16: Desviación cuadrática media (RMSD) de los $C\alpha$ de TbTIM y TcTIM en función del tiempo de simulación. Se muestran las cinco temperaturas empleadas.

de 7 ns. Pero aún a esa temperatura se observa una mayor labilidad en las hélices α . Esto indica que el desplegamiento inducido por temperatura del barril TIM comienza desde el exterior o la “envoltura” y continua con el centro o la parte interna de hojas β .

A 400 y 450 K, la pérdida de la estructura de las hélices α se observa a un ritmo definido para ambas proteínas. En contraste, a 490 y 510 K la pérdida de la estructura secundaria es más evidente y se observa a tiempos más cortos de simulación, aunque con algunas diferencias notables para ambas proteínas. Las hélices de TcTIM se despliegan por completo a 490 K y 510 K después de 12 y 8 ns, respectivamente, mientras que en TbTIM, las hélices muestran más resistencia. A estas temperaturas, la pérdida de las hélices de TbTIM es comparable a la que se observa en TcTIM durante los primeros 4 ns de simulación, pero después continúa a un ritmo mucho menor. Esta diferencia entre ambas enzimas puede atribuirse a un proceso más cooperativo para TcTIM que para TbTIM, en concordancia con los valores de W obtenidos de las trazas calorimétricas (tabla 3.2). El ancho del pico es una medida de la cooperatividad del proceso de desnaturalización térmica. Cuando tiene un valor pequeño, el perfil de la traza es más angosto y la cooperatividad del proceso de desplegamiento es mayor [13, 16].

También se llevaron a cabo simulaciones al equilibrio a 298 K de ambas enzimas silvestres. Con el propósito de no perder continuidad en este escrito, los resultados del análisis general de estas simulaciones se presentan en el Apéndice C.

3.2.1 Agrupación conformacional de las estructuras en clústers

Para este análisis se tomaron las estructuras generadas por las simulaciones cada 4 ps y se agruparon en clústers según su conformación, utilizando como criterio la posición de los $C\alpha$ (ver sección 2.4.3). La proyección en tres dimensiones del espacio generado por la agrupación de las estructuras se muestra en la figura 3.18. Las simulaciones de 700 K no se tomaron en cuenta para este análisis porque el desplegamiento observado en ellas es muy rápido y la información que se genera de agrupar las estructuras no es útil para los propósitos del análisis posterior.

Las estructuras representativas de cada clúster se definieron con la ayuda de una metodología desarrollada para este trabajo que se nombró dispersión progresiva de movilidad (PMD)

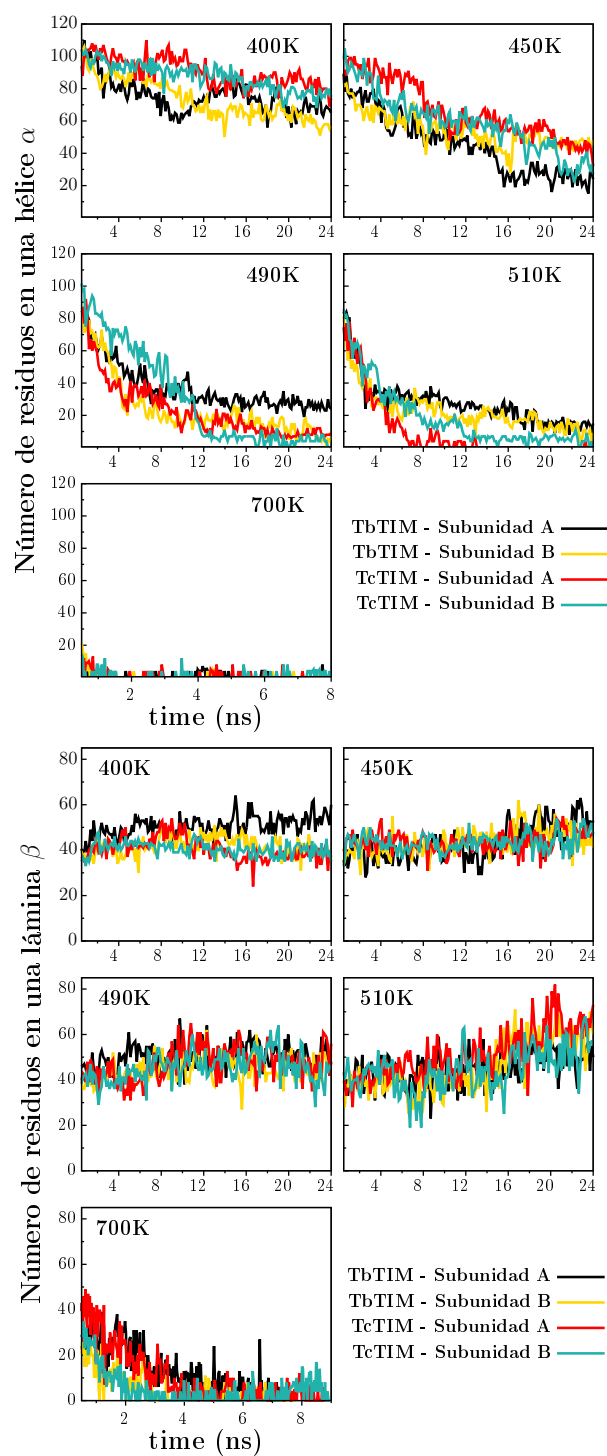


Figura 3.17: Número de residuos que forman parte de las hélices α y las hebras β de TbTIM y TcTIM, a lo largo de las trayectorias. Las subunidades de cada proteína se muestran por separado

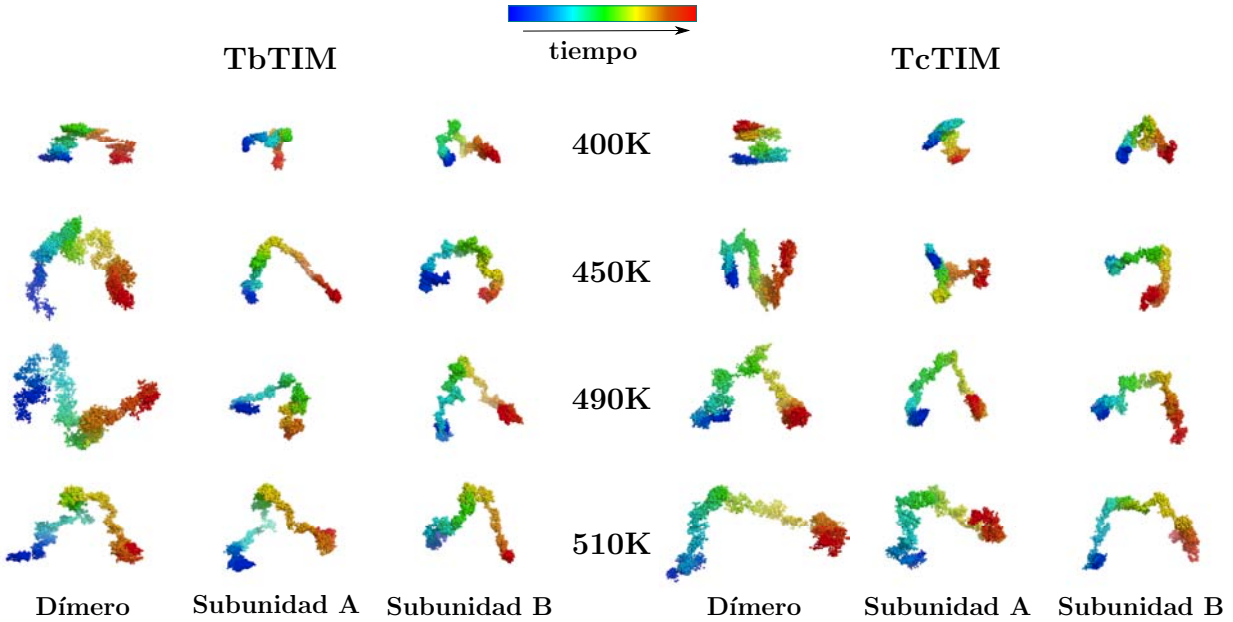


Figura 3.18: Proyección de las trayectorias en un espacio 3D de los valores de desviación cuadrática media de los $C\alpha$ del dímero y de las subunidades de TbTIM y TcTIM a 400, 450, 490 y 510 K. Se usó la misma escala y la misma orientación en todas las imágenes. El tiempo de simulación se representa con el gradiente de color indicado.

(ver la sección de 2.4). Los resultados arrojados por este análisis se muestran en la figura 3.19. Para elegir las estructuras representativas de cada clúster, se seleccionaron los tiempos en donde los mínimos correspondientes al dímero y a ambas subunidades para cada proteína coinciden. El total de clústers identificado de esta manera se encuentra en un intervalo de 11 a 16 para todas las temperaturas de ambas proteínas.

3.2.2 Desplazamiento de residuos y regiones

Los desplazamiento de cada residuo ($\Delta d_{j,c}$) obtenidos para cada clúster de cada trayectoria, son una medida cuantitativa de la posición de cada residuo con respecto al resto de la proteína (ver sección 2.4.3). Los valores de $\Delta d_{j,c}$ son negativos cuando el residuo se acerca en promedio al resto de la proteína y son positivos cuando el residuo se aleja en promedio de la proteína. Como se esperaba, los valores de $\Delta d_{j,c}$ crecen en magnitud cuando la temperatura aumenta (ver el ejemplo mostrado en la figura 3.20). Sin embargo, a este nivel es difícil encontrar tendencias claras comparando ambas proteínas y sus subunidades.

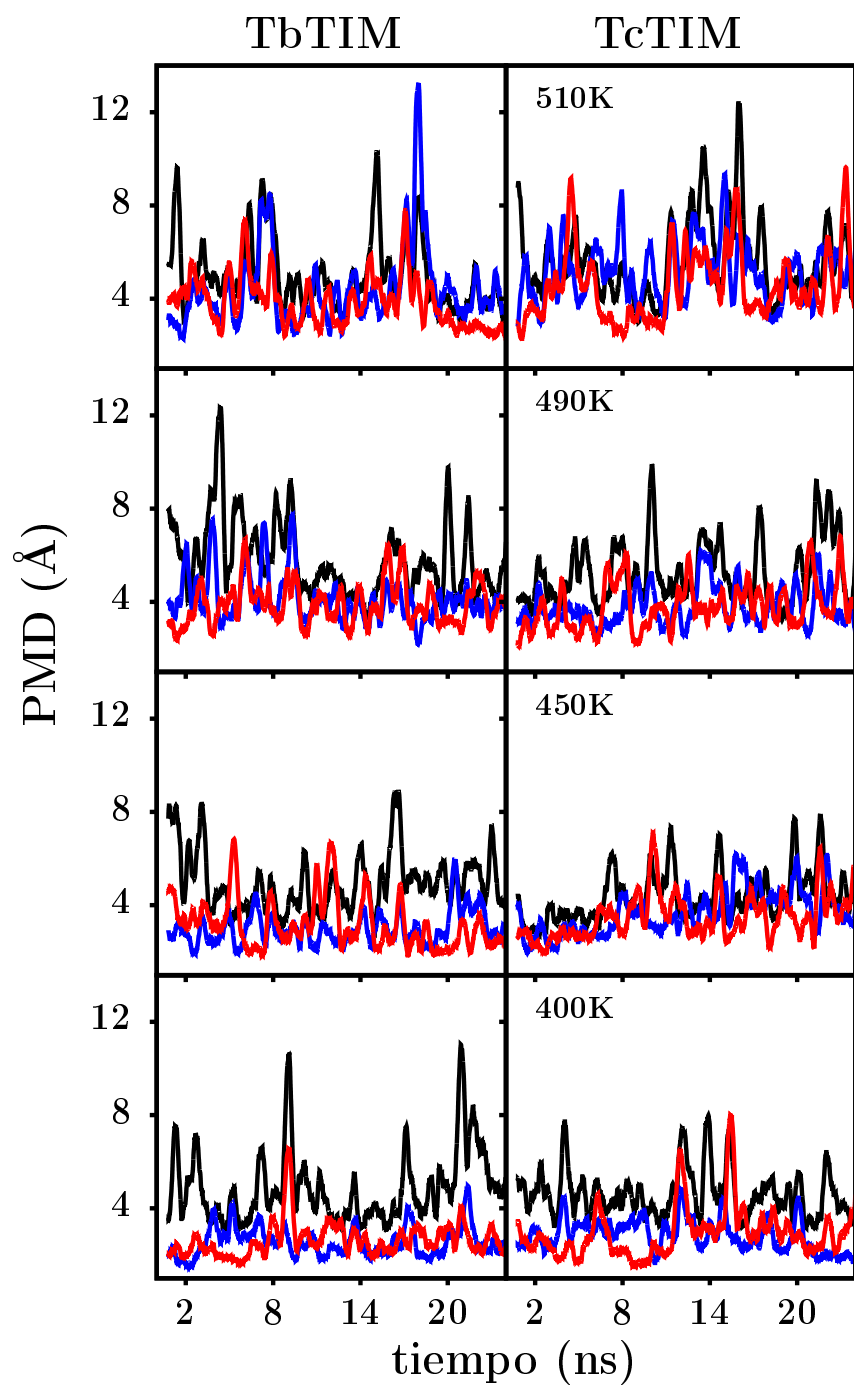


Figura 3.19: Gráficas de la dispersión progresiva de movilidad sobre grupos de 200 puntos para todas las trayectorias. Los mínimos representan estructuras que permanecen relativamente estables durante una parte significativa de la trayectoria, mientras que los máximos representan segmentos de la trayectoria donde la estructura cambia rápidamente. En azul se muestra la subunidad A, en rojo la subunidad B y en negro el dímero.

La situación cambia cuando los residuos son agrupados en las ocho regiones mostradas en la figura 3.1. Los valores de desplazamientos por región, $D_{r,T}$ (ver sección 2.4.3), revelan movimientos globales que sufre la proteína con el aumento de la temperatura. En la figura 3.21 se muestra el valor de $D_{r,T}$ a 400 y 510 K. En esta figura se puede apreciar que el aumento de $D_{r,T}$ es moderado para TbTIM y de mayor magnitud para TcTIM. Es decir, la diferencia entre los desplazamientos de residuos para TbTIM y TcTIM se hace evidente cuando se comparan las trayectorias a diferentes temperaturas de simulación.

3.2.3 Dependencia con temperatura de los desplazamientos (TGD)

El siguiente paso consiste en considerar los desplazamientos de los residuos en las proteínas completas. Al obtener los valores de desplazamiento total por proteína a cada temperatura ($D_{wt,T}$) las diferencias entre TbTIM y TcTIM son más claras. En la figura 3.22 se observa que a 400 K, la temperatura más baja, el valor de $D_{wt,T}$ es más grande para TbTIM que para TcTIM, mientras que a 510 K, la temperatura más alta, la situación es la inversa. Ajustando la variación de $D_{wt,T}$ con la temperatura a una línea recta, se obtiene un nuevo parámetro llamado TGD (temperature gradient of the displacement). El valor de TGD total para TbTIM es $1.5 \pm 0.10 \text{ \AA/K}$ y para TcTIM es $2.12 \pm 0.19 \text{ \AA/K}$, lo que se interpreta como una mayor variación de la posición de los residuos de TcTIM que de TbTIM, al aumentar la temperatura.

El valor de TGD de cada proteína se puede descomponer en las ocho regiones en las que se han dividido las estructuras. Esto se logra ajustando a una línea recta los desplazamientos englobados en regiones, $D_{r,T}$, en función de la temperatura (figura 3.23) para cada proteína. El resultado de comparar los TGD_r entre TbTIM y TcTIM se muestra en la figura 3.24. En esta representación resaltan claramente las regiones 4 y 5 como aquellas que presentan una mayor diferencia en los valores de TGD_r entre TbTIM y TcTIM. El número de residuos que son diferentes entre ambas proteínas (Tabla 3.1) son 5 para la región 4 y 15 para la región 5. Ambas regiones ocupan la parte central de la secuencia de las proteínas silvestres (Figura 3.1) y la región 4 también forma parte de la intercara del dímero (Tabla 3.4).

Estos resultados refuerzan las conclusiones obtenidas en el estudio de calorimetría de las

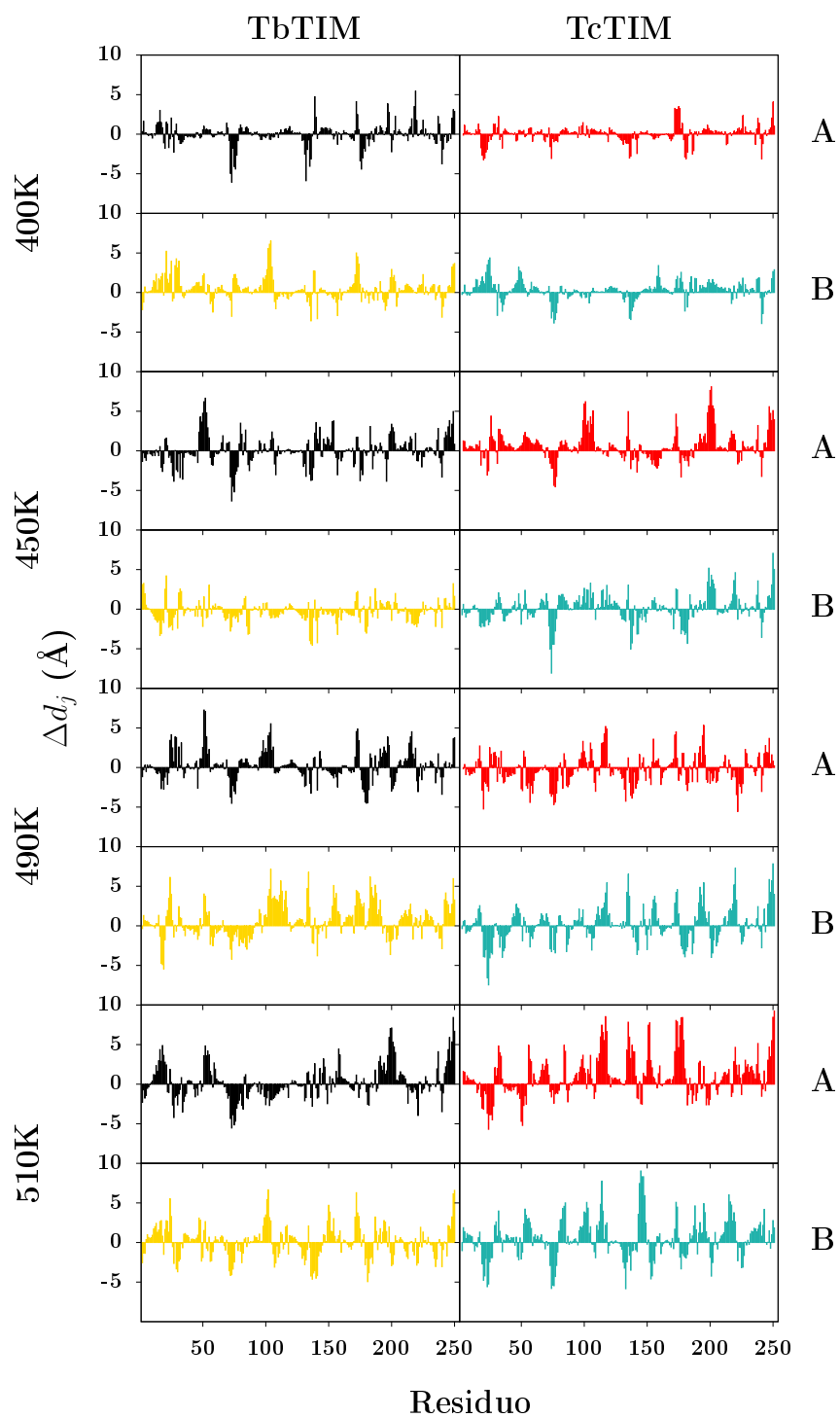


Figura 3.20: Diferencias en las distancias promedio de cada $C\alpha$ entre la estructura del último clúster de cada trayectoria y la estructura inicial de TbTIM y TcTIM para ambas subunidades (A y B). Estos valores se calcularon utilizando el análisis por subunidad.

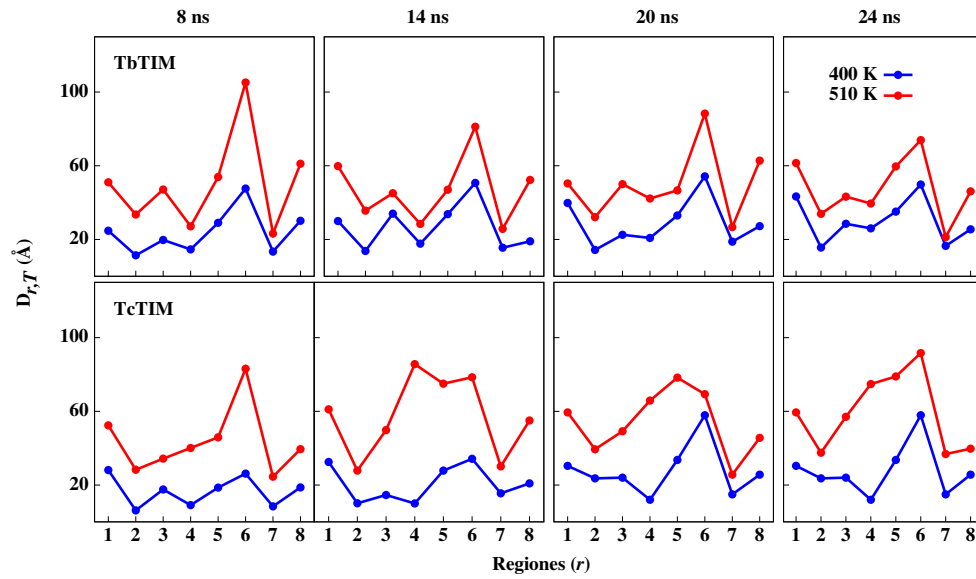


Figura 3.21: Desplazamiento de las regiones de los clústers que se encuentran alrededor de 2, 8, 14, 20 y 24 ns. Se muestra un promedio de ambas subunidades para cada proteína.

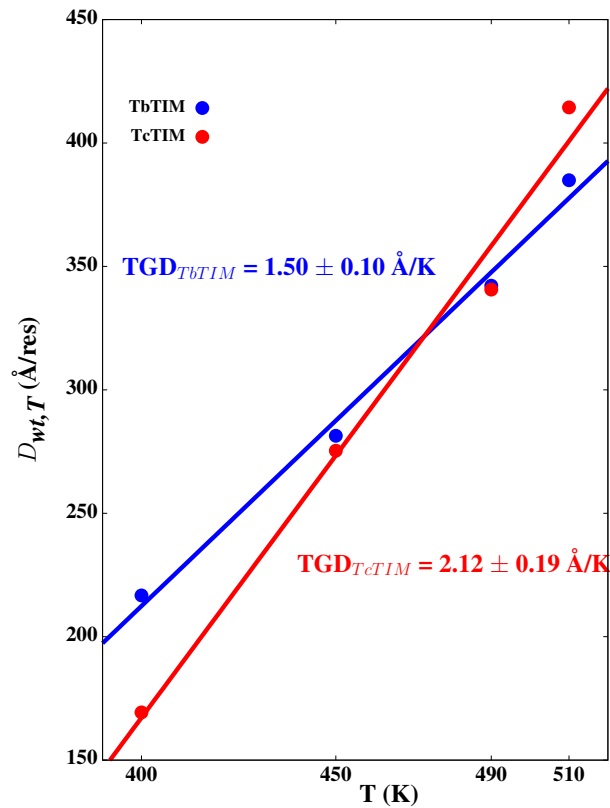


Figura 3.22: Desplazamiento total por proteína a cada temperatura para TbTIM y TcTIM

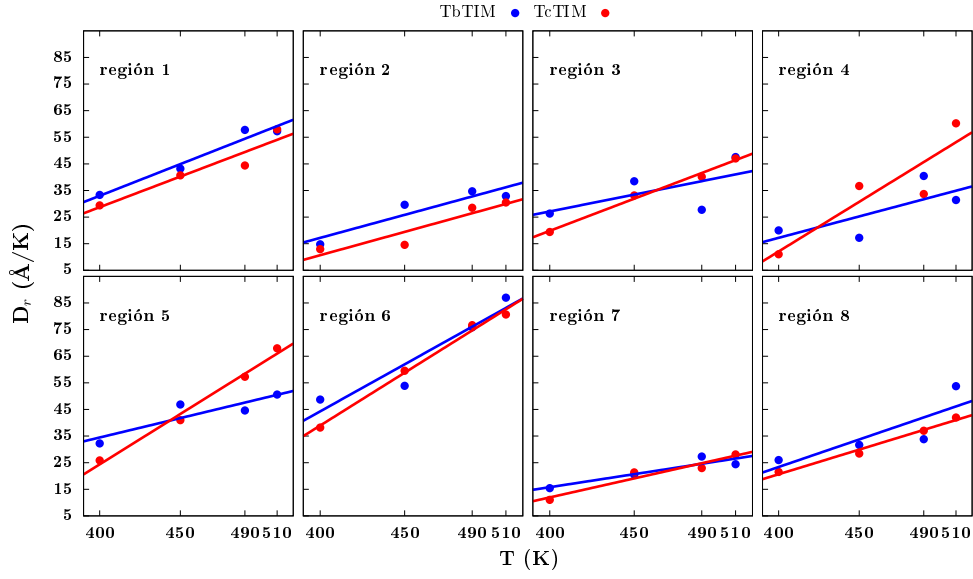


Figura 3.23: Desplazamiento total por región a cada temperatura para las ocho regiones de TbTIM y TcTIM. Las líneas rectas son el mejor ajuste lineal de los datos de desplazamiento (puntos).

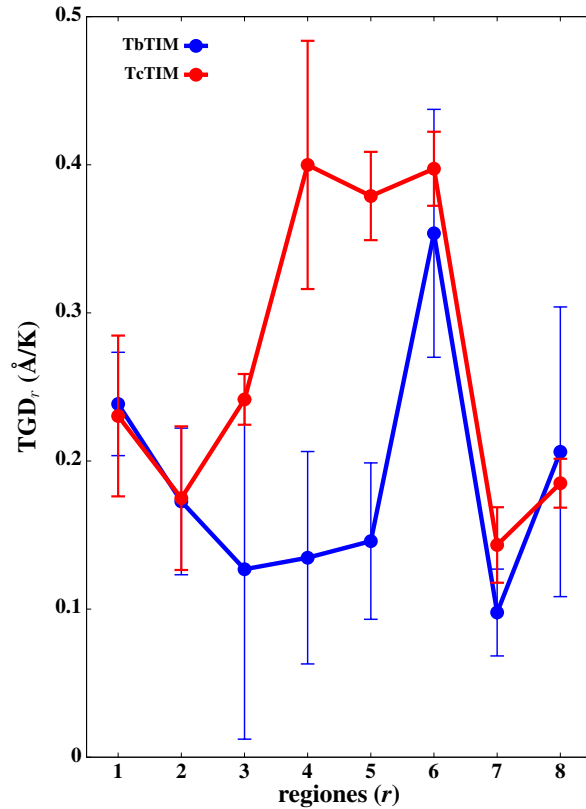


Figura 3.24: Dependencia con la temperatura de los desplazamientos por región (TGD_r). Las barras de error se obtuvieron de las incertidumbres de las pendientes de los ajustes lineales de $D_{r,T}$ vs. T .

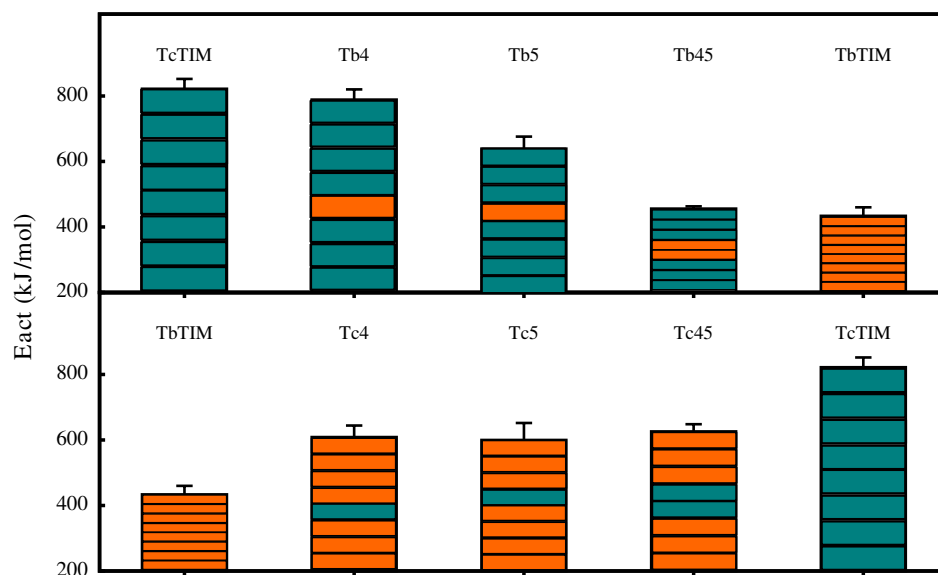


Figura 3.25: Efecto del intercambio de las regiones 4 y 5 en la E_{act} de TbTIM y TcTIM.

quimeras (sección 3.1.2): la región 5 contribuye en gran medida a la estabilidad cinética de TcTIM, y las regiones 4 y 5 juegan un papel importante en la baja estabilidad cinética de TbTIM. De hecho, el efecto de colocar las regiones 4 y 5 de TbTIM en TcTIM es una gran disminución en la E_{act} de TcTIM (Figura 3.25), es decir, la E_{act} de la quimera resultante (Tc(1-3,6-8)Tb(4,5)) es prácticamente la misma que la de TbTIM. El camino inverso es tratar de estabilizar a TbTIM intercambiando sus regiones 4 y 5 por las de TcTIM. La figura 3.25 muestra que la contribución individual de las regiones en el valor de la E_{act} es significativa. Sin embargo, el efecto no es aditivo al colocar ambas regiones juntas, ya que la E_{act} sube a 626 kJ/mol pero no alcanza el valor de 822 kJ/mol que tiene TcTIM. Esto puede deberse a la forma arbitraria en la que se eligieron las regiones de ambas estructuras para diseñar las quimeras. En la última sección de este trabajo (3.4) se presentará la contribución de cada residuo al TGD de cada región, en donde la definición de región no es relevante y de donde se pueden obtener pequeños grupos de residuos que contribuyen a la diferencia en TGD de ambas proteínas y probablemente a su diferencia en E_{act} . Es necesario volver a diseñar quimeras basándonos en los resultados de TGD_{res} o en otro tipo de análisis para tratar de obtener una proteína cuya secuencia sea predominantemente la de TbTIM, pero con la E_{act} de TcTIM.

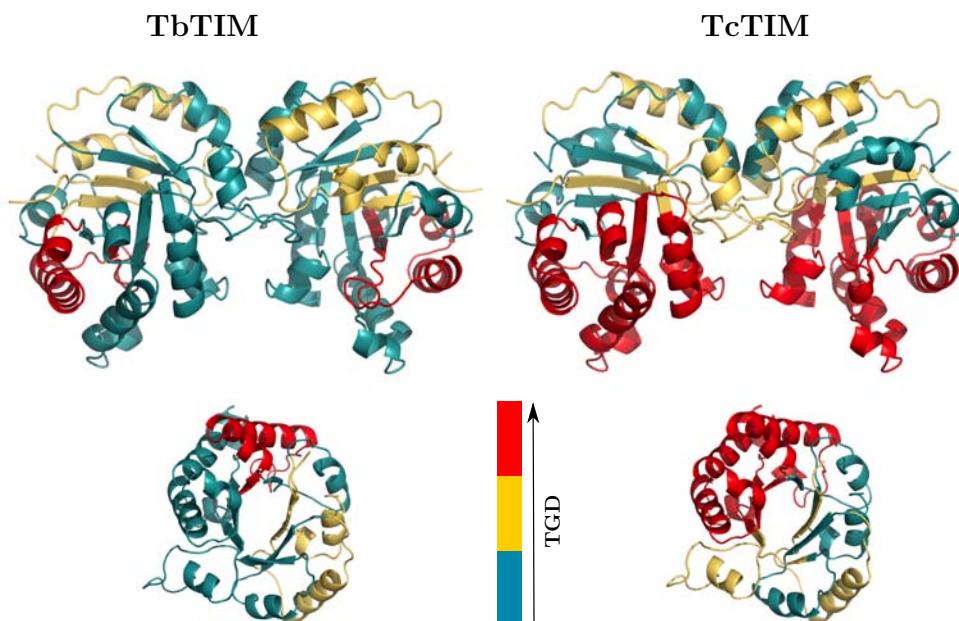


Figura 3.26: Unidades de desplegamiento de TbTIM y TcTIM. En azul se muestran las regiones que tienen un valor de TGD de entre 0 y 2 Å/K, en amarillo las que tienen un valor de TGD entre 2 y 3 Å/K y en rojo las que tienen un valor de TGD mayor a 3 Å/K

Otra diferencia interesante entre TbTIM y TcTIM que se desprende de la figura 3.24 son las unidades de desplegamiento constituidas por regiones que tienen un valor de TGD similar. En la figura 3.26 se representan las regiones de ambas proteínas en tres colores según su valor de TGD : 0-2 Å/K en azul, 2-3 Å/K en amarillo y mayor a 3 Å/K en rojo. Es interesante resaltar que las regiones con valores similares de TGD se encuentran cercanas en la estructura. TcTIM tiene tres regiones con alto TGD (4, 5 y 6) que se encuentran a la mitad de la secuencia. En TbTIM, la única región con un valor de TGD alto es la región 6.

3.3 Flexibilidad térmica y estabilidad cinética

3.3.1 Definición de flexibilidad térmica

La flexibilidad estática de una proteína puede ser cualitativamente entendida como el tamaño del espacio conformacional accesible para su estructura, mientras que la flexibilidad dinámica refleja la rapidez con la que la proteína visita las distintas conformaciones accesibles [111]. Los

desplazamientos calculados a cada temperatura en este trabajo (D_X) se pueden considerar una medida cuantitativa de las flexibilidades estáticas locales y globales de las proteínas silvestres, TbTIM y TcTIM. Cuando se aumenta la temperatura de la caja de simulación, se está suministrando energía al sistema, lo que aumenta los valores de $\Delta d_{j,c}$, $D_{r,T}$ y $D_{wt,T}$. La dependencia del aumento de la temperatura de estos desplazamientos son cuantificados por sus respectivos valores de TGD . A este parámetro se le denotará en este trabajo como flexibilidad térmica, un término inspirado en el de flexibilidad dinámica de Dill y colaboradores [111], con la distinción de que en este caso la variable relevante es la temperatura y no el tiempo. Por lo tanto, la flexibilidad térmica puede entenderse como el incremento en el espacio conformacional accesible a la proteína como resultado de un aumento en la temperatura.

3.3.2 Correlación entre la estabilidad cinética y la flexibilidad térmica

De acuerdo a los modelos simples de Lumry-Eyring [103,104], la estabilidad cinética está ligada a la escala de tiempo en la que una proteína perderá su estructura funcional irreversiblemente. Para que esto ocurra la proteína debe de pasar por un estado conformacional “límite” a partir del cual la desnaturalización ocurrirá irreversiblemente. Este estado es el estado de transición. Para acceder a él es necesario que la proteína cruce una barrera energética de magnitud E_{act} , que representa la diferencia en entalpía entre el estado nativo de la proteína y su estado de transición. Por otra parte, TGD o la flexibilidad térmica refleja un aumento del espacio conformacional accesible a la proteína cuando se suministra energía (térmica) al sistema. El hecho de que TcTIM tenga un valor mayor de E_{act} y de TGD que TbTIM y que específicamente las regiones 4 y 5 con valores muy distintos de TGD entre ambas enzimas, modifiquen sus E_{act} (Figura 3.25) sugiere que la estabilidad cinética y la flexibilidad térmica pueden estar correlacionadas.

Para comprobar esta hipótesis se calcularon los valores de TGD de una serie de quimeras que van transformando una enzima silvestre en la otra, mediante el reemplazo progresivo de sus regiones. Estos valores se graficaron en función de los valores de E_{act} presentados en la tabla 3.2. El resultado es una correlación entre la estabilidad cinética (E_{act}) y la flexibilidad térmica (TGD) de la proteína (Figura 3.27a).

La habilidad predictiva de esta correlación fue probada con las otras veintitrés quimeras presentadas en la tabla 3.2. En la figura 3.27b se compara los valores de E_{act} con la línea recta predicha por la correlación en la figura 3.27a. Pese a la dispersión que se observa, se puede concluir que los valores experimentales tienen una tendencia clara y concuerdan razonablemente bien con los predichos. La dispersión puede deberse a muchos factores, principalmente dos: 1) asumir aditividad para calcular el valor de TGD de las quimeras y 2) que el análisis solo se basa en el desplazamiento de los C_α de las estructuras.

Por otro lado, se ha encontrado una relación inversa entre la E_{act} obtenida por calorimetría y la cooperatividad de la traza calorimétrica [19]. Esta cooperatividad se traduce como el ancho de la endoterma a la mitad de su altura (W). La cooperatividad de desplegamiento se puede entender como un desplazamiento concertado de los residuos que en última instancia conducen a un estado desplegado. Entre más concertado sea el desplazamiento de los residuos, la cooperatividad del proceso será mayor. A través de simulaciones de dinámica molecular de alta temperatura se ha sugerido que la alta cooperatividad de desplegamiento de la proteasa α -lítica, está ligada a su alta estabilidad cinética [101]. Por estas razones, los valores de W (tabla 3.2) deben también correlacionarse con los valores de TGD , como de hecho ocurre (figuras 3.28a y 3.28b).

Se ha probado que la estabilidad cinética de una proteína y la cooperatividad de su desplegamiento están ligadas a las barreras de solvatación que existen debido a la asincronía entre la ruptura de las interacciones presentes en el estado nativo y la formación de nuevas interacciones con el solvente en un proceso de desplegamiento inducido por un suministro de energía al sistema [19, 68, 97]. Los eventos de ruptura y formación de nuevas interacciones en el proceso de desplegamiento de una proteína necesariamente incluye el desplazamiento de los residuos, por lo que, pese a que no se logró ver una diferencia significativa en el área accesible al solvente (figura 3.15) ni en los puentes de hidrógeno entre la proteína y el solvente (figura 3.14), es probable que las barreras de solvatación estén relacionadas con la flexibilidad térmica cuantificada por TGD . Un análisis más detallado de las interacciones que se forman y rompen entre el solvente y la proteína durante las simulaciones de dinámica molecular es necesario para comprobar esta hipótesis.

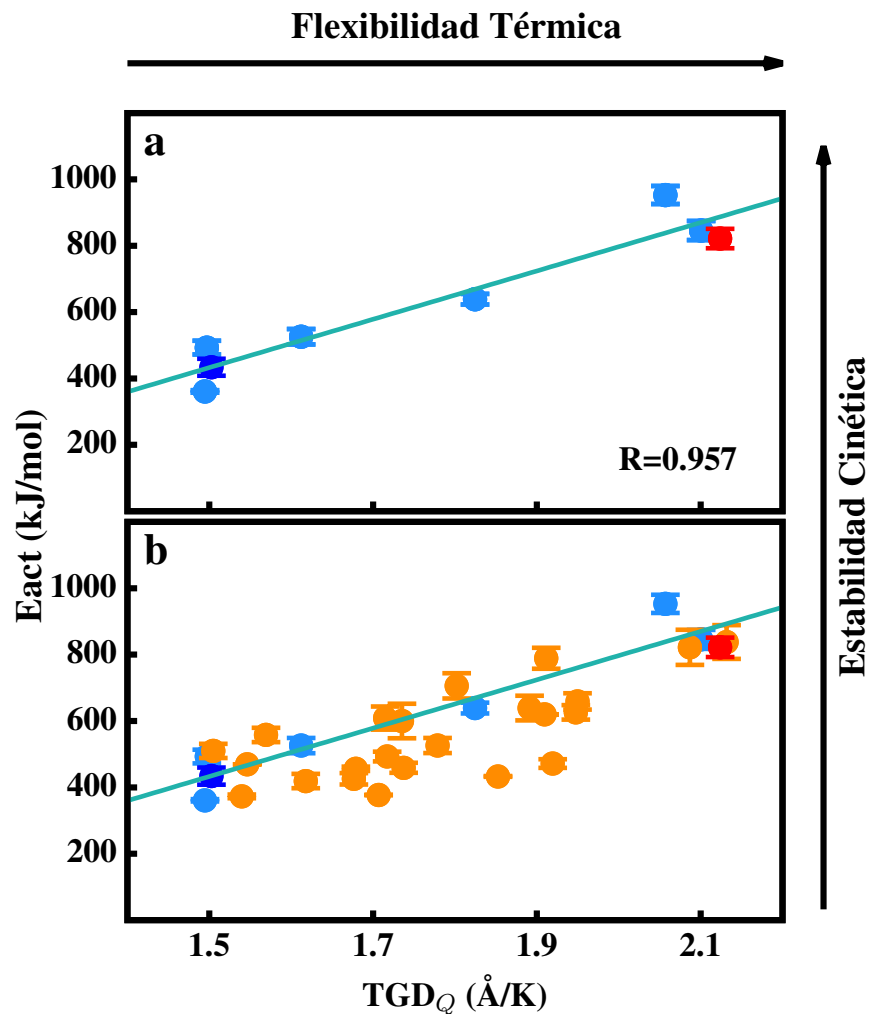


Figura 3.27: (a) Gráfica de E_{act} contra TGD para seis quimeras en donde TbTIM es progresivamente convertida en TcTIM: Tc(1)Tb(2-8), Tc(1-2)Tb(3-8), Tc(1-3)Tb(4-8), Tc(1-4)Tb(4-8), Tc(1-5)Tb(6-8), Tc(1-6)Tb(7-8) (círculos azul claro). El círculo color azul rey corresponde a TbTIM y el rojo a TcTIM. Las líneas rectas corresponden al mejor ajuste de los datos ($E_{act} = -660.01 + 728.54TGD$). (b) Los símbolos naranja corresponden a la E_{act} contra TGD para las demás quimeras mostradas en la tabla 3.2. La línea recta mostrada en (b) es la misma que en (a). Las flechas indican la dirección en la que crece la estabilidad cinética y la flexibilidad térmica.

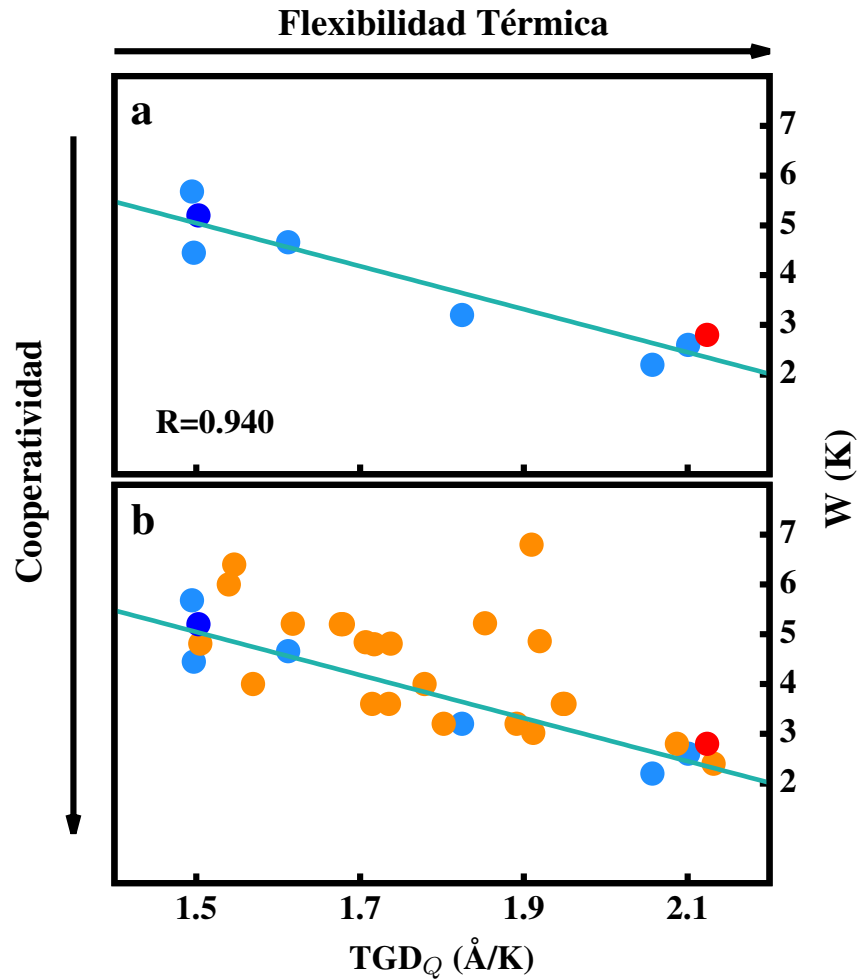


Figura 3.28: (a) Gráfica de W contra TGD para seis quimeras en donde TbTIM es progresivamente convertida en TcTIM: Tc(1)Tb(2-8), Tc(1-2)Tb(3-8), Tc(1-3)Tb(4-8), Tc(1-4)Tb(4-8), Tc(1-5)Tb(6-8), Tc(1-6)Tb(7-8) (círculos azul claro). El círculo color azul rey corresponde a TbTIM y el rojo a TcTIM. La línea recta corresponde al mejor ajuste de los datos ($W = 11.53 - 4.33TGD$). (b) Los símbolos naranja corresponden a los valores de W contra TGD para las demás quimeras mostradas en la tabla 3.2. La línea recta mostrada en (b) es la misma que en (a). Las flechas indican la dirección en la que crece la cooperatividad de la transición de desplegamiento y la flexibilidad térmica.

Los resultados mostrados en la figura 3.27 parecen contradecir la afirmación general de que entre más flexible es una proteína, menor es su estabilidad cinética [75]. Sin embargo, esta afirmación está en concordancia con nuestros resultados, siempre y cuando se haga referencia a su flexibilidad estática a baja temperatura y no a su flexibilidad térmica. De hecho, en la figura 3.22 es evidente que a bajas temperaturas la flexibilidad estática ($D_{X,T}$) de TcTIM, que tienen una E_{act} muy alta, es menor que la de TbTIM. Esta situación se revierte a altas temperaturas. Por lo tanto, es interesante evaluar si la estabilidad cinética se correlaciona también con la flexibilidad estática a temperatura constante. Calculamos el valor de $D_{X,T}$ a 400 y 510 K para las quimeras y los graficamos contra los valores de E_{act} . Los coeficientes de correlación resultantes son 0.78 y 0.32, respectivamente. Estos valores son inferiores a la R de 0.96 obtenida de la gráfica mostrada en la figura 3.27 que se obtuvo con los valores de TGD . Por consiguiente, la flexibilidad que determina principalmente la estabilidad cinética de una proteína es la térmica y no la estática.

La estabilidad térmica de las proteínas es un parámetro que se usa regularmente como una medida de su estabilidad. Esta estabilidad térmica es cuantificada en DSC como la temperatura máxima de la transición (T_m) cuyos valores se encuentran en la tabla 3.2. Dado que el TGD es una medida de la variación de la flexibilidad estática al aumentar la temperatura, se puede pensar que la flexibilidad térmica podría estar correlacionada también con la estabilidad térmica. Sin embargo, este no es el caso como se demuestra en la figura 3.29. Este resultado no es una sorpresa dado que la estabilidad térmica y la estabilidad cinética son fenómenos con fundamentos moleculares distintos [104].

En conclusión la estabilidad cinética y la cooperatividad del desplegamiento están ligadas a una propiedad que hemos definido aquí como la flexibilidad térmica de una proteína. La figura 3.30 es una representación esquemática de un paisaje de energía térmica contra espacio conformacional que muestra la relación entre la flexibilidad térmica y la estabilidad cinética de TbTIM y TcTIM. A una temperatura dada (energía térmica), la flexibilidad estática es proporcional al tamaño del espacio conformacional accesible para la proteína (líneas punteadas horizontales). A medida que la temperatura aumenta, el espacio conformacional también aumenta pero de forma distinta para ambas proteínas debido a sus diferentes flexibilidades

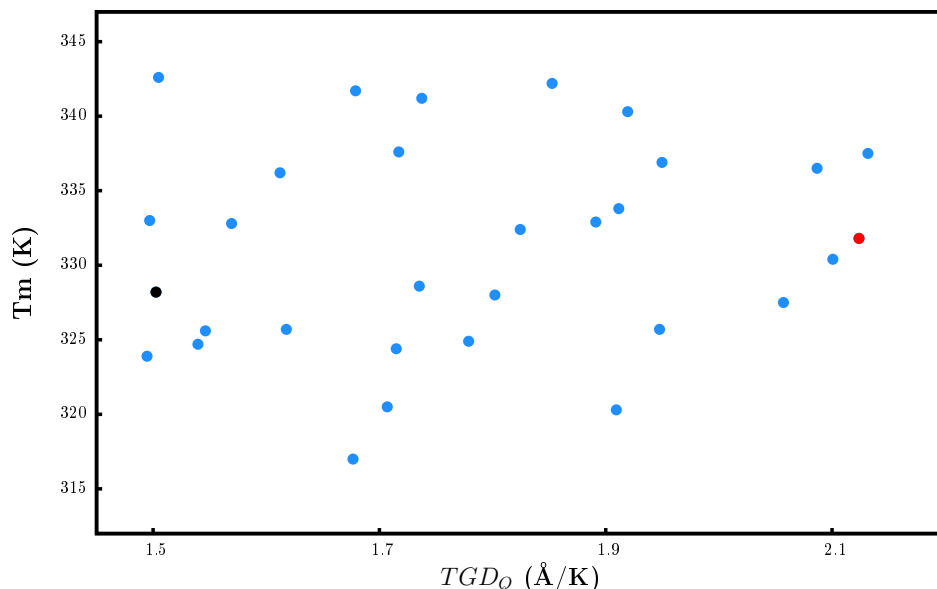


Figura 3.29: Valores de T_m en función de TGD . TbTIM (punto negro), TcTIM (punto rojo) y las quimeras (punto azul).

térmicas. La flexibilidad térmica de TbTIM (en $\text{\AA}/\text{K}$) es menor que la de TcTIM, por lo tanto la pendiente de la gráfica de la energía térmica (en K) contra el espacio conformacional (en \AA) será mayor. La barrera energética que separa el estado nativo del estado de transición es pequeña y se alcanza con facilidad debido a la pendiente pronunciada. El caso de TcTIM es el contrario, una flexibilidad térmica mayor se traduce en una pendiente menor en la gráfica de la energía térmica contra el espacio conformacional. En este caso la barrera es más grande y se alcanza con dificultad.

¿Porqué una proteína que se vuelve más flexible al aumentar la temperatura es cinéticamente más estable que otra que permanece más rígida?. La respuesta que surge de este trabajo es que una proteína con un estado de transición de alta energía (con valor alto de E_{act}) requiere de una alta flexibilidad para distribuir la energía que se le suministra en estados que posean un mayor grado de libertad. Habrá un punto en el que un conjunto de conformaciones la barrera energética del estado de transición y la proteína se desnaturalice irreversiblemente. Esto implica que el estado de transición tenga una conformación muy alejada a la estructura nativa. Pese a que en este trabajo no se identificó el conjunto estructural (o clúster) correspondiente al estado de transición, si se observa que al final de las simulaciones a las temperaturas más

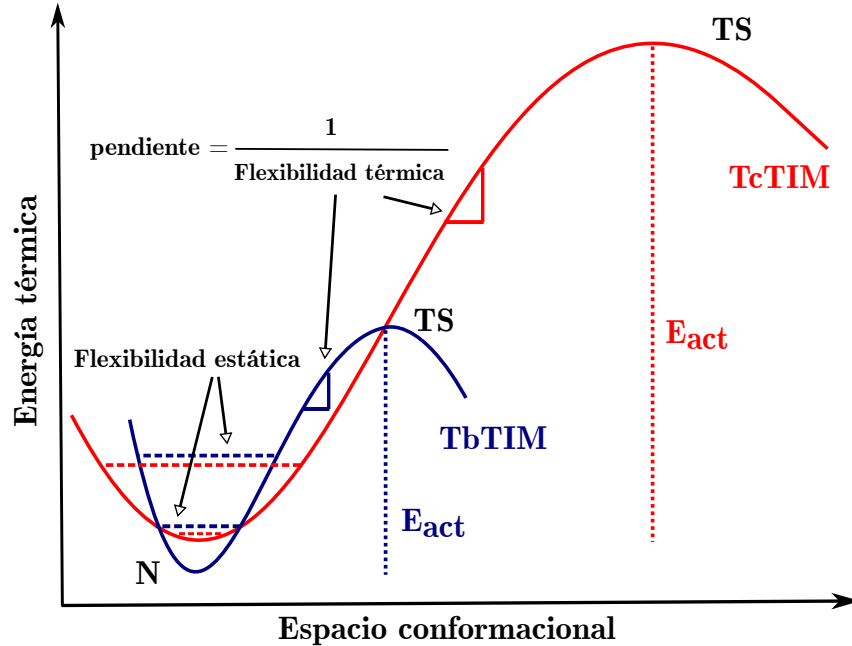


Figura 3.30: Representación esquemática de un paisaje de energía térmica (ET) contra el espacio conformacional (CS) que muestra los efectos de las flexibilidades estática y térmica sobre la estabilidad cinética (E_{act}) de TbTIM y TcTIM.

altas (490 y 510 K), TcTIM está visiblemente más desnaturalizada que TbTIM (figuras 3.11 y 3.17). Una proteína que a través de la selección natural reúna el conjunto de residuos adecuado para distribuir la energía en residuos que la hagan más flexible ante condiciones desnaturalizantes, sin perder funcionalidad, logrará incrementar la barrera energética que la separa de su estado desnaturalizado.

3.4 Contribución de cada residuo a la diferencia en flexibilidad térmica entre TbTIM y TcTIM

Con el propósito de discernir la contribución de cada residuo a los valores de TGD totales de las proteínas silvestres y al de las regiones, se llevó a cabo el análisis de desplazamiento por residuo en función de la temperatura (ecuaciones 2.10 y 2.7). La comparación de los valores de TGD para cada residuo (TGD_{res}) de TcTIM y TbTIM se muestra en la figura 3.31. En ella se desglosa la contribución individual que tienen los residuos al TGD de las regiones 4 y 5, que son las que se destacaron en la figura 3.24. Los residuos en estas regiones que son

visiblemente diferentes entre ambas proteínas se han enmarcado y se enlistan en la tabla 3.5.

Los residuos de la tabla 3.5 forman parte de las hélices α de ambas regiones. De los treinta y un residuos identificados, veinticuatro son iguales en ambas proteínas, seis son una diferencia conservativa y cinco una no-conservativa. En los once residuos distintos entre TbTIM y TcTIM se observa que no hay una predominancia en el tipo de sustitución, lo que indica que la flexibilidad térmica es más el resultado de un efecto concertado entre muchos residuos, que ocurre en un cierto contexto espacial, que de la identidad química de los residuos involucrados.

En la figura 3.31 también se observa que además de los que se encuentran en las regiones 4 y 5, hay otros residuos de TcTIM con valores de TGD mayores que el correspondiente residuo en TbTIM. Destacan los que se pueden observar en la región 3. Esta región también presenta un valor diferente de TGD_r para ambas proteínas, como se puede advertir en la figura 3.24. Sin embargo las barras de error asociadas a estos puntos se traslapan, por lo que no se contó como una diferencia más entre ambas enzimas. No obstante, con más simulaciones esas barras de error podrían ser más pequeñas y la región 3 también podría considerarse un dominio estructural clave para la diferencia en estabilidad cinética de ambas enzimas. En futuros estudios se podría explorar esta posibilidad.

Este escrutinio a nivel residuo ha dado como resultado una conexión con dos resultados publicados: (1) quince de los residuos identificados de la región 5 (del 141 al 155) están dentro de un fragmento de veintidós aminoácidos en TbTIM que se ha reportado como crítico para transportar una proteína reportera dentro del glicosoma de este parásito [33]. (2) Se ha encontrado que las regiones 1 y 4 de TcTIM juegan un papel importante en la susceptibilidad de esta proteína a un agente derivatizante de cisteínas [35]. La región 4 de TcTIM fue identificada en este trabajo como una de las regiones que tiene un alto valor de TDG (figura 3.24) y aunque los valores de TGD de la región 1 no son muy distintos entre TbTIM y TcTIM, la figura 3.31 muestra que hay tres residuos (Val22, Pro23 y Glu26) en esta región de TcTIM con un valor de TGD_{res} significativamente más alto que los correspondientes en TbTIM (Ser22, Glu23, Asp26). Dos de estos residuos (Val y Pro) se introdujeron en la quimera Tc(2,3,5-8)Tb(1,4) para producir una mutante de cinco residuos que muestra un

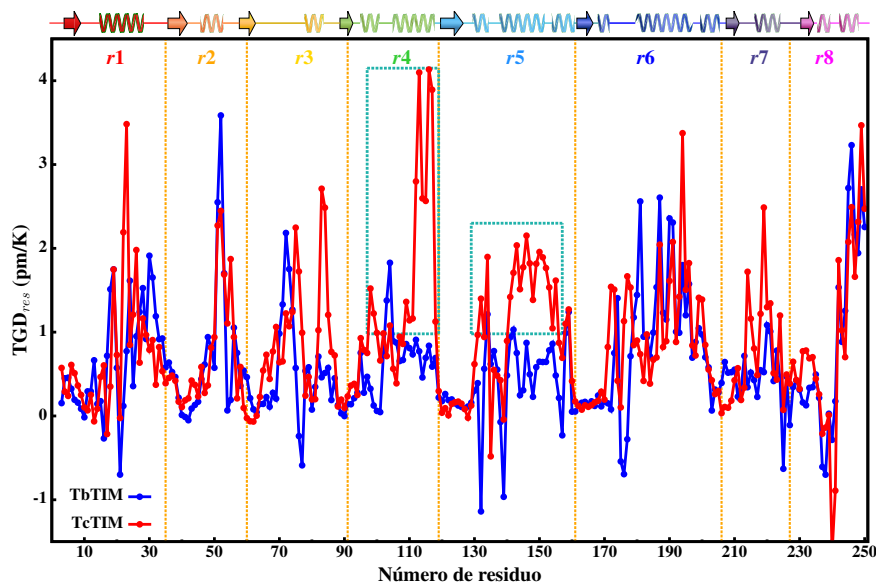


Figura 3.31: Valores de TGD_{res} para cada residuo de TbTIM y TcTIM. Las regiones están separadas por líneas punteadas. Se resaltan dentro de cajas color cian, los residuos con valores significativamente diferentes entre ambas enzimas silvestres.

decremento significativo en la susceptibilidad al agente derivatizante de cisteínas, comparada con la TcTIM silvestre [35]. Por lo tanto, parece que los residuos enumerados en la tabla 3.5 no solo parecen jugar un papel importante en la estabilidad cinética de TcTIM y TbTIM, sino en varias de las diferencias experimentales observadas hasta ahora entre estas enzimas evolutivamente cercanas.

Tabla 3.5: Residuos de TcTIM con valores de TGD_{res} significativamente mayores que TbTIM. Los residuos que son diferentes para ambas enzimas se muestran en rojo. Se indica el tipo de sustitución.

TcTIM	TbTIM	Sustitución	TcTIM	TbTIM	Sustitución
<i>region 4</i>			<i>region 5</i>		
99 ARG	98 ARG		132 ASN	131 LEU	Non-conservative
100 ARG	99 ARG		133 GLU	132 GLN	Non-conservative
101 LEU	100 ALA	Conservative	135 ARG	134 ARG	
110 VAL	109 VAL		142 ALA	141 VAL	Conservative
111 ALA	110 ALA		143 VAL	142 VAL	
112 GLU	111 ASP	Conservative	144 VAL	143 VAL	
113 LYS	112 LYS		145 LEU	144 LEU	
114 VAL	113 VAL		146 THR	145 THR	
115 ALA	114 ALA		147 GLN	146 GLN	
116 GLN	115 ALA	Non-conservative	148 LEU	147 ILE	Conservative
117 ALA	116 ALA		149 ALA	148 ALA	
118 CYS	117 VAL	Conservative	150 ALA	149 ALA	
119 ALA	118 ALA		151 VAL	150 ILE	Conservative
			152 ALA	151 ALA	
			153 GLN	152 LYS	Non-conservative
			154 LYS	153 LYS	
			155 LEU	154 LEU	
			156 SER	155 LYS	Non-conservative

Conclusiones

La flexibilidad es un concepto muy usado para explicar muchos aspectos del comportamiento de las proteínas. En general, se refiere a la habilidad de la estructura de una proteína de visitar diferentes conformaciones. Pero, ¿qué debemos entender cuando se afirma que una proteína es más o menos flexible que otra? En este trabajo, proponemos un nuevo concepto de flexibilidad de proteínas, al que hemos llamado flexibilidad térmica. Este concepto hace referencia al desplazamiento de los residuos debido a un gradiente de temperatura. Como tal, la flexibilidad térmica mide cuantitativamente la habilidad de la estructura de una proteína de visitar diferentes conformaciones cuando se le provee de energía.

Usando como sistema modelo la triosafosfato isomerasa de dos parásitos evolutivamente cercanos, se ha demostrado que hay una correlación entre la estabilidad cinética y la cooperatividad del desplegamiento, determinadas a través de experimentos de calorimetría diferencial de barrido, con la flexibilidad térmica obtenida por un método novedoso para calcular el desplazamiento de residuos en simulaciones de dinámica molecular a alta temperatura. Entre mayor estabilidad cinética tenga una proteína, su flexibilidad térmica será mayor.

Esta correlación contribuye al conocimiento de las bases moleculares de la estabilidad cinética de las proteínas. La estrategia aquí descrita puede ser de gran ayuda en el diseño de proteínas con estabilidades cinéticas altas, lo que es un objetivo importante desde un punto de vista biotecnológico. Desde el punto de vista evolutivo, contribuye a dar luz al entendimiento del delicado balance entre la estructura, la función y el tiempo de vida media de una proteína en tiempos relevantes dentro de la célula.

Perspectivas

Considerando los resultados más importantes reportados en este trabajo, se hacen las siguientes sugerencias para un trabajo futuro:

- Realizar el cálculo de TGD de otras proteínas de estabilidad cinética conocida, con la finalidad de observar si la relación entre la flexibilidad térmica y la estabilidad cinética es una característica general de las proteínas.
- Calcular los valores de TGD utilizando todos los átomos de la proteína (excepto los hidrógenos) para obtener una cuantificación más precisa de su flexibilidad y así reducir la dispersión que se observa en la correlación entre la flexibilidad térmica y la estabilidad cinética (figura 3.27).
- Llevar a cabo simulaciones de TbTIM y TcTIM a otras temperaturas para descartar o confirmar la relevancia de la región 3 en la diferencia en estabilidad cinética entre TbTIM y TcTIM (ver figura 3.24).
- Repetir las simulaciones de dinámica molecular de alta temperatura con diferentes modelos de agua para tratar de caracterizar las interacciones que se forman y rompen entre el solvente y la proteína con el objetivo de establecer una relación entre las barreras de solvatación y la flexibilidad térmica.
- Diseñar y medir la estabilidad cinética de la proteína conformada por los residuos que tienen un valor mayor de TGD en la figura 3.31 y la proteína conformada por los residuos que tienen un valor menor de TGD en dicha figura.

- Diseñar y medir la estabilidad cinética de una proteína que sea mayoritariamente TbTIM pero con los residuos de TcTIM que aparecen en la tabla 3.5 y de la proteína inversa, es decir, la proteína que sea mayoritariamente TcTIM pero con los residuos de TbTIM que se enlistan en dicha tabla.
- Obtener la estructura cristalina de algunas quimeras y cuantificar su flexibilidad térmica mediante MDS de alta temperatura, para evaluar que tan correcta es la aproximación de aditividad para calcular los valores de TGD de las quimeras que se describió en este trabajo.

Bibliografía

- [1] Y. AGUIRRE, N. CABRERA, B. AGUIRRE, R. PÉREZ-MONTFORT, A. HERNANDEZ-SANTOYO, H. REYES-VIVAS, S. ENRÍQUEZ-FLORES, M. DE GÓMEZ-PUYOU, A. GÓMEZ-PUYOU, J. M. SANCHEZ-RUIZ, AND M. COSTAS, *Different contribution of conserved amino acids to the global properties of triosephosphate isomerases*, *Proteins*, 82 (2014), pp. 323–335.
- [2] T. ALBER, D. BANNER, A. C. BLOOMER, G. A. PETSKO, D. PHILLIPS, R. SEC, P. S. RIVERS, AND I. A. WILSON, *On the three-dimensional structure and catalytic mechanism of triosephosphate isomerase*, *Phil. Trans. R. Soc. Lond. B*, 293 (1981), pp. 159–171.
- [3] M. ALVAREZ, J. P. ZEELLEN, V. MAINFROID, F. RENTIER-DELRUE, J. A. MARTIAL, L. WYNS, R. K. WIERENGA, AND D. MAES, *Triose-phosphate Isomerase (TIM) of the Psychrophilic Bacterium Vibrio marinus: Kinetic and Structural Properties*, *Journal of Biological Chemistry*, 273 (1998), pp. 2199–2206.
- [4] P. A.M. MICHELS, F. BRINGAUD, M. HERMAN, AND V. HANNAERT, *Metabolic functions of glycosomes in trypanosomatids*, *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1763 (2006), pp. 1463–1477.
- [5] C. B. ANFINSEN AND E. HABER, *ARTICLE : Studies on the Reduction and Re-formation of Protein Disulfide Bonds Studies on the Reduction and Re-formation of Protein Disulfide Bonds*, *The Journal of biological chemistry*, 236 (1961), pp. 1361–1363.

- [6] B. M. BAKKER, F. I. C. MENSORIDES, B. TEUSINK, P. V. HOEK, P. A. M. MICHELS, AND H. V. WESTERHOFF, *Compartmentation protects trypanosomes from the dangerous design of glycolysis*, 97 (2000), pp. 2087–2092.
- [7] M. P. BARRETT, R. J. S. BURCHMORE, A. STICH, J. O. LAZZARI, A. C. FRASCH, AND J. J. CAZZULO, *The trypanosomiases*, 362 (2003), pp. 1469–1480.
- [8] N. BEAUCAMP, A. HOFMANN, B. KELLERER, AND R. JAENICKE, *Dissection of the gene of the bifunctional PGK-TIM fusion protein from the hyperthermophilic bacterium Thermotoga maritima: design and characterization of the separate triosephosphate isomerase.*, Protein science : a publication of the Protein Society, 6 (1997), pp. 2159–65.
- [9] N. C. BENSON AND V. DAGGETT, *Dynameomics: large-scale assessment of native protein flexibility.*, Protein science : a publication of the Protein Society, 17 (2008), pp. 2038–50.
- [10] H. J. C. BERENDSEN, J. P. M. POSTMA, W. F. VAN GUNSTEREN, AND J. HERMANS, *Interaction models for water in relation to protein hydration*, 1981.
- [11] H. J. C. BERENDSEN, D. VAN DER SPOEL, AND R. VAN DRUNEN, *GROMACS: A message-passing parallel molecular dynamics implementation*, Computer Physics Communications, 91 (1995), pp. 43–56.
- [12] C. A. BIPPES, A. ZELTINA, F. CASAGRANDE, M. RATERA, M. PALACIN, D. J. MULLER, AND D. FOTIADIS, *Substrate binding tunes conformational flexibility and kinetic stability of an amino acid antiporter*, Journal of Biological Chemistry, 284 (2009), pp. 18651–18663.
- [13] G. BRUYLANTS, J. WOUTERS, AND C. MICHAUX, *Differential scanning calorimetry in life science: thermodynamics, stability, molecular recognition and application in drug design.*, Current medicinal chemistry, 12 (2005), pp. 2011–2020.
- [14] S. CANSU AND P. DORUKER, *Dimerization affects collective dynamics of triosephosphate isomerase.*, Biochemistry, 47 (2008), pp. 1358–68.

- [15] M. E. CHÁNEZ-CÁRDENAS, G. PÉREZ-HERNÁNDEZ, B. G. SÁNCHEZ-REBOLLAR, M. COSTAS, AND E. VÁZQUEZ-CONTRERAS, *Reversible equilibrium unfolding of triosephosphate isomerase from Trypanosoma cruzi in guanidinium hydrochloride involves stable dimeric and monomeric intermediates*, *Biochemistry*, 44 (2005), pp. 10883–10892.
- [16] M. H. CHIU AND E. J. PRENNER, *Differential scanning calorimetry: An invaluable tool for a detailed thermodynamic characterization of macromolecules and their interactions.*, *Journal of Pharmacy and Bioallied Sciences*, 3 (2011), pp. 39–59.
- [17] C.-H. CHU, Y.-J. LAI, H. HUANG, AND Y.-J. SUN, *Kinetic and structural properties of triosephosphate isomerase from Helicobacter pylori*, *Proteins: Structure, Function, and Bioinformatics*, 71 (2008), pp. 396–406.
- [18] S. CONNELLY, S. CHOI, S. M. JOHNSON, J. W. KELLY, AND I. A. WILSON, *Structure-based design of kinetic stabilizers that ameliorate the transthyretin amyloidoses*, *Current Opinion in Structural Biology*, 20 (2010), pp. 54–62.
- [19] M. COSTAS, D. RODRÍGUEZ-LARREA, L. DE MARIA, T. V. BORCHERT, A. GÓMEZ-PUYOU, AND J. M. SANCHEZ-RUIZ, *Between-Species Variation in the Kinetic Stability of TIM Proteins Linked to Solvation-Barrier Free Energies*, *Journal of Molecular Biology*, 385 (2009), pp. 924–937.
- [20] E. L. CUNNINGHAM, S. S. JASWAL, J. L. SOHL, AND D. A. AGARD, *Kinetic stability as a mechanism for protease longevity*, *Proc. Natl. Acad. Sci. U. S. A.*, 96 (1999), pp. 11008–11014.
- [21] V. DAGGETT, *Molecular Dynamics Simulations of the Protein Unfolding / Folding Reaction*, *Accounts of Chemical Research*, 35 (2002), pp. 422–429.
- [22] T. DARDEN, D. YORK, AND L. PEDERSEN, *Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems*, *The Journal of Chemical Physics*, 98 (1993), pp. 10089–10092.

- [23] D. R. DEE, B. MYERS, AND R. Y. YADA, *Dynamics of thermodynamically stable, kinetically trapped, and inhibitor-bound states of pepsin*, *Biophysical Journal*, 101 (2011), pp. 1699–1709.
- [24] M. A. DEPRISTO, D. M. WEINREICH, AND D. L. HARTL, *Missense meanderings in sequence space: a biophysical view of protein evolution*, *Nature Reviews Genetics*, 6 (2005), pp. 678–687.
- [25] N. DÍAZ-VERGARA AND A. PIÑEIRO, *Molecular dynamics study of triosephosphate isomerase from Trypanosoma cruzi in water/decane mixtures.*, *The journal of physical chemistry. B*, 112 (2008), pp. 3529–3539.
- [26] A. Y. DUNBAR, Y. KAMADA, G. J. JENKINS, E. R. LOWE, AND S. S. BILLECKE, *Ubiquitination and Degradation of Neuronal Nitric-Oxide Synthase in Vitro : Dimer Stabilization Protects the Enzyme from Proteolysis*, *Molecular Pharmacology*, 66 (2004), pp. 964–969.
- [27] U. ESSMANN, L. PERERA, M. L. BERKOWITZ, T. DARDEN, H. LEE, AND L. G. PEDERSEN, *A smooth particle mesh Ewald method*, *The Journal of Chemical Physics*, 103 (1995), pp. 8577–8593.
- [28] J. A. FARIÁS-RICO, S. SCHMIDT, AND B. HÖCKER, *Evolutionary relationship of two ancient protein superfolds*, *Nature Chemical Biology*, 10 (2014), pp. 710–715.
- [29] E. E. FIGUEROA-ANGULO, P. ESTRELLA-HERNÁNDEZ, H. SALGADO-LUGO, A. OCHOA-LEYVA, A. GÓMEZ PUYOU, S. S. CAMPOS, G. MONTERO-MORAN, J. ORTEGA-LÓPEZ, G. SAAB-RINCÓN, R. ARROYO, C. G. BENÍTEZ-CARDOZA, AND L. G. BRIEBA, *Cellular and biochemical characterization of two closely related triosephosphate isomerases from Trichomonas vaginalis*, *Parasitology*, 139 (2012), pp. 1729–1738.
- [30] J. FITTER AND J. HEBERLE, *Structural Equilibrium Fluctuations in Mesophilic and Thermophilic α -Amylase*, *Biophysical Journal*, 79 (2000), pp. 1629–1636.

- [31] P. FORRER, C. CHANG, D. OTT, A. WLODAWER, AND A. PLÜCKTHUN, *Kinetic stability and crystal structure of the viral capsid protein SHP*, Journal of Molecular Biology, 344 (2004), pp. 179–193.
- [32] T. R. FOSS, M. S. KELKER, R. L. WISEMAN, I. A. WILSON, AND J. W. KELLY, *Kinetic Stabilization of the Native State by Protein Engineering : Implications for Inhibition of Transthyretin Amyloidogenesis*, (2005), pp. 841–854.
- [33] N. GALLAND, S. DE WALQUE, F. G. VONCKEN, C. L. VERLINDE, AND P. A.M. MICHELS, *An internal sequence targets Trypanosoma brucei triosephosphate isomerase to glycosomes*, Molecular and Biochemical Parasitology, 171 (2010), pp. 45–49.
- [34] I. GARCÍA-TORRES, *Quimeras de la Triosafosfato Isomerasa: transformación progresiva de la triosafosfato isomerasa de Trypanosoma cruzi en la triosafosfato isomerasa de Trypanosoma brucei.*, PhD thesis, Universidad Nacional Autónoma de México, 2012.
- [35] I. GARCÍA-TORRES, N. CABRERA, A. TORRES-LARIOS, M. RODRÍGUEZ-BOLAÑOS, S. DÍAZ-MAZARIEGOS, A. GÓMEZ-PUYOU, AND R. PEREZ-MONTFORT, *Identification of Amino Acids that Account for Long-Range Interactions in Two Triosephosphate Isomerases from Pathogenic Trypanosomes*, PLoS ONE, 6 (2011), p. e18791.
- [36] V. GUZMAN-LUNA AND G. GARZA-RAMOS, *The folding pathway of glycosomal triosephosphate isomerase: Structural insights into equilibrium intermediates*, Proteins: Structure, Function, and Bioinformatics, 80 (2012), pp. 1669–1682.
- [37] P. HAMMARSTRÖM, R. L. WISEMAN, E. T. POWERS, AND J. W. KELLY, *Prevention of Transthyretin Amyloid Disease by Changing Protein Misfolding Energetics*, 713 (2009).
- [38] G. HERNANDEZ, F. E. JENNEY, M. W. ADAMS, AND D. M. LEMASTER, *Milli-second time scale conformational flexibility in a hyperthermophile protein at ambient temperature.*, Proceedings of the National Academy of Sciences of the United States of America, 97 (2000), pp. 3166–3170.

- [39] B. HESS, H. BEKKER, H. J. C. BERENDSEN, AND J. G. E. M. FRAAIJE, *LINCS: A linear constraint solver for molecular simulations*, J. Comput. Chem., 18 (1997), pp. 1463–1472.
- [40] B. HESS, C. KUTZNER, D. VAN DER SPOEL, AND E. LINDAHL, *GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation*, J. Chem. Theory Comput., 4 (2008), pp. 435–447.
- [41] B. HÖCKER, *Engineering chimaeric proteins from fold fragments: ‘hopeful monsters’ in protein design*, Biochemical Society Transactions, 41 (2013), pp. 1137–1140.
- [42] B. HÖCKER, S. BEISMANN-DRIEMEYER, S. HETTWER, A. LUSTIG, AND R. STERNER, *Dissection of a (betaalpha)8-barrel enzyme into two folded halves.*, Nature structural biology, 8 (2001), pp. 32–36.
- [43] B. HÖCKER, J. CLAREN, AND R. STERNER, *Mimicking enzyme evolution by generating new (betaalpha)8-barrels from (betaalpha)4-half-barrels.*, Proceedings of the National Academy of Sciences of the United States of America, 101 (2004), pp. 16448–53.
- [44] B. HÖCKER, C. JÜRGENS, M. WILMANN, AND R. STERNER, *Stability, catalytic versatility and evolution of the (β α)8-barrel fold*, Current Opinion in Biotechnology, 12 (2001), pp. 376–381.
- [45] B. HÖCKER, S. SCHMIDT, AND R. STERNER, *A common evolutionary origin of two elementary enzyme folds*, FEBS Letters, 510 (2002), pp. 133–135.
- [46] R. W. HOCKNEY AND J. W. EASTWOOD, *Computer Simulation Using Particles*, Adam Hilger, 1988.
- [47] P.-S. HUANG, K. FELDMEIER, F. PARMEGGIANI, D. A. FERNANDEZ VELASCO, B. HÖCKER, AND D. BAKER, *De novo design of a four-fold symmetric TIM-barrel protein with atomic-level accuracy*, Nature Chemical Biology, 12 (2016), pp. 29–34.
- [48] T. ISOMERASE, F. MASSI, C. WANG, AND A. G. PALMER, *Solution NMR and computer simulation studies of active site loop motion in triosephosphate isomerase.*, Biochemistry, 45 (2006), pp. 10787–94.

- [49] M. JAMROZ, A. KOLINSKI, AND S. KMIECIK, *CABS-flex: Server for fast simulation of protein structure fluctuations.*, Nucleic acids research, 41 (2013), pp. W427–W431.
- [50] R. E. JEFFERSON, T. M. BLOIS, AND J. U. BOWIE, *Membrane proteins can have high kinetic stability*, Journal of the American Chemical Society, 135 (2013), pp. 15183–15190.
- [51] L. JIMÉNEZ, D. A. FERNÁNDEZ-VELASCO, K. WILLMS, AND A. LANDA, *A comparative study of biochemical and immunological properties of triosephosphate isomerase from Taenia Solium and Sus Scrofa*, Journal of Parasitology, 89 (2003), pp. 209–214.
- [52] L. JIMÉNEZ, N. VIBANCO-PÉREZ, L.ÑAVARRO, AND A. LANDA, *Cloning, expression and characterisation of a recombinant triosephosphate isomerase from Taenia solium*, International Journal for Parasitology, 30 (2000), pp. 1007–1012.
- [53] F. JOUBERT, A. W. H. NEITZ, AND A. I. LOUW, *Structure-based inhibitor screening: a family of sulfonated dye inhibitors for malaria parasite triosephosphate isomerase.*, Proteins: Structure, Function, and Genetics, 45 (2001), pp. 136–143.
- [54] T. J. KAMERZELL AND C. R. MIDDAUGH, *The complex Inter-relationships between protein flexibility and stability*, Journal of pharmaceutical sciences, 97 (2008), pp. 3494–3517.
- [55] A. KARSHIKOFF, L.ÑILSSON, AND R. LADENSTEIN, *Rigidity versus flexibility: The dilemma of understanding protein thermal stability*, FEBS Journal, 282 (2015), pp. 3899–3917.
- [56] D. KNOBELOCH, A. SCHMIDT, P. SCHEERER, N. KRAUSS, H. WESSNER, C. SCHOLZ, G. KÜTTNER, T. VON RINTELEN, A. WESSEL, AND W. HÖHNE, *A coleopteran triosephosphate isomerase: X-ray structure and phylogenetic impact of insect sequences.*, Insect molecular biology, 19 (2010), pp. 35–48.
- [57] T. P. J. KNOWLES, M. VENDRUSCOLO, AND C. M. DOBSON, *The amyloid state and its association with protein misfolding diseases.*, Nature reviews. Molecular cell biology, 15 (2014), pp. 384–96.

- [58] L. KOHL, M. CALLENS, R. K. WIERENGA, F. R. OPPERDOES, AND P. A. M. MICHELS, *Triosephosphate isomerase of Leishmania mexicana mexicana. Cloning characterization of the gene, overexpression in Escherichia coli and analysis of the protein*, European Journal of Biochemistry, 220 (1994), pp. 331–338.
- [59] W. K. KRIETSCH, P. G. PENTCHEV, H. KLINGENBÜRG, T. HOFSTÄTTER, AND T. BÜCHER, *The isolation and crystallization of yeast and rabbit liver triose phosphate isomerase and a comparative characterization with the rabbit muscle enzyme.*, European journal of biochemistry / FEBS, 14 (1970), pp. 289–300.
- [60] K. KUMAR, P. BHARGAVA, AND U. ROY, *Cloning, overexpression and characterization of Leishmania donovani triosephosphate isomerase*, Experimental Parasitology, 130 (2012), pp. 430–436.
- [61] O. KURKCUOGLU, R. L. JERNIGAN, AND P. DORUKER, *Loop motions of triosephosphate isomerase observed with elastic networks.*, Biochemistry, 45 (2006), pp. 1173–82.
- [62] Z. KURKCUOGLU, A. BAKAN, D. KOCAMAN, I. BAHAR, AND P. DORUKER, *Coupling between Catalytic Loop Motions and Enzyme Global Dynamics*, PLoS Computational Biology, 8 (2012), p. e1002705.
- [63] Z. KURKCUOGLU AND P. DORUKER, *Substrate Effect on Catalytic Loop and Global Dynamics of Triosephosphate Isomerase*, Entropy, 15 (2013), p. 1085.
- [64] Z. KURKCUOGLU, D. FINDIK, E. AKTEN, AND P. DORUKER, *How an Inhibitor Bound to Subunit Interface Alters Triosephosphate Isomerase Dynamics*, Biophysical Journal, 109 (2015), pp. 1169–1178.
- [65] A.-M. LAMBEIR, F. R. OPPERDOES, AND R. K. WIERENGA, *Kinetic properties of triosephosphate isomerase. A comparison with the rabbit muscle and yeast enzymes*, European Journal of Biochemistry, FEBS, 168 (1987), pp. 69–74.
- [66] A. LANDA, A. ROJO-DOMÍNGUEZ, L. JIMÉNEZ, AND D. A. FERNÁNDEZ-VELASCO, *Sequencing, expression and properties of triosephosphate isomerase from Entamoeba histolytica.*, European journal of biochemistry / FEBS, 247 (1997), pp. 348–355.

- [67] E. LINDAHL, B. HESS, AND D. VAN DER SPOEL, *GROMACS 3.0: a package for molecular simulation and trajectory analysis*, Molecular modeling annual, 7 (2001), pp. 306–317—.
- [68] Z. LIU AND H. S. CHAN, *Solvation and desolvation effects in protein folding: native flexibility, kinetic cooperativity and enthalpic barriers under isostability conditions.*, Physical biology, 2 (2005), pp. S75–S85.
- [69] G. LÓPEZ-VELÁZQUEZ, D. MOLINA-ORTIZ, N. CABRERA, G. HERNÁNDEZ-ALCÁNTARA, J. PEON-PERALTA, L. YÉPEZ-MULIA, R. PÉREZ-MONTFORT, AND H. REYES-VIVAS, *An unusual triosephosphate isomerase from the early divergent eukaryote Giardia lamblia.*, Proteins: Structure, Function, and Bioinformatics, 55 (2004), pp. 824–34.
- [70] S. M. LYNCH, S. A. BOSWELL, AND W. COLÓN, *Kinetic stability of Cu/Zn superoxide dismutase is dependent on its metal ligands: Implications for ALS*, Biochemistry, 43 (2004), pp. 16525–16531.
- [71] M. MACHIUS, N. DECLERCK, R. HUBER, AND G. WIEGAND, *Kinetic stabilization of Bacillus licheniformis α -amylase through introduction of hydrophobic residues at the surface*, Journal of Biological Chemistry, 278 (2003), pp. 11546–11553.
- [72] V. MAINFROID, S. C. MANDE, W. G. HOL, J. A. MARTIAL, AND K. GORAJ, *Stabilization of human triosephosphate isomerase by improvement of the stability of individual alpha-helices in dimeric as well as monomeric forms of the protein.*, Biochemistry, 35 (1996), pp. 4110–7.
- [73] V. MAINFROID, P. TERPSTRA, M. BEAUREGARD, J. M. FRÈRE, S. C. MANDE, W. G. HOL, J. A. MARTIAL, AND K. GORAJ, *Three hTIM mutants that provide new insights on why TIM is a dimer.*, Journal of molecular biology, 257 (1996), pp. 441–456.
- [74] S. C. MANDE, V. MAINFROID, K. H. KALK, K. GORAJ, J. A. MARTIAL, AND W. G. HOL, *Crystal structure of recombinant human triosephosphate isomerase at 2.8 Å resolution. Triosephosphate isomerase-related human genetic disorders and comparison*

- with the trypanosomal enzyme.*, Protein science : a publication of the Protein Society, 3 (1994), pp. 810–21.
- [75] M. MANNING AND W. COLÓN, *Structural basis of protein kinetic stability: Resistance to sodium dodecyl sulfate suggests a central role for rigidity and a bias toward β -sheet structure*, Biochemistry, 43 (2004), pp. 11248–11254.
- [76] J. MANSFELD, G. VRIEND, W. BAUKE, O. R. VELTMAN, B. VAN DEN, G. VENEMA, V. G. H. EIJSINK, B. W. DIJKSTRA, B. V. D. BURG, AND R. ULBRICH-HOFMANN, *Protein Chemistry and Structure : Extreme Stabilization of a Thermolysin-like Protease by an Engineered Disulfide Bond Extreme Stabilization of a Thermolysin-like Protease by an Engineered Disulfide Bond **, The Journal of biological chemistry, 272 (1997), pp. 11152–11156.
- [77] D. MATHUR, G. MALIK, AND L. C. GARG, *Biochemical and functional characterization of triosephosphate isomerase from Mycobacterium tuberculosis H37Rv.*, FEMS microbiology letters, 263 (2006), pp. 229–35.
- [78] S. MIYAMOTO AND P. A. KOLLMAN, *Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models*, J. Comput. Chem., 13 (1992), pp. 952–962.
- [79] E. MONSELLIER AND F. CHITI, *Prevention of amyloid-like aggregation as a driving force of protein evolution.*, EMBO reports, 8 (2007), pp. 737–42.
- [80] J. MORAES, R. ARREOLA, N. CABRERA, L. SARAMAGO, D. FREITAS, A. MASUDA, I. DA SILVA VAZ, M. TUENA DE GOMEZ-PUYOU, R. PEREZ-MONTFORT, A. GOMEZ-PUYOU, AND C. LOGULLO, *Structural and biochemical characterization of a recombinant triosephosphate isomerase from Rhipicephalus (Boophilus) microplus.*, Insect biochemistry and molecular biology, 41 (2011), pp. 400–9.
- [81] H. ÑÁJERA, M. COSTAS, AND D. A. FERNÁNDEZ-VELASCO, *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 (2003), pp. 785–792.

- [82] Z. NEVIN GEREK, S. KUMAR, AND S. BANU OZKAN, *Structural dynamics flexibility informs function and evolution at a proteome scale*, *Evolutionary Applications*, 6 (2013), pp. 423–433.
- [83] S. ÑOSE, *A molecular dynamics method for simulations in the canonical ensemble*, *Molecular Physics*, 52 (1984), pp. 255–268.
- [84] C. OOSTENBRINK, A. VILLA, A. E. MARK, AND W. F. VAN GUNSTEREN, *A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6*, *J. Comput. Chem.*, 25 (2004), pp. 1656–1676.
- [85] F. R. OPPERDOES, *Glycosomes may provide clues to the import of peroxisomal proteins*, *Trends in Biochemical Sciences*, 13 (1988), pp. 255–260.
- [86] F. R. OPPERDOES AND P. A. M. MICHELS, *The glycosomes of the Kinetoplastida*, (1993), pp. 231–234.
- [87] M. PARRINELLO AND A. RAHMAN, *Polymorphic transitions in single crystals: A new molecular dynamics method*, *Journal of Applied Physics*, 52 (1981), pp. 7182–7190.
- [88] A. B. PATEL, S. ALLEN, M. C. DAVIES, C. J. ROBERTS, S. J. B. TENDLER, AND P. M. WILLIAMS, *Influence of Architecture on the Kinetic Stability of Molecular Assemblies*, *Journal of the American Chemical Society*, 126 (2004), pp. 1318–1319.
- [89] H. M. PETRASSI, S. M. JOHNSON, H. E. PURKEY, K. P. CHIANG, T. WALKUP, X. JIANG, E. T. POWERS, AND J. W. KELLY, *Potent and Selective Structure-Based Dibenzofuran Inhibitors of Transthyretin Amyloidogenesis : Kinetic Stabilization of the Native State*, (2005), pp. 6662–6671.
- [90] B. PLAUT AND J. R. KNOWLES, *pH-dependence of the triose phosphate isomerase reaction.*, *The Biochemical journal*, 129 (1972), pp. 311–20.
- [91] I. M. PLAZA DEL PINO, B. IBARRA-MOLERO, AND J. M. SANCHEZ-RUIZ, *Lower kinetic limit to protein thermal stability: A proposal regarding protein stability in vivo*

- and its relation with misfolding diseases*, Proteins: Structure, Function and Genetics, 40 (2000), pp. 58–70.
- [92] I. POZDNYAKOVA, J. GUIDRY, AND P. WITTUNG-STAFSHEDE, *Copper stabilizes azurin by decreasing the unfolding rate.*, Archives of biochemistry and biophysics, 390 (2001), pp. 146–148.
- [93] S. J. PUTMAN, A. F. COULSON, I. R. FARLEY, B. RIDDLESTON, AND J. R. KNOWLES, *Specificity and kinetics of triose phosphate isomerase from chicken muscle.*, The Biochemical journal, 129 (1972), pp. 301–310.
- [94] H. REYES-VIVAS, 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, *Disulfide bridges in the mesophilic triosephosphate isomerase from Giardia lamblia are related to oligomerization and activity.*, Journal of molecular biology, 365 (2007), pp. 752–63.
- [95] H. REYES-VIVAS, E. MARTINEZ-MARTINEZ, G. MENDOZA-HERNANDEZ, G. LOPEZ-VELAZQUEZ, R. PEREZ-MONTFORT, M. TUENA DE GOMEZ-PUYOU, AND A. GOMEZ-PUYOU, *Susceptibility to proteolysis of triosephosphate isomerase from two pathogenic parasites: Characterization of an enzyme with an intact and a nicked monomer*, Proteins: Structure, Function, and Genetics, 48 (2002), pp. 580–590.
- [96] A. W. M. RIETVELD AND S. T. FERREIRA, *Kinetics and energetics of subunit dissociation/unfolding of TIM: The importance of oligomerization for conformational persistence and chemical stability of proteins*, Biochemistry, 37 (1998), pp. 933–937.
- [97] D. RODRIGUEZ-LARREA, S. MINNING, T. V. BORCHERT, AND J. M. SANCHEZ-RUIZ, *Role of Solvation Barriers in Protein Kinetic Stability*, Journal of Molecular Biology, 360 (2006), pp. 715–724.
- [98] S. ROMERO-ROMERO, *Control cinético y termodinámico en el desplegamiento térmico de la Triosafosfato isomerasa proveniente de especies bacterianas*, tesis de maestría, Universidad Nacional Autónoma de México, 2013.

- [99] S. ROMERO-ROMERO, M. COSTAS, A. RODRIGUEZ-ROMERO, AND D. A. FERNANDEZ-VELASCO, *Reversibility and two state behaviour in the thermal unfolding of oligomeric TIM barrel proteins*, *Phys. Chem. Chem. Phys.*, 17 (2015), pp. 20699–20714.
- [100] G. SAAB-RINCÓN, L. OLVERA, M. OLVERA, E. RUDIÑO-PIÑERA, E. BENITES, X. SOBERÓN, AND E. MORETT, *Evolutionary Walk between $(\beta/\alpha)_8$ Barrels: Catalytic Migration from Triosephosphate Isomerase to Thiamin Phosphate Synthase*, *Journal of Molecular Biology*, 416 (2012), pp. 255–270.
- [101] N. L. SALIMI, B. HO, AND D. A. AGARD, *Unfolding Simulations Reveal the Mechanism of Extreme Unfolding Cooperativity in the Kinetically Stable α -Lytic Protease*, *PLoS Computational Biology*, 6 (2010), p. e1000689.
- [102] I. SANCHEZ-ROMERO, A. ARIZA, K. S. WILSON, M. SKJØT, J. VIND, L. DE MARIA, L. K. SKOV, AND J. M. SANCHEZ-RUIZ, *Mechanism of Protein Kinetic Stabilization by Engineered Disulfide Crosslinks*, *PLoS ONE*, 8 (2013), p. e70013.
- [103] J. M. SANCHEZ-RUIZ, *Theoretical analysis of Lumry-Eyring models in differential scanning calorimetry.*, *Biophysical journal*, 61 (1992), pp. 921–935.
- [104] J. M. SANCHEZ-RUIZ, *Protein kinetic stability*, *Biophysical Chemistry*, 148 (2010), pp. 1–15.
- [105] S. SELVARAJ AND M. M. GROMIHA, *An analysis of the amino acid clustering pattern in $(\alpha/\beta)_8$ barrel proteins.*, *Journal of protein chemistry*, 17 (1998), pp. 407–15.
- [106] J. A. SILVERMAN, R. BALAKRISHNAN, AND P. B. HARBURY, *Reverse engineering the $(\beta/\alpha)_8$ barrel fold.*, *Proceedings of the National Academy of Sciences of the United States of America*, 98 (2001), pp. 3092–3097.
- [107] M. A. SMITH, P. A. ROMERO, T. WU, E. M. BRUSTAD, AND F. H. ARNOLD, *Chimeragenesis of distantly-related proteins by noncontiguous recombination*, *Protein Science*, 22 (2013), pp. 231–238.
- [108] R. STERNER AND B. HÖCKER, *Catalytic versatility, stability, and evolution of the $(\beta\alpha)_8$ -barrel enzyme fold*, *Chemical Reviews*, 105 (2005), pp. 4038–4055.

- [109] W. G. STREET, F. PARAK, AND R. D. YOUNG, *Conformational substates in proteins*, (1988).
- [110] W. SUN, S. LIU, P. J. BRINDLEY, AND D. P. MCMANUS, *Bacterial expression and characterization of functional recombinant triosephosphate isomerase from Schistosoma japonicum.*, Protein expression and purification, 17 (1999), pp. 410–413.
- [111] K. E. TANG AND K. A. DILL, *Native protein fluctuations: the conformational-motion temperature and the inverse correlation of protein flexibility with protein stability.*, Journal of biomolecular structure & dynamics, 16 (1998), pp. 397–411.
- [112] M. L. TEODORO, G.Ñ. PHILLIPS, AND L. E. KAVRAKI, *Understanding protein flexibility through dimensionality reduction.*, Journal of Computational Biology, 10 (2003), pp. 617–34.
- [113] D. R. TRENTHAM, C. H. McMURRAY, AND C. I. POGSON, *The active chemical state of D-glyceraldehyde 3-phosphate in its reactions with D-glyceraldehyde 3-phosphate dehydrogenase, aldolase and triose phosphate isomerase.*, The Biochemical journal, 114 (1969), pp. 19–24.
- [114] M. TUENA DE GOMEZ-PUYOU AND A. GOMEZ-PUYOU, *Enzymes in low water systems.*, Critical reviews in biochemistry and molecular biology, 33 (1998), pp. 53–89.
- [115] H. WALDEN, G. L. TAYLOR, E. LORENTZEN, E. POHL, H. LILIE, A. SCHRAMM, T. KNURA, K. STUBBE, B. TJADEN, AND R. HENSEL, *Structure and function of a regulated archaeal triosephosphate isomerase adapted to high temperature.*, Journal of Molecular Biology, 342 (2004), pp. 861–875.
- [116] R. K. WIERENGA, *The TIM-barrel fold: A versatile framework for efficient enzymes*, FEBS Letters, 492 (2001), pp. 193–198.
- [117] R. K. WIERENGA, E. G. KAPETANIOU, AND R. VENKATESAN, *Triosephosphate isomerase : a highly evolved biocatalyst*, Cellular and Molecular Life Sciences, 67 (2010), pp. 3961–3982.

- [118] R. L. WISEMAN, S. M. JOHNSON, M. S. KELKER, T. FOSS, I. A. WILSON, AND J. W. KELLY, *Kinetic stabilization of an oligomeric protein by a single ligand binding event*, Journal of the American Chemical Society, 127 (2005), pp. 5540–5551.
- [119] K. XIA, M. MANNING, H. HESHAM, Q. LIN, C. BYSTROFF, AND W. COLÓN, *Identifying the subproteome of kinetically stable proteins via diagonal 2D SDS/PAGE.*, Proceedings of the National Academy of Sciences of the United States of America, 104 (2007), pp. 17329–34.
- [120] Y. XIE, J. AN, G. YANG, G. WU, Y. ZHANG, L. CUI, AND Y. FENG, *Enhanced enzyme kinetic stability by increasing rigidity within the active site*, Journal of Biological Chemistry, 289 (2014), pp. 7994–8006.

Apéndices

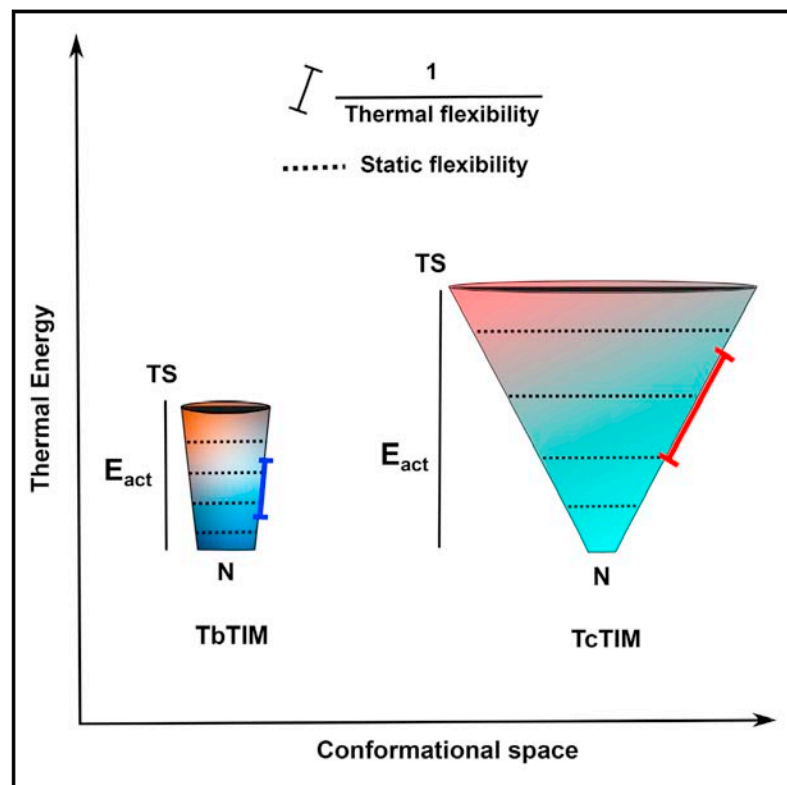
Apéndice A

Artículo publicado

Structure

Interplay between Protein Thermal Flexibility and Kinetic Stability

Graphical Abstract



Authors

Andrea G. Quezada,
A. Jessica Díaz-Salazar,
Nallely Cabrera, Ruy Pérez-Montfort,
Ángel Piñeiro, Miguel Costas

Correspondence

angel.pineiro@usc.es (Á.P.),
costasmi@unam.mx (M.C.)

In Brief

Quezada et al. propose a quantitative measure of protein flexibility using a residue displacement analysis of temperature-induced unfolding molecular dynamics simulations. Experimentally unfolding activation energies and cooperativities correlate with thermal flexibility. This correlation increases our knowledge on the molecular basis of protein kinetic stability.

Highlights

- Kinetic stability and unfolding cooperativity are linked to thermal flexibility
- Thermal flexibility is given by residue displacements due to a temperature gradient
- There are concerted effects among residues within a certain global spatial context
- Protein kinetic stability can be enhanced modifying its thermal flexibility



Interplay between Protein Thermal Flexibility and Kinetic Stability

Andrea G. Quezada,¹ A. Jessica Díaz-Salazar,¹ Nallely Cabrera,² Ruy Pérez-Montfort,² Ángel Piñeiro,^{3,*} and Miguel Costas^{1,4,*}

¹Laboratorio de Biofísicoquímica, Departamento de Físicoquímica, Facultad de Química, Universidad Nacional Autónoma de México, México City 04510, México

²Departamento de Bioquímica y Biología Estructural, Instituto de Fisiología Celular, Universidad Nacional Autónoma de México, Mexico City 04510, México

³Soft Matter and Molecular Biophysics Group, Department of Applied Physics, University of Santiago de Compostela, Santiago de Compostela 15782, Spain

⁴Lead Contact

*Correspondence: angel.pineiro@usc.es (Á.P.), costasmi@unam.mx (M.C.)

<http://dx.doi.org/10.1016/j.str.2016.11.018>

SUMMARY

Kinetic stability is a key parameter to comprehend protein behavior and it plays a central role to understand how evolution has reached the balance between function and stability in cell-relevant timescales. Using an approach that includes simulations, protein engineering, and calorimetry, we show that there is a clear correlation between kinetic stability determined by differential scanning calorimetry and protein thermal flexibility obtained from a novel method based on temperature-induced unfolding molecular dynamics simulations. Thermal flexibility quantitatively measures the increment of the conformational space available to the protein when energy is provided. The $(\beta/\alpha)_8$ barrel fold of two closely related by evolution triosephosphate isomerases from two trypanosomes are used as model systems. The kinetic stability-thermal flexibility correlation has predictive power for the studied proteins, suggesting that the strategy and methodology discussed here might be applied to other proteins in biotechnological developments, evolutionary studies, and the design of protein based therapeutics.

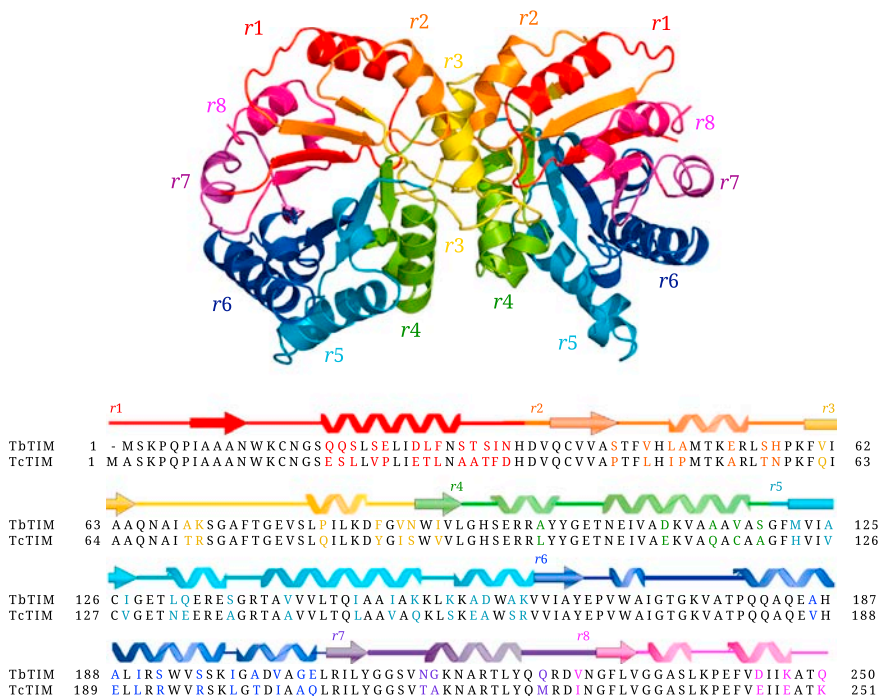
INTRODUCTION

It is generally accepted that there is a relationship between the flexibility, stability, and catalytic activity of enzymes: the higher the flexibility, the higher the catalytic activity and the lower the stability (Kamerzell and Middaugh, 2008; Tuena de Gómez-Puyou and Gómez-Puyou, 1998). The concept of stability might be interpreted as thermodynamic or kinetic. The former refers to the free energy change for the unfolding process, which is positive at a given temperature. The latter refers to the time dependence of the irreversible process of protein denaturation, which is related to an activation energy barrier between the native state and the non-functional forms (unfolded states, irreversible dena-

tured protein) (Sanchez-Ruiz, 2010). The concept of protein kinetic stability has emerged, not only to fulfill a global view of protein stability, but also to understand how evolution has reached the balance between function and stability in cell-relevant timescales as well as to predict and enhance the half-life of proteins with biotechnological applications (Jefferson et al., 2013; Manning and Colón, 2004; Patel et al., 2004; Sanchez-Ruiz, 2010; Xia et al., 2007). The disruption of the function/stability balance in the crowded cell medium may lead to situations (aggregation, proteolysis, amyloidogenesis, etc.) that can result in pathological conditions. In fact, it has been suggested that preventing aggregation in the cell is an important driving force for protein evolution (Monsellier and Chiti, 2007). The experimental quantitative measure of kinetic stability is the activation energy that can be obtained using differential scanning calorimetry (DSC) and has been interpreted in terms of desolvation/solvation barriers in protein-folding/unfolding processes (Costas et al., 2009; Rodriguez-Larrea et al., 2006). These barriers reduce protein flexibility and increase folding/unfolding cooperativity (Cheung et al., 2002; Liu and Chan, 2005). Protein flexibility has been evaluated in several forms (Benson and Daggett, 2008; Jamroz et al., 2013; Ma, 2005; Teodoro et al., 2003), and conceptualized in two broad definitions (Tang and Dill, 1998): static flexibility refers to the number and structural diversity of fluctuation conformations in the equilibrium ensemble, while dynamic flexibility refers to how quickly the protein can hop from one conformation to another. From these definitions, a correlation between dynamic flexibility and kinetic stability might be expected. Such correlation has also been inferred from single-molecule dynamic force microscopy unfolding experiments using proteoliposomes deposited on mica (Bippes et al., 2009).

Triosephosphate isomerase (TIM) is a dimeric enzyme with a $(\beta/\alpha)_8$ barrel fold. This fold is present in five out of the six types of reactions known for enzymes (Höcker, 2013). As a result of its versatility, it has been widely used in protein engineering research (Höcker et al., 2004; Huang et al., 2016). Construction of chimeras of the TIM barrel fold has helped answer questions in the area of protein science (Höcker, 2013; Smith et al., 2013). Of importance for the present work are the TIMs of the evolutionary closely related parasites, *Trypanosoma brucei* (TbTIM) and *Trypanosoma cruzi* (TcTIM). These two enzymes have 74% sequence identity and their structures superpose





with a root-mean-square deviation (RMSD) value less than 1 Å. Despite this, they show strikingly different behaviors in several aspects: susceptibility to a sulfhydryl reagent (García-Torres et al., 2011), sensitivity to proteolysis (Reyes-Vivas et al., 2002), chemically induced unfolding pathways (Chávez-Cárdenas et al., 2005; Guzman-Luna and Garza-Ramos, 2012), and kinetic stability (the activation energy value for TcTIM is twice that of TbTIM) (Costas et al., 2009). Given that natural selection has shaped these two enzymes with highly different energetic barriers for irreversible denaturation, without losing catalytic efficiency, and with only a few changes in sequence and structure, they are excellent model systems to increase our knowledge of the molecular basis of kinetic stability and its relation to flexibility. To reach this goal, we used a synergistic approach (Das et al., 2013; Ladumer et al., 1998) combining protein engineering, DSC and molecular dynamics (MD) simulations.

RESULTS

Based on its $(\beta/\alpha)_8$ barrel fold, TIM can be divided into eight interchangeable regions, each containing a β sheet, a β - α loop and an α helix (Figure 1). Using the eight regions, it is possible to construct chimeras where one or several regions of one protein are grafted into the other protein. Here, we studied 29 chimeras. They have been named as Tc(R₁)Tb(R₂), where Tc and Tb are short forms for TcTIM and TbTIM and R₁ and R₂ indicate the regions of each protein that form the chimera (see Experimental Procedures). The 29 TcTIM/TbTIM chimeras showed TIM activity (Table 1) in the GAP to DHAP direction. Their catalytic efficiencies (cat/K_M) are within the $\mu \pm \sigma$ interval (μ being the mean and σ the SD) estimated using data for 23 wild-type TIMs reported in the literature (Table S1 and Figure S1). In consequence, although no structural data are available for the majority of the

Figure 1. Separation of the TIM Structure in Eight Regions

The TIM barrel was divided into eight equivalent regions (r1 to r8), each comprising an α helix, a loop, and a β strand. The dimer structure for TbTIM is shown. The number of residues are: 35, 25, 31, 28, 42, 45, 21, and 23 for regions r1 to r8, respectively. The aligned sequences of TbTIM and TcTIM and the secondary structure are displayed. Residues that are identical for both proteins are in black, while colored ones indicate the differences between them. There are 13, 7, 8, 5, 15, 10, 3, and 4 such differences in regions r1 to r8, respectively. This figure was generated with PyMOL.

chimeral enzymes (except for Tc(2,3,5-8)Tb(1,4) (PDB: 3Q37) (García-Torres et al., 2011), it is concluded that the chimeras studied here are correctly folded and catalytically competent proteins.

Thermal Denaturation, Activation Energies, and Kinetic Stabilities

For the TIM chimeras and the wild-type enzymes, scanning-rate-dependent

DSC was used to characterize the kinetics of the denaturation and to determine the activation energy values. For the wild-type TcTIM and TbTIM proteins these parameters have been reported previously and discussed in detail (Costas et al., 2009). The DSC profiles of all the chimeras showed that their thermal denaturation is irreversible, i.e., no signal was observed after cooling and reheating the samples, and scan-rate dependent. Figures 2A and 2B show examples of DSC profiles for two chimeras. These results imply that the thermal denaturation of all the chimeras is under kinetic control. As shown in Figures 2A and 2B, the profiles are well fitted by the two-state irreversible model (see Experimental Procedures). The compliance of the model with the calorimetric profiles obtained experimentally indicates that the kinetically relevant transition is dimeric, as was the case for wild-type TcTIM and TbTIM (Costas et al., 2009) and the E104D mutants of several TIMs (Aguirre et al., 2014). The activation energies (E_{act}) were obtained (see Experimental Procedures) from the slopes of the Arrhenius plots (Figures 2C and 2D). For the other 27 chimeras, the fitted calorimetric profiles and the corresponding Arrhenius plots are given in the Supplemental Information (Figures S2–S5). The E_{act} values for all chimeras are shown in Table 1, together with the temperatures corresponding to the maximum of each transition (T_m) and the widths of the calorimetric transitions at half the maximum height (W). W is a direct experimental measure of the cooperativity of the unfolding transition: the smaller the W value (sharp transition) the greater the cooperativity and the larger the E_{act} value.

The E_{act} values in Table 1 provide information about the relative importance of some of the regions upon the kinetic stability of the enzymes. Three examples are: (1) the chimera Tc(5)Tb(1-4,6-8) has an $E_{act} = 600 \pm 52$ kJ/mol, which is 40% higher than that for wild-type TbTIM ($E_{act} = 434 \pm 26$ kJ/mol), indicating that region 5 of TcTIM might play an important role in the high kinetic stability of TcTIM ($E_{act} = 822 \pm 30$ kJ/mol),

Table 1. Parameters Derived from DSC and Catalytic Activity Experiments^a

Protein	E_{act} (kJ/mol)	W (K)	T_m (K) ^b	$K_{cat}/K_M \times 10^{-8}$ (M/s) ^c
Wild-type				
TbTIM	434 ± 26	5.2	328.2	1.24
TcTIM	822 ± 30	2.8	331.8	1.28
One region of TcTIM grafted into TbTIM				
Tc(1)Tb(2–8)	361 ± 4 ^d	5.7	323.9	2.75
Tc(2)Tb(1,3–8)	510 ± 21	4.8	342.6	0.93
Tc(3)Tb(1,2,4–8)	419 ± 22	5.2	325.7	ND ^h
Tc(4)Tb(1–3,5–8)	609 ± 35	3.6	324.4	2.7 ^g
Tc(5)Tb(1–4,6–8)	600 ± 52	3.6	328.6	0.75
Tc(6)Tb(1–5,7–8)	470 ± 47 ^e	6.4	325.6	1.4
One region of TbTIM grafted into TcTIM				
Tc(2–8)Tb(1)	838 ± 51	2.4	337.5	ND
Tc(1–3,5–8)Tb(4)	789 ± 31	3.0	333.8	2.46 ^g
Tc(1–4,6–8)Tb(5)	639 ± 37	3.2	332.9	0.68
Two regions of TcTIM grafted into TbTIM				
Tc(2,4)Tb(1,3,5–8)	493 ± 15	4.8	337.6	1.05
Tc(1–2)Tb(3–8)	493 ± 21	4.5	333.0	2.40 ^g
Tc(4,5)Tb(1–3,6–8)	626 ± 22	3.6	325.7	1.1
Tc(2,5)Tb(1–3,6–8)	459 ± 15	4.8	341.2	0.96
Tc(5,6)Tb(1–4,7–8)	526 ± 23	4.0	324.9	1.28
Tc(1,7)Tb(2–6,8)	373 ± 5	6.0	324.7	0.7
Tc(1,4)Tb(2,3,5–8)	377 ± 0 ^f	4.8	320.5	ND
Two regions of TbTIM grafted into TcTIM				
Tc(1,3,5–8)Tb(2,4)	620 ± 62 ^e	6.8	320.3	ND
Tc(2–6,8)Tb(1,7)	822 ± 53	2.8	336.5	1.24
Tc(2,3,5–8)Tb(1,4)	472 ± 13	4.9	340.3	5.75 ^g
Tc(1–3,6–8)Tb(4,5)	456 ± 7	5.2	341.7	1.58
Tc(1–6)Tb(7,8)	846 ± 29	2.6	330.4	3.08 ^g
Three regions of TcTIM grafted into TbTIM				
Tc(1–3)Tb(4–8)	526 ± 23	4.7	336.2	2.38 ^g
Tc(2,4,5)Tb(1,3,6–8)	659 ± 24	3.6	336.9	0.99
Tc(2,3,5)Tb(1,4,6–8)	433 ± 1	5.2	342.2	ND
Tc(6–8)Tb(1–5)	558 ± 22 ^f	4.0	327.5	ND
Three regions of TbTIM grafted into TcTIM				
Tc(1–5)Tb(6–8)	953 ± 27	2.2	332.8	1.70 ^g
Tc(1,3,6–8)Tb(2,4,5)	426 ± 17	5.2	317.0	0.91
Four regions of TcTIM and four regions of TbTIM				
Tc(1–4)Tb(5–8)	639 ± 16	3.2	332.4	2.20 ^g
Tc(5–8)Tb(1–4)	706 ± 38 ^f	3.2	328.0	ND

Activation energy (E_{act}), temperature corresponding to the maximum of the transition (T_m), width of the calorimetric transition at half the maximum height (W), derived from DSC experiments and catalytic efficiency for wild-type and chimerical proteins.

^aThe range of scan rates employed was 0.5 to 3 K/min, unless otherwise stated.

^bValues correspond to scan rate of 1.5 K/min.

^c K_M corrected to consider that there is only 4% of free GAP in solution.

^dScan rates ranged from 0.5 to 2.5 K/min.

^eFrom a single experiment at a scan rate of 1.5 K/min.

^fScan rates ranged from 0.5 to 1.5 K/min.

^gFrom [García-Torres et al. \(2011\)](#).

^hND, not determined.

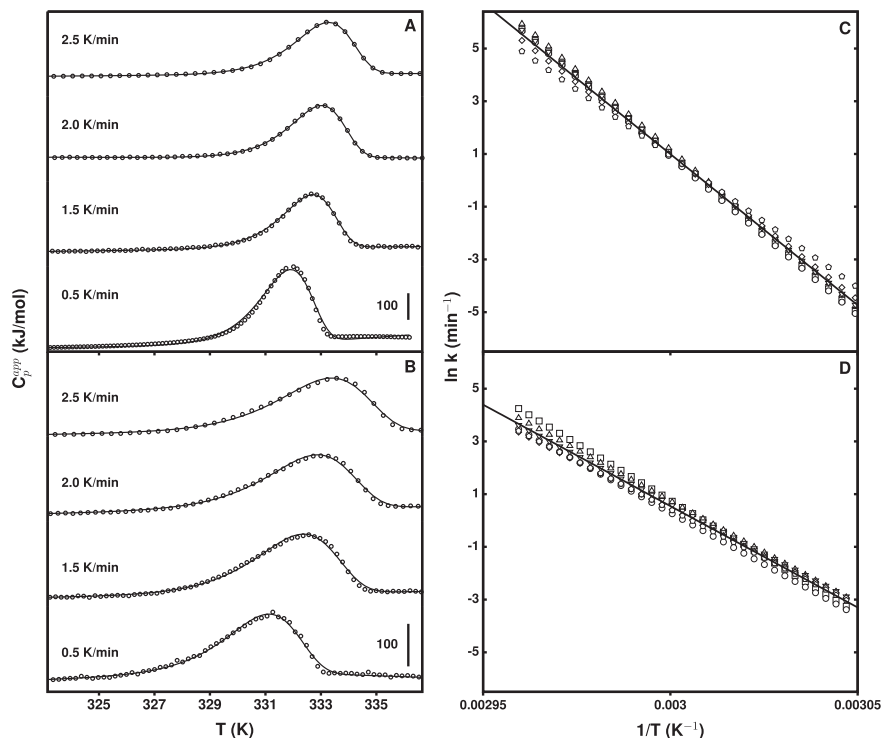


Figure 2. Examples of Experimental DSC Thermograms

Results for two chimeral proteins are shown: (A) Tc(1–5)Tb(6–8) and (B) Tc(1–4)Tb(5–8) at the indicated scan rates and at 0.4 mg/mL. Symbols represent the apparent heat capacity and continuous lines represent the fitting to the two-state irreversible model. The profiles have been shifted in the y axis for clarity. Arrhenius plots for the same chimeral proteins are shown in (C) and (D), respectively. The symbols correspond to the different scan rates employed: 0.5, 1, 1.5, 2.0, 2.5, and 3.0 K/min. Straight lines are the best fit to all the points. The values for T_m , widths of the calorimetric transitions at half the maximum height (W) and E_{act} derived from (C) and (D) for all the chimeral proteins are given in Table 1.

MD Simulations of Wild-type TbTIM and TcTIM

The thermal unfolding of the wild-type TcTIM and TbTIM was studied with 24-ns-long MD simulations at 400, 450, 490, and 510 K and with a 7-ns-long simulation at 700 K (see [Experimental Procedures](#)). The temperatures employed in the

(2) the E_{act} for the chimera Tc(1–3,6–8)Tb(4,5) is very close (456 ± 7 kJ/mol) to that for wild-type TbTIM, signaling that regions 4 and 5 of TbTIM might be crucial for the low kinetic stability of wild-type TbTIM, and (3) the E_{act} values for the chimeras Tc(1–4)Tb(5–8) and Tc(5–8)Tb(1–4), i.e., those where the dimer interface (constituted by residues that are located within regions 1 to 4) was swapped, are close to each other (639 ± 16 and 706 ± 38 kJ/mol, respectively), but halfway between the values for the parental wild-type TIMs. This shows that the interface has an important contribution to the kinetic stability but it is not uniquely responsible for either the low or the high kinetic stability of TbTIM and TcTIM, respectively. Despite this and other similar valuable inferences from the data in Table 1 it appears that, irrespective of the number of regions or their identity, exchanging regions between the wild-type TcTIM to TbTIM enzymes does not produce a smooth and systematic change in the E_{act} values for the chimeric proteins. In fact, there are chimeras with higher and lower kinetic stability than those of wild-type TcTIM and TbTIM, respectively: for Tc(1–5)Tb(6–8) $E_{act} = 953 \pm 27$ kJ/mol and for Tc(1)Tb(2–8) $E_{act} = 361 \pm 4$ kJ/mol. Clearly, the non-smooth and non-systematic variation of E_{act} with the exchanged regions (which extends to W and T_m) is not the result of the fact that the 29 chimeral TIMs studied here are only a fraction of the total possible number of chimeras that can be constructed in this way. In general, from the experimental results in Table 1 it is concluded that the kinetic stability of TIM barrel proteins results from a complex network of short (local, within a region) and long (global, between regions) range interactions in concerted action. To gain insight into the possible molecular level determinants of the kinetic stability and unfolding cooperativity, MD simulations at several temperatures were performed.

MD simulations are artificially high in order to speed up the protein-unfolding process. Therefore, it is assumed that the structure of the proteins in the simulations at temperatures between 400 and 510 K, for which the unfolding is incomplete, corresponds to structures at much lower temperatures. In the simulations, as the temperature increases the unfolding becomes more significant and is faster (Figure 3A). At 700 K the perturbation is much more serious and takes place in a much shorter timescale, the secondary structure vanishing in less than 7 ns. The monomers remained together in both proteins during all simulations, and the radius of gyration does not exhibit important changes along the trajectories (Figure S6). Although we do not expect to determine the unfolding pathway in our simulations, it is worth mentioning that the joint unfolding of both subunits, without separation into monomers, is consistent with the fitting of the calorimetric data to a model that does not include a dimer dissociation step (see [Experimental Procedures](#)).

The RMSD of the TbTIM $C\alpha$ atoms is independent and larger than the RMSD values of the corresponding subunits at all temperatures, while for TcTIM this only happens in the simulations at 510 and 700 K (Figure S7). This could be attributed to a decoupling of both subunits in the protein structure or simply to the relative movement between them, as for instance the counter-rotation of subunits observed previously in chicken TIM and in TcTIM (Cansu and Doruker, 2008; Kurkcuglu and Doruker, 2013). The eight α helices surrounding the β barrel core of the TIM structure are easily disrupted with increasing temperature (Figure 3). In contrast, the central part of the barrel formed by eight parallel β sheets was much more resistant at all temperatures (they were observed to unfold only at 700 K) for both proteins and subunits (Figures 3 and 4). This indicates that TIM barrels denature with temperature starting from their envelope, or external part (α helices), and continue with their core, or internal part (β sheets).

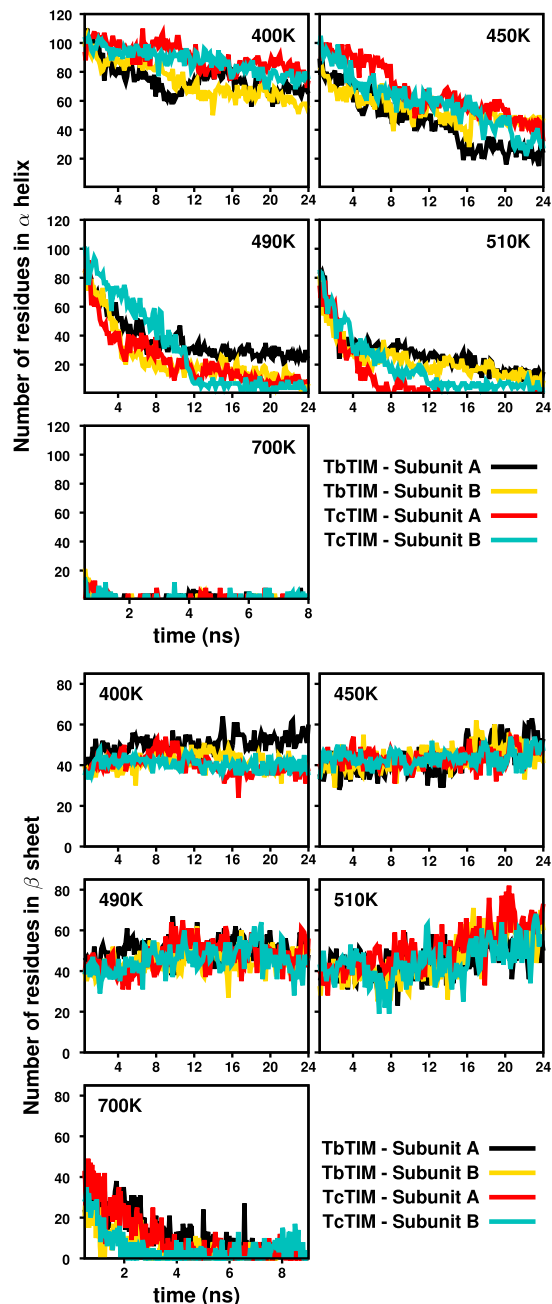


Figure 3. Secondary Structure

The number of residues forming part of the α helices (top panel) and beta strands (bottom panel) along the trajectories for subunits A and B of TbTIM and TcTIM at the temperatures indicated in the plots.

At 400 and 450 K the loss of α -helical structure progresses with time at a constant rate throughout all the simulations (Figure 3), with no significant differences between TcTIM and TbTIM. In contrast, at 490 and 510 K the loss of secondary structure is more evident at shorter simulation times and some differences appear between the structural evolution of both species. The α helices in TcTIM are completely unfolded at 490 and 510 K after ~ 12 and ~ 8 ns, respectively, while the TbTIM α helices seem to be slightly more resistant (Figure 3). At these tempera-

tures, the loss of α -helical structure in TbTIM is comparable with that in TcTIM during the first 4 ns of the simulations, but then it continues at a lower rate. The differences observed in the unfolding pattern of both proteins might be related to a more cooperative process for TcTIM than for TbTIM, in agreement with the experimental W values (Table 1).

Conformational Clustering and Residue Displacement Analysis

The projections on the 3D C α -RMSD space of the trajectories for the dimer and for its subunits were employed to calculate (see Experimental Procedures) the progressive moving dispersion (PMD) (Figures 4B, 4C, S8, and S9). This analysis was performed at every temperature except at 700 K, at which the unfolding process was too fast. Clusters were defined when simultaneous minima in the PMD were observed for the dimer and the two subunits. The total number of clusters identified in this manner ranged from 11 to 16 for all temperatures and both enzymes.

At each temperature, residue displacement values ($\Delta d_{j,c}$) obtained from the last cluster of each trajectory provide a quantitative expression of the position change of every residue with respect to the rest of the protein (see Experimental Procedures). Typical $\Delta d_{j,c}$ values are of a few Å, negative when the residue (on average) gets closer to the rest of the protein and positive in the opposite case; as expected, the $\Delta d_{j,c}$ values increase in magnitude from 400 to 510 K for both proteins (Figure 5). However, at the residue level it is difficult to find clear tendencies when comparing both proteins and their subunits. Nonetheless, when the residues are grouped into the eight regions depicted in Figure 1, the corresponding region displacements $D_{r,T}$ (Equation 6 in Experimental Procedures) reveal the global movements that the protein undergoes upon temperature change; as the trajectory advances and the temperature increases, $D_{r,T}$ rises moderately for TbTIM and more intensely for TcTIM (Figure 6A). The differences between TbTIM and TcTIM are more clearly reflected in the protein total displacement $D_{wt,T}$ at each temperature (Figure 6B). At the lowest temperature (400 K) $D_{wt,T}$ for TbTIM is larger than for TcTIM, while at the highest temperature (510 K) the situation is reversed. The variation of $D_{wt,T}$ with temperature, i.e., the temperature gradient of the total displacement (TGD), was estimated using a linear fitting for both wild-type proteins: 2.12 ± 0.19 Å/K (0.0086 ± 0.0008 Å/K per residue) and 1.50 ± 0.10 Å/K (0.0061 ± 0.0004 Å/K per residue) for TcTIM and TbTIM, respectively. The decomposition of the TGDs for each protein region (Figure 6C) reveals clear differences between TcTIM and TbTIM. Regions 4 and 5 are seen in Figure 6 to have the largest differences between the TGD values for TcTIM and TbTIM. The comparison between both proteins (Figure 1) shows that the number of residues which are different are 5 for region 4, and 15 for region 5. Therefore, TGD values reflect concerted effects among residues within a certain global spatial context.

DISCUSSION

Protein Thermal Flexibility, Kinetic Stability, and Cooperativity of the Unfolding Transition

Static flexibility may be qualitatively understood as the ability of a protein to change its structure by visiting different

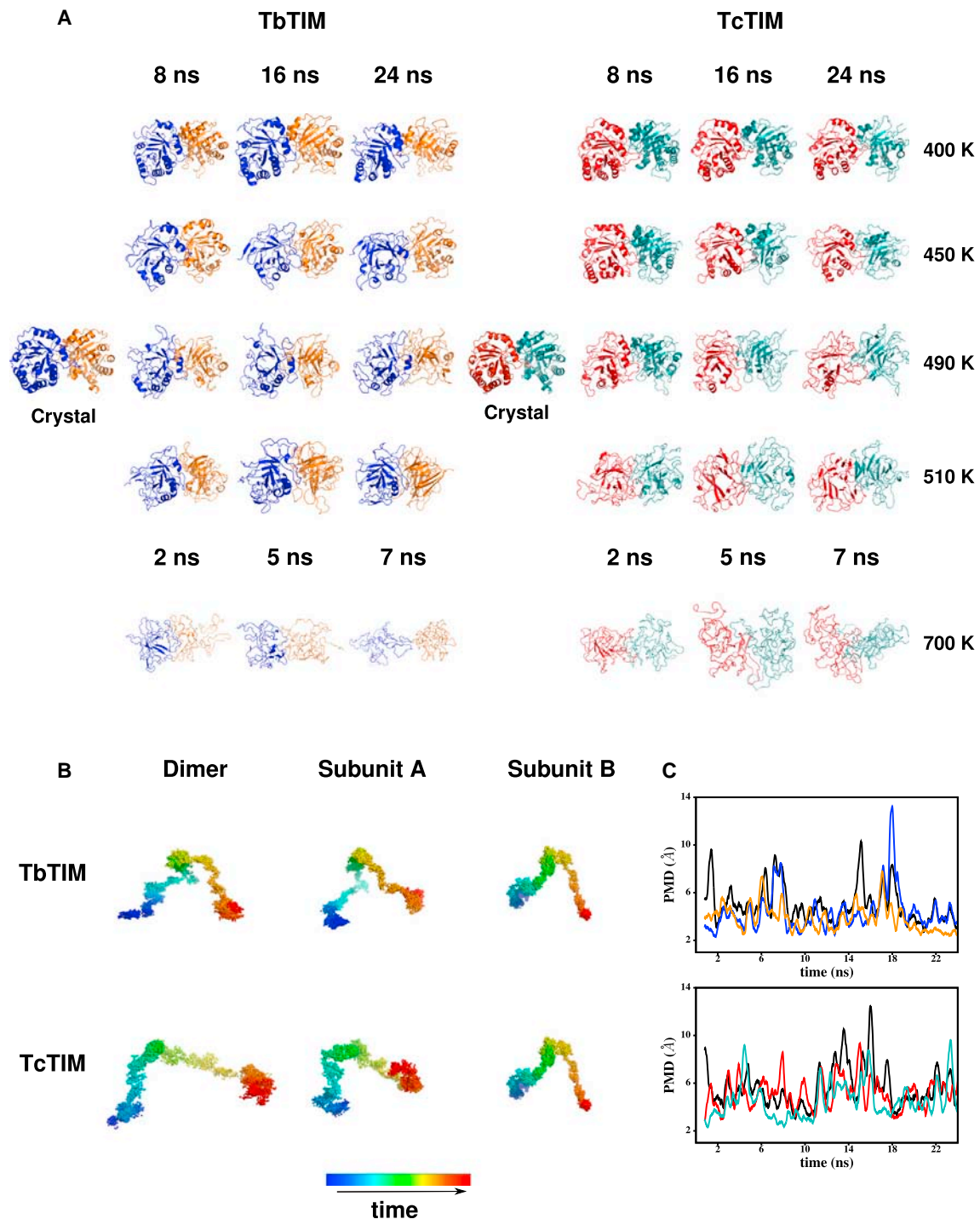
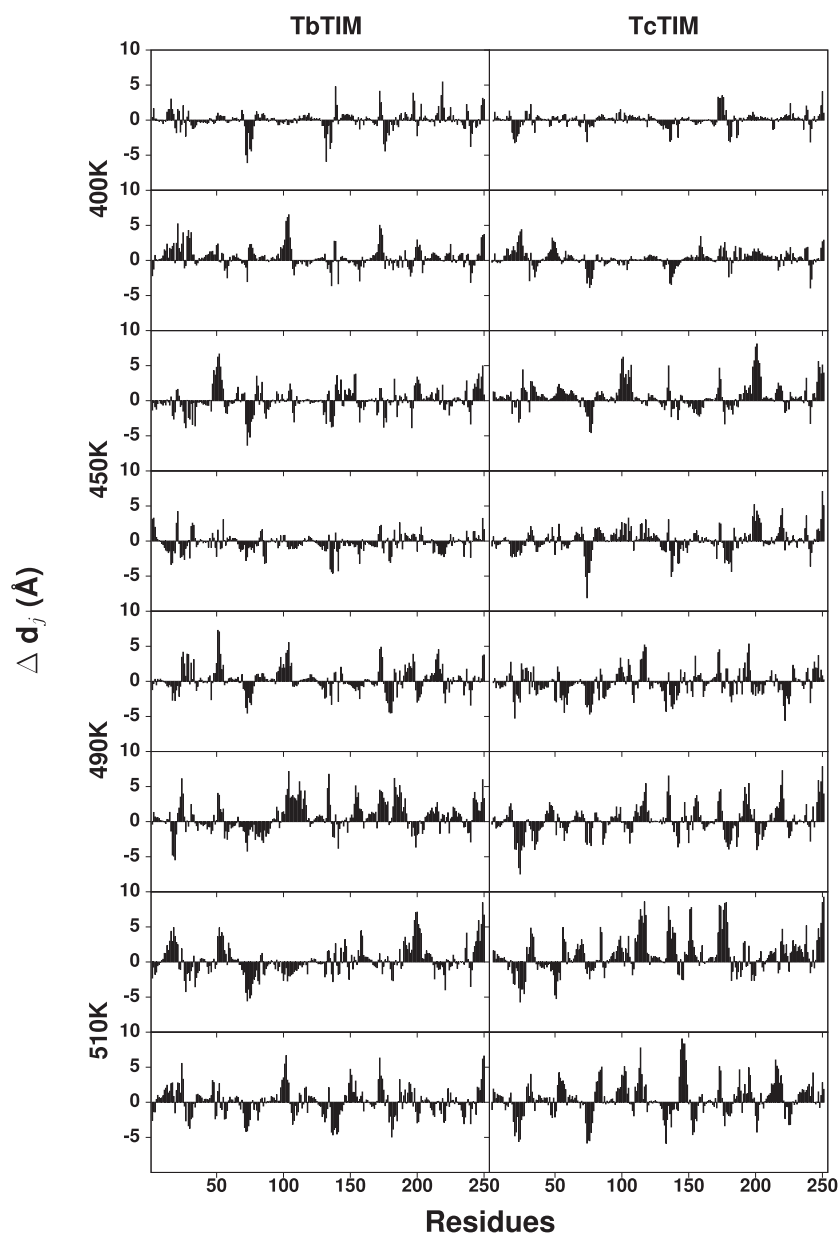


Figure 4. Evolution with the Simulation Time of the TbTIM and TcTIM Structures

(A) Snapshots of the simulation trajectories at 8, 16, and 24 ns for 400, 450, 490, and 510 K. For the simulations at 700 K, snapshots are at 2, 5, and 7 ns. Subunits A and B of TbTIM are colored blue and orange, while those of TcTIM are colored red and cyan.

(B) Projection of the trajectories on a 3D C α -RMSD space for the dimer and for both subunits of TbTIM and TcTIM at 510 K. The simulation time is represented by the colors, as indicated by the blue-red gradient bar.

(C) Progressive moving dispersion (PMD) plots over sets of 200 points for the trajectories shown in (B). Subunits A and B are colored as in (A), while the dimer is colored in black for both proteins. The minima represent structures that remain relatively stable during a significant part of the trajectory, while the maxima represent trajectory segments where the protein structure is evolving quickly.



conformations, whereas dynamic flexibility indicates the rate at which the conformational change takes place with time (Tang and Dill, 1998). Thus, the displacements $\Delta d_{j,c}$, $D_{r,T}$, and $D_{wt,T}$ calculated from the MD simulations can be considered as a quantitative measurement of local and global static flexibilities at a given temperature. When energy is provided, the protein undergoes partial or total denaturation increasing these magnitudes (Figures 5, 6A, and 6B). The temperature dependence of these displacements is quantified by their TGD values, which we denote here as thermal flexibilities, a term inspired by the definition of dynamic flexibility since, in this case, temperature instead of time is the relevant variable. On the other hand, according to simple Lumry-Eyring models (Costas et al., 2009; Rodriguez-Larrea et al., 2006; Sanchez-Ruiz, 2010) protein kinetic stability is linked to the timescale of irreversible denaturation and the corresponding kinetically relevant transition state

upon heating, ultimately lead to the unfolded state. E_{act} is directly related to cooperativity that, in turn, is inversely related to the width of the calorimetric transition (W) at half the maximum height (Costas et al., 2009). Hence, the TGD values obtained from our structural analysis should correlate well with the experimental W values in Table 1. This is the case, as shown in Figure 7B, corresponding to the same set of proteins used in Figure 7A. In short, protein kinetic stability and unfolding cooperativity are linked to the thermal flexibility, i.e., the temperature dependence of protein static flexibility.

The predictive ability of the correlations linking protein thermal flexibility with kinetic stability and with cooperativity of the unfolding transition was tested with the remaining 23 chimeras (see Table 1). We found that the predicted E_{act} and W values are in reasonably good agreement with the experimental ones (Figures 7C and 7D). The dispersion of points in Figures 7C

Figure 5. Residue Displacements

Difference of the average distance of each $C\alpha$ atom (Δd_j) between the last cluster structure of each trajectory and the crystal structures of TbTIM and TcTIM, using analysis of subunits A and B.

A

B

A

B

A

B

A

B

is expected to be analogous to that of protein-unfolding processes. The experimental quantitative measurement of kinetic stability, i.e., the activation energy E_{act} , can be viewed as the difference in enthalpy between the transition state and the native state, and it has been interpreted in terms of desolvation/solvation barriers in protein-folding/unfolding processes (Liu and Chan, 2005; Rodriguez-Larrea et al., 2006). In this interpretation, as energy is given to the system a network of water-deficient (broken or weakened) internal contacts between residues is created. Therefore, E_{act} reflects the evolution of this network with temperature and hence a working hypothesis would be that TGD values obtained from the MD simulations and E_{act} values obtained from the calorimetric experiments, i.e., thermal flexibility and kinetic stability, must be correlated. Figure 7A shows that there is a good correlation between E_{act} and TGD for the wild-type enzymes and a series of six chimeras, where TbTIM is progressively “converted” into TcTIM by replacing sequentially each region of TbTIM by the corresponding region in TcTIM. TGD values for these chimerical proteins were calculated using Equations 7 and 9, and their E_{act} values are reported in Table 1. On the other hand, the cooperativity of the unfolding transition refers to the concerted residue displacements that,

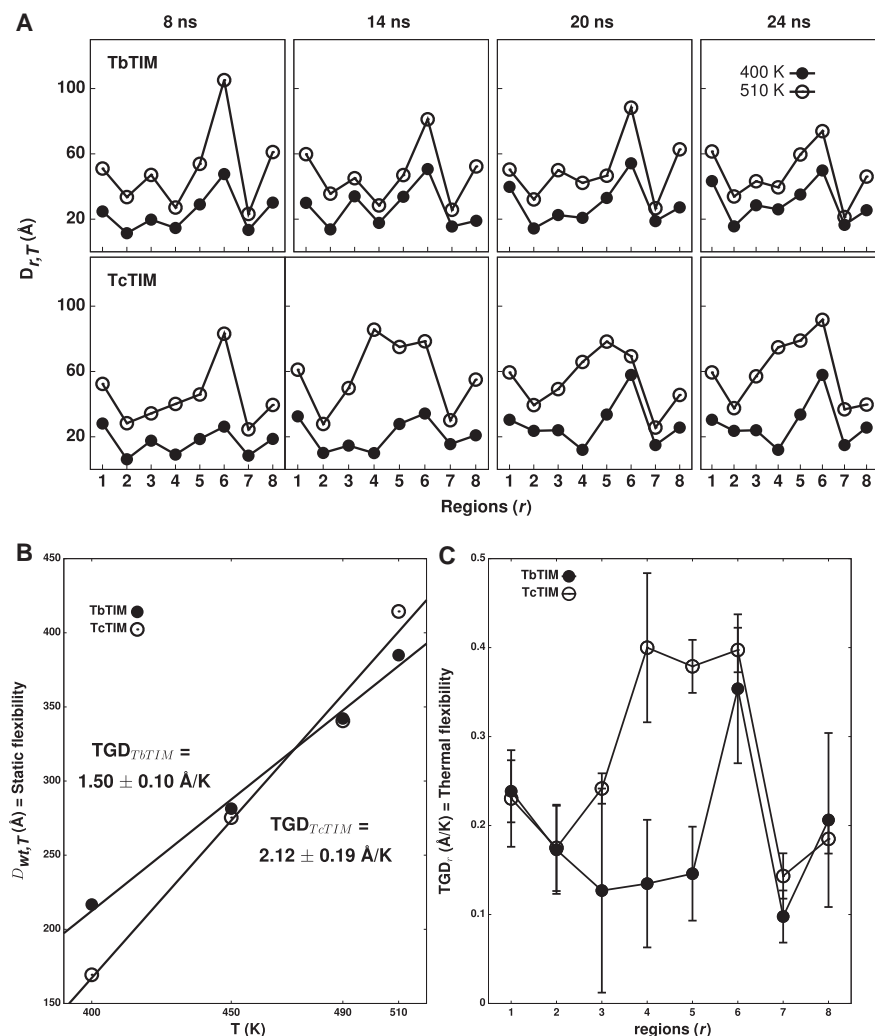


Figure 6. Displacements and Their Temperature Gradients

(A) Region displacements (average for both subunits) of the clusters structures around 2, 8, 14, 20, and 24 ns.

(B) Total $D_{wt,T}$ at each temperature for TbTIM and TcTIM.

(C) Temperature gradient of the displacement (TGD) for each region of both wild-type proteins. $D_{wt,T}$ and TGD are the average for both subunits and for all clusters. The error bars in (C) are the uncertainties of the slopes of the straight lines fitted to the $D_{r,T}$ versus T data for each region of both proteins.

high temperature (510 K). It is then important to assess whether the activation energies might be correlated with the static flexibility ($D_{X,T}$ values) at constant temperature. Calculations of $D_{X,T}$ values at 400 and 510 K showed that the correlation coefficients of the straight lines fitted to the data (E_{act} versus $D_{X,T}$) are 0.78 and 0.32, respectively, which are much lower than the value of 0.96 in Figure 7A obtained with TGD. Hence, the flexibility that mainly determines the protein kinetic stability is the thermal flexibility (TGD) and not the static one. Finally, no correlation was observed between the thermal stability of the proteins, as measured by their T_m values (Table 1), and the TGD values (see Figure S10). This is not a surprise since thermal and kinetic stabilities refer to different processes (Sanchez-Ruiz, 2010).

The fact that protein thermal flexibility is related to the cooperativity of the unfold-

ing transition (Figures 7B and 7D), reveals another interesting difference between the wild-type TcTIM and TbTIM proteins. Figure 6C indicates that for TbTIM region 6 has a TGD value which is much higher than for the other regions and hence appears to behave as an independent cooperative unfolding unit. On the other hand, for TcTIM regions 4, 5, and 6, sharing high TGD values, constitute a cooperative unfolding unit. Differential unfolding cooperativity involving several protein regions have also been found during unfolding simulations (Salimi et al., 2010). Although the grouping of regions into cooperative units is somewhat arbitrary, it is clear that the TGD values for some regions are quite similar in both proteins, while for other regions they differ greatly. The net result is that the TGD value for TcTIM is larger than for TbTIM. The adjacent central regions 4 and 5 contribute more than the extreme regions 1–3 and 6–8, which ultimately translates into a more cooperative unfolding transition for TcTIM, as has been found experimentally (W values in Table 1).

The statements that the higher the flexibility the lower the stability (Tuena de Gómez-Puyou and Gómez-Puyou, 1998) and that a rigid protein structure might be the basis for kinetic stability (Manning and Colón, 2004) appear to be in contradiction with the findings displayed in Figure 7. It is our contention that these statements are correct but they refer the static flexibility ($D_{X,T}$ values) at constant temperature. As indicated above in regard to Figure 6B, TbTIM is a more flexible protein (higher $D_{wt,T}$ value) than TcTIM at low temperature (400 K), while the reverse is true at

ing transition (Figures 7B and 7D), reveals another interesting difference between the wild-type TcTIM and TbTIM proteins. Figure 6C indicates that for TbTIM region 6 has a TGD value which is much higher than for the other regions and hence appears to behave as an independent cooperative unfolding unit. On the other hand, for TcTIM regions 4, 5, and 6, sharing high TGD values, constitute a cooperative unfolding unit. Differential unfolding cooperativity involving several protein regions have also been found during unfolding simulations (Salimi et al., 2010). Although the grouping of regions into cooperative units is somewhat arbitrary, it is clear that the TGD values for some regions are quite similar in both proteins, while for other regions they differ greatly. The net result is that the TGD value for TcTIM is larger than for TbTIM. The adjacent central regions 4 and 5 contribute more than the extreme regions 1–3 and 6–8, which ultimately translates into a more cooperative unfolding transition for TcTIM, as has been found experimentally (W values in Table 1).

Residue Individual Contributions to the Thermal Flexibility

The TGD values for each residue (TGD_{res}) of TcTIM and TbTIM are shown in Figure 8. This refinement allows the identification

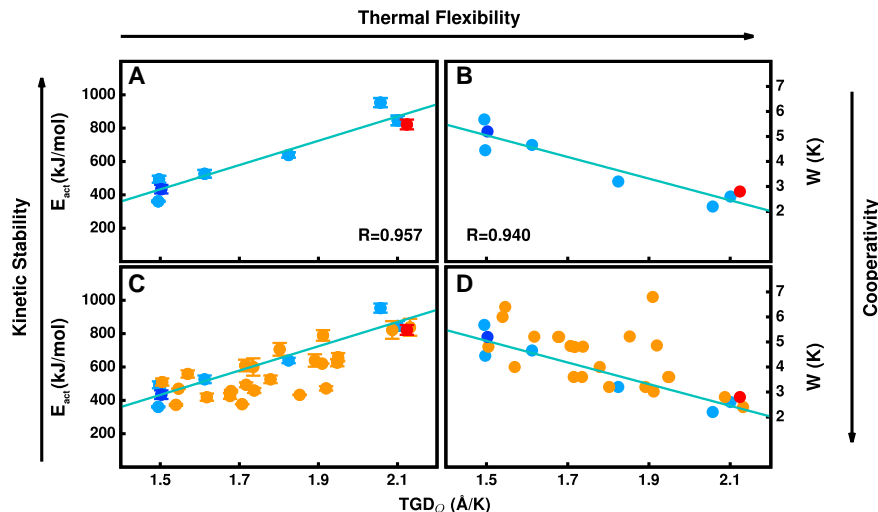


Figure 7. Correlation between the Calorimetric Experimental Results and MD Simulations

E_{act} (A) and W (B) versus TGD for the wild-type proteins and a series of six chimeras where TbTIM is progressively converted into TcTIM: Tc(1)Tb(2–8), Tc(1–2)Tb(3–8), Tc(1–3)Tb(4–8), Tc(1–4)Tb(5–8), Tc(1–5)Tb(6–8), and Tc(1–6)Tb(7–8). TGD was calculated using Equations 7 and 9, and the E_{act} values are those reported in Table 1. Red and dark blue symbols are for wild-type TcTIM and TbTIM, respectively. The straight lines ($E_{act} = -660.01 + 728.54 \times TGD$ and $W = 11.53 - 4.33 \times TGD$) correspond to the best fit of the data and the correlation coefficients (R) are shown in the plots. The orange symbols are E_{act} (C) and W (D) versus TGD for the remaining 23 chimeras in Table 1. The straight lines in (C) and (D) are the same as those in (A) and (B), respectively, showing that the predicted E_{act} and W values from the correlation are in reasonable agreement with the experimental ones. Arrows indicate the direction of increasing cooperativity of the unfolding transition, kinetic stability, and thermal flexibility.

of the individual contributions of every lateral chain to the TGD values of the regions shown in Figure 6C. For regions 4 and 5, the sets of residues responsible for the high TGD value for TcTIM (compared with that for TbTIM) are identified. The total number of TcTIM residues in this case (i.e., those within the boxes in Figure 8) is 31, and these are listed in Table S2. Figure 8 also indicates that in these two regions these residues are located in the α helices, the protein domains whose thermal flexibility is predicted to be the largest. Eleven of these 31 identified residues are different in TcTIM compared with TbTIM (Table S2). Among these 11 differences there is no predominance of conservative, non-conservative, or semi-conservative substitutions, indicating that thermal flexibility is more the result of concerted effects between many residues that occur in a global spatial context, than of the chemical identity of the involved amino acids. This scrutiny at the residue level is justified by the following two clear connections to published results. First, 15 of the residues identified in region 5 (141–155) are within a 22-amino-acid long fragment in TbTIM that has been reported as critical to route a reporter protein to glycosomes in transfected trypanosomes (Galland et al., 2010). Second, it has been reported that regions 1 and 4 of TcTIM play a critical role in the susceptibility of the enzymes exposed to high and low concentrations of a sulfhydryl reagent (methylmethane sulfonate), respectively (García-Torres et al., 2011). Region 4 is identified as one with high TGD values for TcTIM and although the TGD values for region 1 are statistically the same for TcTIM and TbTIM (Figure 6C), Figure 8 shows there are three residues in TcTIM with markedly higher TGD values (Val22, Pro23, Glu26) than for the corresponding residues in TbTIM (Ser22, Glu23, Asp26); two of these amino acids (Val, Pro) were introduced in the chimera Tc(2,3,5–8)Tb(1,4) substituting those in TbTIM, to produce a five-residue mutant that showed an important decrease in the susceptibility to the sulfhydryl reagent (methylmethane sulfonate) compared with wild-type TcTIM (García-Torres et al., 2011). Therefore, it appears that these residues and those in Table S2 do play an important role in the observed experimental behavior of TcTIM and TbTIM.

Concluding Remarks

Protein flexibility is a widely used concept to understand many aspects of protein behavior. However, what should be understood by the statement that a protein is more or less flexible than another? In this work, we propose a new concept regarding protein flexibility, namely thermal flexibility, which refers to the change of residue displacements due to a temperature gradient. As such, thermal flexibility quantitatively measures the increment of the size of the conformational space available to the protein when energy is provided. Figure 9 shows schematically the effects of the static and thermal flexibilities on the thermal energy versus the conformational space landscape for TcTIM and TbTIM, producing different kinetics stabilities for each protein.

Using as model systems the $(\beta/\alpha)_8$ barrel fold of two TIMs closely related by evolution, it has been demonstrated that there is a correlation between kinetic stability and thermal flexibility. The latter was obtained by a novel method based on temperature-induced unfolding MD simulations, whereas the former was experimentally determined through calorimetry. This correlation increases our knowledge of the molecular basis of protein kinetic stability. Although further work would be needed to generalize this correlation, the employed strategy can be helpful in the design of proteins, through single or multiple mutations, with small or large kinetic stabilities, the latter being a desirable goal from the biotechnological perspective.

EXPERIMENTAL PROCEDURES

Construction, Gene Design, Expression, and Purification of the Chimerical Proteins

For the design of the chimerical proteins, DNA sequences NCBI: X03921 and U53867 for wild-type TbTIM and TcTIM, respectively, were used. All the genes were synthesized by GenScript, except for the genes of the chimeras Tc(4)Tb(1–3,5–8), Tc(1,2)Tb(3–8), Tc(1)Tb(2–8), and Tc(2,3,5–8)Tb(1,4) that were obtained by different PCR reactions using the ACCUZYME DNA polymerase (Bioline) and the external T7 promoter and terminator oligonucleotides (for the specific mutagenic oligonucleotides and the template used to obtain

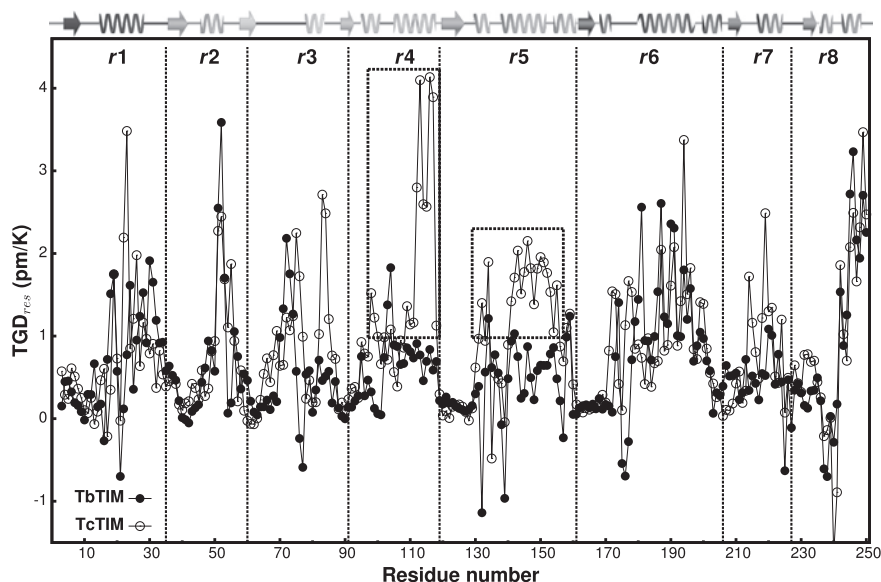


Figure 8. Comparison of TGD_{res} Values for Each of the Residues of TcTIM and TbTIM

Regions are separated by dashed lines. Residues with a significant difference in TGD values between TbTIM and TcTIM are highlighted. The secondary structure is depicted at the top of the plot.

DSC

DSC experiments were performed using a VP-DSC capillary calorimeter (Malvern). Protein solutions were prepared by exhaustive dialysis against 100 mM TEA and 10 mM EDTA (pH 7.4). The protein concentration was 0.4 mg/mL, determined from the absorbance at 280 nm using the Cary 50 spectrophotometer. The scanning rates ranged from 0.5 to 3 K/min. The calorimetric traces were fitted to the two-state irreversible model (Aguirre et al., 2014; Costas et al., 2009; Rodriguez-Larrea et al., 2006) using homemade software. In this model it is assumed that, upon heating, the native protein (N) undergoes a conformational transition to a final state (F) which is unable to fold back.

The kinetic conversion between these two states (N → F) is described by a first-order rate constant that is assumed to change with temperature according to the Arrhenius equation. The apparent heat capacity in this model is given by (Rodriguez-Larrea et al., 2006):

$$C_p^{APP} = C_p^{PRE} + (C_p^{POST} - C_p^{PRE})(1 - X_N) - \Delta H \cdot \left(\frac{dX_N}{dT} \right), \quad (\text{Equation 1})$$

with

$$X_N = \exp \left[- \exp \left(\frac{E_{act} \cdot \Delta T}{R \cdot T_m^2} \right) \right], \quad (\text{Equation 2})$$

$$\frac{dX_N}{dT} = - \frac{E_{act}}{R \cdot T_m^2} \cdot \exp \left(\frac{E_{act} \cdot \Delta T}{R \cdot T_m^2} \right) \cdot \exp \left[- \exp \left(\frac{E_{act} \cdot \Delta T}{R \cdot T_m^2} \right) \right], \quad (\text{Equation 3})$$

where X_N is the mole fraction of native state, ΔH is the denaturation enthalpy (N to F transition), C_p^{PRE} and C_p^{POST} are the pre- and post-transition baselines which were taken to be linear functions of temperature, T_m is the temperature corresponding to the maximum of the unfolding transition, $\Delta T = T - T_m$, and E_{act} is the activation energy separating the native state from the transition state. Each individual experimental DSC profile was fitted with Equations 1–3.

MD Simulations

Simulation Boxes and Parameters

MD simulations of the wild-type proteins TbTIM and TcTIM (starting from the crystal structures with PDB: 5TIM and 1TCD, respectively) were performed at 400, 450, 490, 510, and 700 K using the GROMACS package (Berendsen et al., 1995; Hess et al., 2008; Lindahl et al., 2001) version 4.5.1 and the GROMOS53a6 force field (Oostenbrink et al., 2004). Truncated dodecahedron boxes with periodic boundary conditions in the three spatial dimensions were employed in all cases. To prevent the direct interaction between periodic images upon protein unfolding within the cutoffs for the electrostatic and van der Waals interactions (see below), the size of the simulation unit cell in the initial configuration was chosen to provide a minimum distance of 0.9 nm between every protein atom and the box walls. The simulation boxes were solvated with 23,287 (TbTIM) and 19,729 (TcTIM) water molecules using an explicit simple point charge model (Berendsen et al., 1981). The energy of each system was minimized using the steepest descent method and the resulting configurations were equilibrated at 298 K by 5-ns-long MD simulations using the NPT ensemble. The heavy atoms (all except hydrogens) were constrained to their initial position by a

each of these four chimeras (see García-Torres et al., 2011). All genes were cloned into the pET-3a expression plasmid using the NdeI and BamHI restriction sites, and completely sequenced and transformed into BL21 (DE3)pLysS cells (Novagen). Bacteria containing the plasmids with each of the chimerical genes were grown in Luria-Bertani medium supplemented with 100 μg/mL of ampicillin and then incubated at 310 K. Once the cell cultures reached an $A_{600\text{ nm}} = 0.6$, a final concentration of 0.4 mM isopropyl-β-D-thiogalactopyranoside was used for induction and the bacteria were incubated for a further 12 hr at 303 K before harvesting them. After centrifugation, the cells were resuspended in 20 mL of lysis buffer containing 100 mM 2-(N-morpholino)ethanesulfonic acid hydrate (MES), 1 mM DTT, 0.5 mM EDTA, 0.2 mM phenylmethylsulfonyl fluoride, and 300 mM NaCl (pH 6.3). The cells were sonicated and centrifuged at $1.44 \times 10^5 \times g$ for 1 hr. The supernatant was diluted to a final concentration of approximately 20 mM NaCl and then applied to a fast flow SP Sepharose column that had been previously equilibrated with 50 mM MES buffer (pH 6.3), and then eluted with a 0–500 mM NaCl gradient in the same buffer. The second step in the purification process consisted in the precipitation with ammonium sulfate, $(\text{NH}_4)_2\text{SO}_4$ (70% w/v), for 3 hr. The precipitate was centrifuged at $2.3 \times 10^4 \times g$ for 15 min and dissolved in 3 mL of 100 mM triethanolamine (TEA), 1 mM EDTA (pH 7.4), and 2.2 M $(\text{NH}_4)_2\text{SO}_4$. This solution was applied to a hydrophobic TOYOPEARL column, previously equilibrated with the same buffer. The proteins were eluted with a linear gradient of 2.2–0 M $(\text{NH}_4)_2\text{SO}_4$. The protein was stored at 277 K in 70% $(\text{NH}_4)_2\text{SO}_4$. SDS-PAGE gels (16% acrylamide) stained with Coomassie blue were used at different points along the purification process to verify the adequacy of each step.

Enzyme Activity

Enzyme activity was indirectly determined in the glyceraldehyde 3-phosphate (GAP) to dihydroxyacetone phosphate (DHAP) direction with the α-glycerol phosphate dehydrogenase (α-GDH) reaction as a coupled-enzyme assay. The reaction started by addition of 5 ng/mL of the enzyme to the cell mixture containing 100 mM TEA, 10 mM EDTA (pH 7.4), 1 mM GAP, 0.2 mM β-nicotinamide adenine dinucleotide (NADH), and 5 μg/mL α-GDH. The oxidation of NADH was followed spectrophotometrically at 340 nm. To calculate the catalytic parameters, GAP concentration was varied between 0.02 and 3.2 mM. The data were adjusted to the Michaelis-Menten model and the values of K_M and V_{max} were calculated by non-linear regression. To calculate the catalytic efficiency, the value of K_M was corrected considering that there is only 4% of free GAP in the solution (Trentham et al., 1969). All the assays were carried out in a Cary 50 spectrophotometer (Varian) with a multi-cell array. The temperature of the cell block was maintained at 298 K with a constant-temperature circulating water bath.

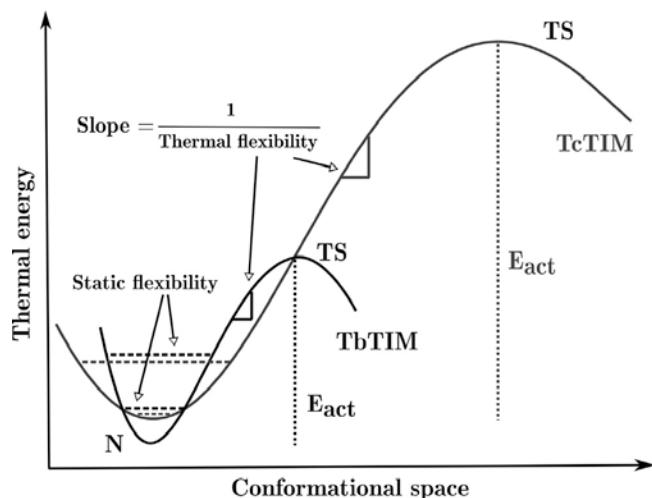


Figure 9. Thermal Energy versus the Conformational Space Landscape

Schematic diagram showing the effects of the static and thermal flexibilities on the thermal energy (E) versus the conformational space (CS) landscape for TcTIM and TbTIM, producing different kinetic stabilities (activation energy, E_{act}) for each protein. The minima corresponding to the native structures have been shifted in the vertical axis to show more clearly the difference in static flexibility between both proteins. At a given temperature (thermal energy) the static flexibility is proportional to the size of the CS available to the protein (dotted horizontal lines). As the temperature increases the static flexibility increases widening the CS available to the protein (see also Figure 6B). The thermal flexibility (TGD) is inversely proportional to the slope of the E versus the CS curve joining the native state (N) and transition state (TS), and represents the increment of the CS available to the protein when energy is provided. As a result of a higher thermal flexibility for TcTIM its kinetic stability is larger than for TbTIM. In this work, these slopes and hence TGD have been taken as temperature independent (see also Figures 6B and 6C).

harmonic potential with a force constant of 1,000 kJ/mol/nm during this equilibration process. The production simulations were run at constant volume and temperature (an NVT ensemble) for 24 ns at temperatures between 400 and 510 K and for 8 ns at 700 K. All these simulations started from the same initial structure but using different random velocities. The temperature was controlled using a v-rescale thermostat (Nose, 1984) with a coupling constant of 0.1 ps. For the simulations at constant pressure (only the 5-ns-long equilibration at 298 K) a Berendsen barostat (Parrinello and Rahman, 1981) was isotropically employed with a coupling constant of 0.5 ps and considering an isothermal compressibility of $4.5 \times 10^{-5} \text{ bar}^{-1}$. Long-range electrostatic interactions were calculated using the particle mesh Ewald method (Darden et al., 1993; Essmann et al., 1995) with a real-space cutoff of 1.2 nm, a 0.15 spaced grid, and a fourth-order B-spline interpolation. The van der Waals interactions were truncated at 1.2 nm. A long-range dispersion correction was applied to the energy and pressure. The neighbor list was updated every five steps. The leapfrog method (Hockney and Eastwood, 1988) was used to integrate the equation of motion with a time step of 2 fs. The linear constraint solver (Hess et al., 1997) algorithm was used to constrain the bond lengths of the protein while the SETTLE (Miyamoto and Kollman, 1992) algorithm was employed to constrain the bond lengths and angles of water molecules. Coordinates and energies were stored every 1 ps for analysis.

Standard Analysis of MD Trajectories

In the simulations at temperatures between 400 and 700 K the protein structures are expected to unfold at different rates. Tools from the GROMACS package were employed to characterize the evolution of the accessible surface area, the secondary structure, the hydrogen bonds, the radius of gyration, and the RMSD of the atomic positions from the X-ray structures for all the trajectories.

Individual Residue Displacements, Total Displacements, and Temperature Gradients

The effect of temperature-induced unfolding on each individual residue, on each protein subunit, on the entire proteins, as well as the specific differences between TbTIM and TcTIM, were assessed by a procedure that consists of the following steps:

- (1) For each trajectory, the RMSD matrix considering only the α carbons ($C\alpha$) of the protein dimers and of each protein subunit separately were built from the alignment of the different conformations to each other taken every 4 ps.
- (2) A multidimensional scaling analysis was carried out from the diagonalization of each $C\alpha$ -RMSD matrix. In all cases, only the three dimensions with larger eigenvalues were considered to be significant. The projection of the trajectory on this reduced RMSD space provided a representation where the similarity between structures is equivalent to their distance.
- (3) The protein structures in the reduced RMSD representation were clustered using a PMD calculation over sets of 200 sequential structures, corresponding to 800 ps. This PMD was determined as the average distance between every structure in each set and their mean position. The set of 200 structures was progressively redefined as $[i, \dots, i + 200]$ with i going from 1 to $n - 200$, n being the last structure of the trajectory. Note that the alignment performed to determine the RMSD might be significantly different for the whole protein and for the individual subunits. Thus, for each trajectory, simultaneous minima in the PMD of the dimer and both subunits were taken to identify the representative structures of the clusters along each trajectory. Minima in the PMD represent structures that remain relatively stable, while the maxima represent trajectory segments where the protein structure is evolving fast.
- (4) For each of these representative structures the average internal distance between each $C\alpha$ atom and every $C\alpha$ atom of the whole structure is then determined, as

$$d_{j,c} = \frac{1}{N} \sum_{i=1}^N |r_j^c - r_i^c|, \quad (\text{Equation 4})$$

where j and c refer to residue and cluster number, respectively, and N is either the total number of residues of the dimer or the number of residues of each monomer. Note that this parameter does not depend on any alignment with other structures since it is intrinsic to each structure and residue.

- (5) For each residue, the difference between this average internal distance and the corresponding value for the crystal structure, i.e., the residue displacement $\Delta d_{j,c}$, is given by

$$\Delta d_{j,c} = d_{j,c} - d_{j,0} \quad (\text{Equation 5})$$

evaluating how this individual residue internally moved from its original position. Positive $\Delta d_{j,c}$ values indicate that, on average, residue j moved “away” from the whole protein while the negative values correspond to the situation where the residue moved “closer” to the protein.

- (6) The total displacement $D_{X,T}$ at each temperature T was determined for X being the wild-type TbTIM or TcTIM proteins, as the average of the sum of the absolute values of $\Delta d_{j,c}$ for subunits A and B, averaged over all the clusters (N_C) except the first three (identified on the initial 4–7 ns) where the structural evolution of the protein is too fast:

$$D_{X,T} = \frac{1}{N_C - 3} \sum_{j=1}^N (|\Delta d_{j,c}|_A + |\Delta d_{j,c}|_B). \quad (\text{Equation 6})$$

- (7) TGD = $dD_{X,T}/dT$ for each protein was calculated fitting a straight line

$$D_{X,T} = a + (\text{TGD}) \cdot T \quad (\text{Equation 7})$$

to the $D_{X,T}$ versus T data (400, 450, 490, and 510 K). The displacement of each region r at every temperature T ($D_{r,T}$) of both wild-type TbTIM and TcTIM, was

calculated assuming that the regions are independent using Equation 6 but restricting the sum to the residues present in each region (see Figure 1). Thus, the TGD for each of the eight regions is also given by Equation 7 using $D_{r,T}$ instead of $D_{X,T}$. Note that the $D_{X,T}$ and $D_{r,T}$ values employed to determine the TGD are the sum of the independent contributions of the residues constituting each protein or region. Thus, it is easy to see that the value of $D_{X,T}$ can be obtained from the sum of the $D_{r,T}$ for the eight regions:

$$TGD_X = \sum_{r=1}^8 TGD_r. \quad (\text{Equation 8})$$

Both the displacements $D_{X,T}$ in Equation 6 and TGD in Equation 7 can be normalized using the total number of residues in the protein, allowing a direct comparison between different proteins.

Estimation of the TGD for the Chimeras and Individual Residues

For each of the chimerical proteins (Q) the displacement $D_{Q,T}$ was estimated assuming region additivity:

$$D_{Q,T} = \sum_{r_{Tc}} D_{r_{Tc}} + \sum_{r_{Tb}} D_{r_{Tb}}, \quad (\text{Equation 9})$$

where r_{Tc} and r_{Tb} indicate the regions of TcTIM and TbTIM that form the chimera, respectively. Then, the TGD for the chimera (TGD_Q) is obtained using Equation 7. For each individual residue, the displacement ($D_{res,T}$) is given by

$$D_{res,T} = \frac{1}{N_{C-3}} \sum_{C=4}^{N_C} \frac{1}{2} (|\Delta d_{j,c}|_A + |\Delta d_{j,c}|_B) \quad (\text{Equation 10})$$

and the TGD_{res} is estimated using Equation 7.

Note that the selection of clusters is based on alignments between C α coordinates of different conformations along the trajectories, and that each RMSD value is a single number representing the difference between two aligned structures typically formed by thousands of atoms. In contrast, the determination of the residue displacements and of the TGD values does not depend on any alignment but only on the structure to be analyzed. All the previous analyses were applied to the dimeric wild-type proteins and chimeras (dimer analysis) and also to each protein subunit (subunit analysis). This is because, using the dimer analysis, potential relative movements between subunits (approaching, separation, or rotation) might screen information on structural changes and also would affect the determination of the displacements and TGD values. Note that the average distance of a residue to the subunit to which it belongs is different from the average distance of the same residue to the entire protein, even when both amounts are internal values and do not depend on any alignment. The results obtained using the dimers are expected to contain information of the relative movement between subunits and of protein unfolding, i.e., changes on the quaternary, tertiary, and secondary structures. On the other hand, the results obtained using the subunits will refer to changes in the tertiary and secondary structures on the corresponding monomer. Thus, the employment of the whole dimer or each monomer separately is expected to provide complementary useful information. Considering our approach regarding the chimerical proteins, using the subunit analysis is a better option than dimer analysis to characterize the unfolding process, since the former is free from the possible ‘‘contamination’’ due to quaternary structural changes of the wild-type enzymes. Hence, the subunit analysis was employed throughout the manuscript.

SUPPLEMENTAL INFORMATION

Supplemental Information includes ten figures and two tables and can be found with this article online at <http://dx.doi.org/10.1016/j.str.2016.11.018>.

AUTHOR CONTRIBUTIONS

A.G.Q., Á.P., and M.C. conceived the project; A.G.Q., N.C., and R.P.-M. obtained and purified the chimerical proteins; A.J.D.-S. performed the enzyme activity assays; A.G.Q. performed the DSC experiments and the MD simulations; A.G.Q., Á.P., and M.C. analyzed the data and wrote the manuscript with contributions from all authors.

ACKNOWLEDGMENTS

We thank Itzhel García-Torres for her kind gift of some chimerical proteins at the onset of this project. A.G.Q. thanks CONACyT-México for a graduate student scholarship. This work was supported by a PAIP-FQ-UNAM grant to M.C. (5000-9018), DGAPA-UNAM-PAPIIT grants to M.C. (IN112813) and to R.P.-M. (IN221812 and IN206816), CONACyT-México grants (254694 and 167823), grants from the Spanish Government (MAT2015-71826-P) and from Xunta de Galicia (AGRUP2015/11) to Á.P. All calculations were carried out at Centro de Supercomputación de Galicia (CESGA).

Received: August 11, 2016

Revised: October 18, 2016

Accepted: November 22, 2016

Published: January 3, 2017

REFERENCES

- Aguirre, Y., Cabrera, N., Aguirre, B., Pérez-Montfort, R., Hernandez-Santoyo, A., Reyes-Vivas, H., Enríquez-Flores, S., de Gómez-Puyou, M., Gómez-Puyou, A., Sanchez-Ruiz, J.M., and Costas, M. (2014). Different contribution of conserved amino acids to the global properties of triosephosphate isomerases. *Proteins* 82, 323–335.
- Benson, N.C., and Daggett, V. (2008). Dynameomics: large-scale assessment of native protein flexibility. *Protein Sci.* 17, 2038–2050.
- Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., and Hermans, J. (1981). Interaction models for water in relation to protein hydration. In *Intermolecular Forces. The Jerusalem Symposia on Quantum Chemistry and Biochemistry*, vol. 14 (Springer), pp. 331–342.
- Berendsen, H.J.C., van der Spoel, D., and van Drunen, R. (1995). GROMACS: a message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* 91, 43–56.
- Bippes, C.A., Zeltina, A., Casagrande, F., Ratera, M., Palacin, M., Muller, D.J., and Fotiadis, D. (2009). Substrate binding tunes conformational flexibility and kinetic stability of an amino acid antiporter. *J. Biol. Chem.* 284, 18651–18663.
- Cansu, S., and Doruker, P. (2008). Dimerization affects collective dynamics of triosephosphate isomerase. *Biochemistry* 47, 1358–1368.
- Chávez-Cárdenas, M.E., Pérez-Hernández, G., Sánchez-Rebollar, B.G., Costas, M., and Vázquez-Contreras, E. (2005). Reversible equilibrium unfolding of triosephosphate isomerase from *Trypanosoma cruzi* in guanidinium hydrochloride involves stable dimeric and monomeric intermediates. *Biochemistry* 44, 10883–10892.
- Cheung, M.S., García, A.E., and Onuchic, J.N. (2002). Protein folding mediated by solvation: water expulsion and formation of the hydrophobic core occur after the structural collapse. *Proc. Natl. Acad. Sci. USA* 99, 685–690.
- Costas, M., Rodríguez-Larrea, D., De María, L., Borchert, T.V., Gómez-Puyou, A., and Sanchez-Ruiz, J.M. (2009). Between-species variation in the kinetic stability of TIM proteins linked to solvation-barrier free energies. *J. Mol. Biol.* 385, 924–937.
- Darden, T., York, D., and Pedersen, L. (1993). Particle mesh Ewald: an N.log(N) method for Ewald sums in large systems. *J. Chem. Phys.* 98, 10089–10092.
- Das, P., Kapoor, D., Halloran, K.T., Zhou, R., and Matthews, C.R. (2013). Interplay between drying and stability of a TIM barrel protein: a combined simulation – experimental study. *J. Am. Chem. Soc.* 135, 1882–1890.
- Essmann, U., Perera, L., Berkowitz, M.L., Darden, T., Lee, H., and Pedersen, L.G. (1995). A smooth particle mesh Ewald method. *J. Chem. Phys.* 103, 8577–8593.
- Galland, N., de Walque, S., Voncken, F.G.J., Verlinde, C.L.M.J., and Michels, P. (2010). An internal sequence targets *Trypanosoma brucei* triosephosphate isomerase to glycosomes. *Mol. Biochem. Parasitol.* 171, 45–49.
- García-Torres, I., Cabrera, N., Torres-Larios, A., Rodríguez-Bolaños, M., Díaz-Mazariegos, S., Gómez-Puyou, A., and Perez-Montfort, R. (2011). Identification of amino acids that account for long-range interactions in two triosephosphate isomerases from pathogenic trypanosomes. *PLoS One* 6, e18791.

- Guzman-Luna, V., and Garza-Ramos, G. (2012). The folding pathway of glycosomal triosephosphate isomerase: structural insights into equilibrium intermediates. *Proteins* *80*, 1669–1682.
- Hess, B., Bekker, H., Berendsen, H.J.C., and Fraaije, J.G.E.M. (1997). LINCS: a linear constraint solver for molecular simulations. *J. Comput. Chem.* *18*, 1463–1472.
- Hess, B., Kutzner, C., van der Spoel, D., and Lindahl, E. (2008). GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theor. Comput.* *4*, 435–447.
- Höcker, B. (2013). Engineering chimaeric proteins from fold fragments: “hopeful monsters” in protein design. *Biochem. Soc. Trans.* *41*, 1137–1140.
- Höcker, B., Claren, J., and Sterner, R. (2004). Mimicking enzyme evolution by generating new (betaalpha)8-barrels from (betaalpha)4-half-barrels. *Proc. Natl. Acad. Sci. USA* *101*, 16448–16453.
- Hockney, R.W., and Eastwood, J.W. (1988). *Computer Simulation Using Particles* (Adam Hilger).
- Huang, P.-S., Feldmeier, K., Parmeggiani, F., Fernandez Velasco, D.A., Höcker, B., and Baker, D. (2016). De novo design of a four-fold symmetric TIM-barrel protein with atomic-level accuracy. *Nat. Chem. Biol.* *12*, 29–34.
- Jamroz, M., Kolinski, A., and Kmiecik, S. (2013). CABS-flex: server for fast simulation of protein structure fluctuations. *Nucleic Acids Res.* *41*, W427–W431.
- Jefferson, R.E., Blois, T.M., and Bowie, J.U. (2013). Membrane proteins can have high kinetic stability. *J. Am. Chem. Soc.* *135*, 15183–15190.
- Kamerzell, T.J., and Middaugh, C.R. (2008). The complex inter-relationships between protein flexibility and stability. *J. Pharm. Sci.* *97*, 3494–3517.
- Kurkcuoglu, Z., and Doruker, P. (2013). Substrate effect on catalytic loop and global dynamics of triosephosphate isomerase. *Entropy* *15*, 1085.
- Ladurner, A.G., Itzhaki, L.S., Daggett, V., and Fersht, A.R. (1998). Synergy between simulation and experiment in describing the energy landscape of protein folding. *Proc. Natl. Acad. Sci. USA* *95*, 8473–8478.
- Lindahl, E., Hess, B., and van der Spoel, D. (2001). GROMACS 3.0: a package for molecular simulation and trajectory analysis. *Mol. Model. Annu.* *7*, 306–317.
- Liu, Z., and Chan, H.S. (2005). Solvation and desolvation effects in protein folding: native flexibility, kinetic cooperativity and enthalpic barriers under iso-stability conditions. *Phys. Biol.* *2*, S75–S85.
- Ma, J. (2005). Usefulness and limitations of normal mode analysis in modeling dynamics of biomolecular complexes. *Structure* *13*, 373–380.
- Manning, M., and Colón, W. (2004). Structural basis of protein kinetic stability: resistance to sodium dodecyl sulfate suggests a central role for rigidity and a bias toward β -sheet structure. *Biochemistry* *43*, 11248–11254.
- Miyamoto, S., and Kollman, P.A. (1992). Settle: an analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J. Comput. Chem.* *13*, 952–962.
- Monsellier, E., and Chiti, F. (2007). Prevention of amyloid-like aggregation as a driving force of protein evolution. *EMBO Rep.* *8*, 737–742.
- Nose, S. (1984). A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.* *52*, 255–268.
- Oostenbrink, C., Villa, A., Mark, A.E., and Van Gunsteren, W.F. (2004). A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *J. Comput. Chem.* *25*, 1656–1676.
- Parrinello, M., and Rahman, A. (1981). Polymorphic transitions in single crystals: a new molecular dynamics method. *J. Appl. Phys.* *52*, 7182–7190.
- Patel, A.B., Allen, S., Davies, M.C., Roberts, C.J., Tendler, S.J.B., and Williams, P.M. (2004). Influence of architecture on the kinetic stability of molecular assemblies. *J. Am. Chem. Soc.* *126*, 1318–1319.
- Reyes-Vivas, H., Martínez-Martínez, E., Mendoza-Hernández, G., López-Velázquez, G., Pérez-Montfort, R., Tuena de Gómez-Puyou, M., and Gómez-Puyou, A. (2002). Susceptibility to proteolysis of triosephosphate isomerase from two pathogenic parasites: characterization of an enzyme with an intact and a nicked monomer. *Proteins* *48*, 580–590.
- Rodríguez-Larrea, D., Minning, S., Borchert, T.V., and Sanchez-Ruiz, J.M. (2006). Role of solvation barriers in protein kinetic stability. *J. Mol. Biol.* *360*, 715–724.
- Salimi, N.L., Ho, B., and Agard, D.A. (2010). Unfolding simulations reveal the mechanism of extreme unfolding cooperativity in the kinetically stable α -lytic protease. *PLoS Comput. Biol.* *6*, e1000689.
- Sanchez-Ruiz, J.M. (2010). Protein kinetic stability. *Biophys. Chem.* *148*, 1–15.
- Smith, M.A., Romero, P.A., Wu, T., Brustad, E.M., and Arnold, F.H. (2013). Chimeragenesis of distantly-related proteins by noncontiguous recombination. *Protein Sci.* *22*, 231–238.
- Tang, K.E., and Dill, K.A. (1998). Native protein fluctuations: the conformational-motion temperature and the inverse correlation of protein flexibility with protein stability. *J. Biomol. Struct. Dyn.* *16*, 397–411.
- Teodoro, M.L., Phillips, G.N., and Kavraki, L.E. (2003). Understanding protein flexibility through dimensionality reduction. *J. Comput. Biol.* *10*, 617–634.
- Trentham, D.R., McMurray, C.H., and Pogson, C.I. (1969). The active chemical state of D-glyceraldehyde 3-phosphate in its reactions with D-glyceraldehyde 3-phosphate dehydrogenase, aldolase and triose phosphate isomerase. *Biochem. J.* *114*, 19–24.
- Tuena de Gómez-Puyou, M., and Gómez-Puyou, A. (1998). Enzymes in low water systems. *Crit. Rev. Biochem. Mol. Biol.* *33*, 53–89.
- Xia, K., Manning, M., Hesham, H., Lin, Q., Bystroff, C., and Colón, W. (2007). Identifying the subproteome of kinetically stable proteins via diagonal 2D SDS/PAGE. *Proc. Natl. Acad. Sci. USA* *104*, 17329–17334.

Apéndice B

Parámetros de DSC de otras quimeras de TbTIM/TcTIM

Tabla B.1: Parámetros de DSC de otras quimeras de TbTIM/TcTIM^a

Proteína	E_{act} (kJ/mol)	W (K)	T_m (K) ^b	$k_{cat}/K_M \times 10^{-8}$ ($M^{-1}s^{-1}$) ^c
Quimeras de regiones completas				
Tc(1,3-8)Tb(2)	367 ± 121	6.8	327.1	0.02
Tc(1,3,4,6-8)Tb(2,5)	304 ± 2.6	8.8	317.2	0.67
Quimeras de residuos puntuales				
TbTIM+9mutsd	360 ± 4	6.0	328.61	0.93
TbTIM-V61Q/M122He	393 ± 39	5.2	326.0	ND ^f
TbTIM-S43P/V46L/A49P	437 ± 1.9	335.7	5.2	ND
TbTIM-A69T/P81Q	519 ± 16.2	4.4	328.3	ND
TbTIM+5mutsg	365 ± 3.1	5.6	329.1	ND
TbTIM+5mutsh,e	416 ± 42	5.6	325.3	ND
TbTIM+7mutsi	427 ± 2.7	5.6	335.5	ND
TbTIM+12mutsj	524 ± 28.4	4.4	332.8	ND

^a Velocidades de barrido de 0.5 a 3 K/min, a menos que se indique otra cosa;

^b Los valores corresponden a 1.5 K/min;

^c Valor de K_M considerando 4% de GAP libre en solución [113];

^d V141A/I147L/I150V/K152Q/K155S/A158E/D158A/A160S/K161R;

^e 1.0 a 2.5 K/min;

^f ND: No determinado;

^g K155S/D158A/A160S/K161R/E205Q

^h *V61Q/M122H/A69T/P81Q/A160S

ⁱ S43P/V46L/A49P/K155S/D158A/K161R/E205Q;

^j V61Q/M122H/A69T/P81Q/A160S/S43P/V46L/A49P/K155S/D158A/K161R/E205Q

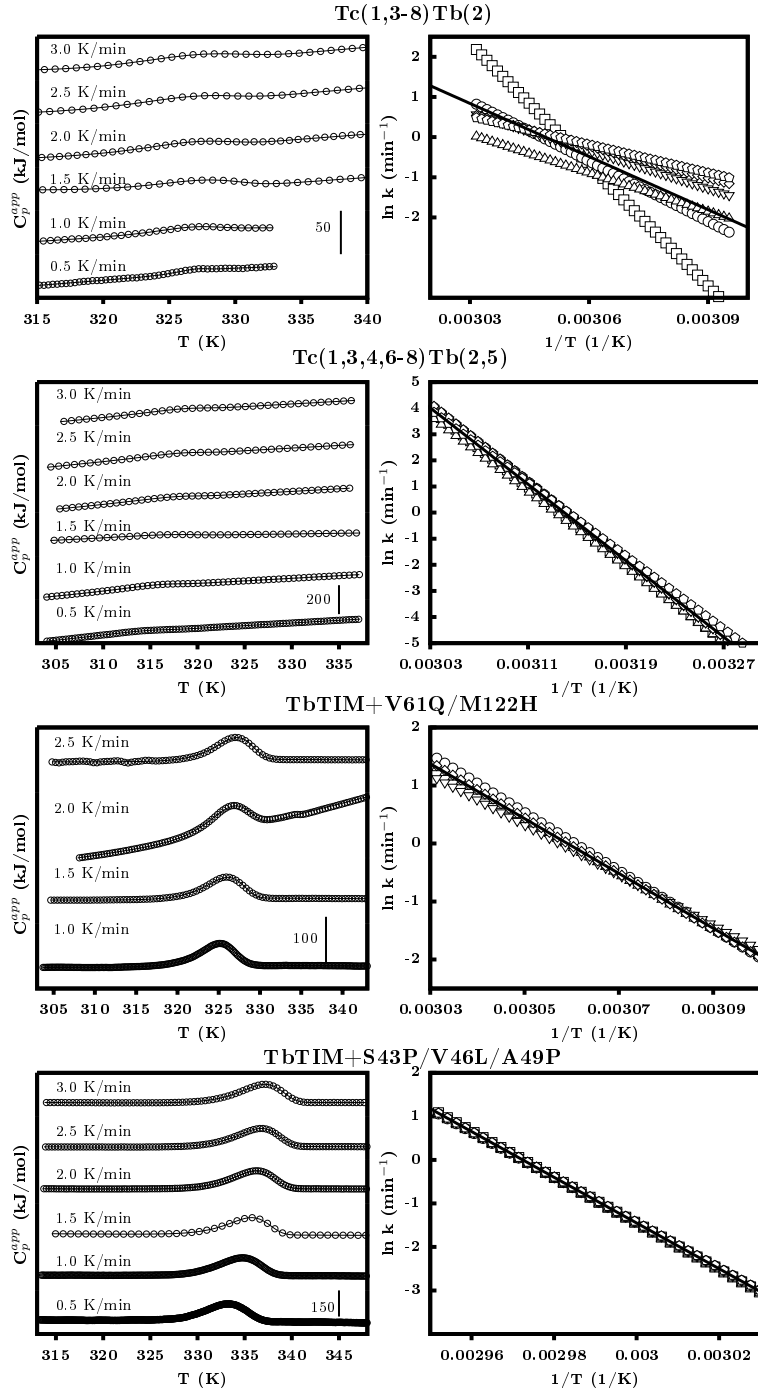


Figura B.1: Se muestran los termogramas de DSC de cuatro quimeras a las velocidades de barrido indicadas (izquierda). La concentración de proteína fue de 0.4 mg/mL en todos los casos. Los símbolos representan la capacidad calorífica aparente y las líneas continuas representan el ajuste al modelo de dos estados irreversible. Los perfiles fueron desplazados en el eje de las ordenadas por efectos de claridad. Se muestran los gráficos de Arrhenius (derecha) para las mismas proteínas. Los símbolos corresponden a las diferentes velocidades de barrido empleadas: 0.5, 1.0, 1.5, 2.0, 2.5 y 3.0 K/min. Las líneas continuas representan el mejor ajuste lineal de todos los puntos.

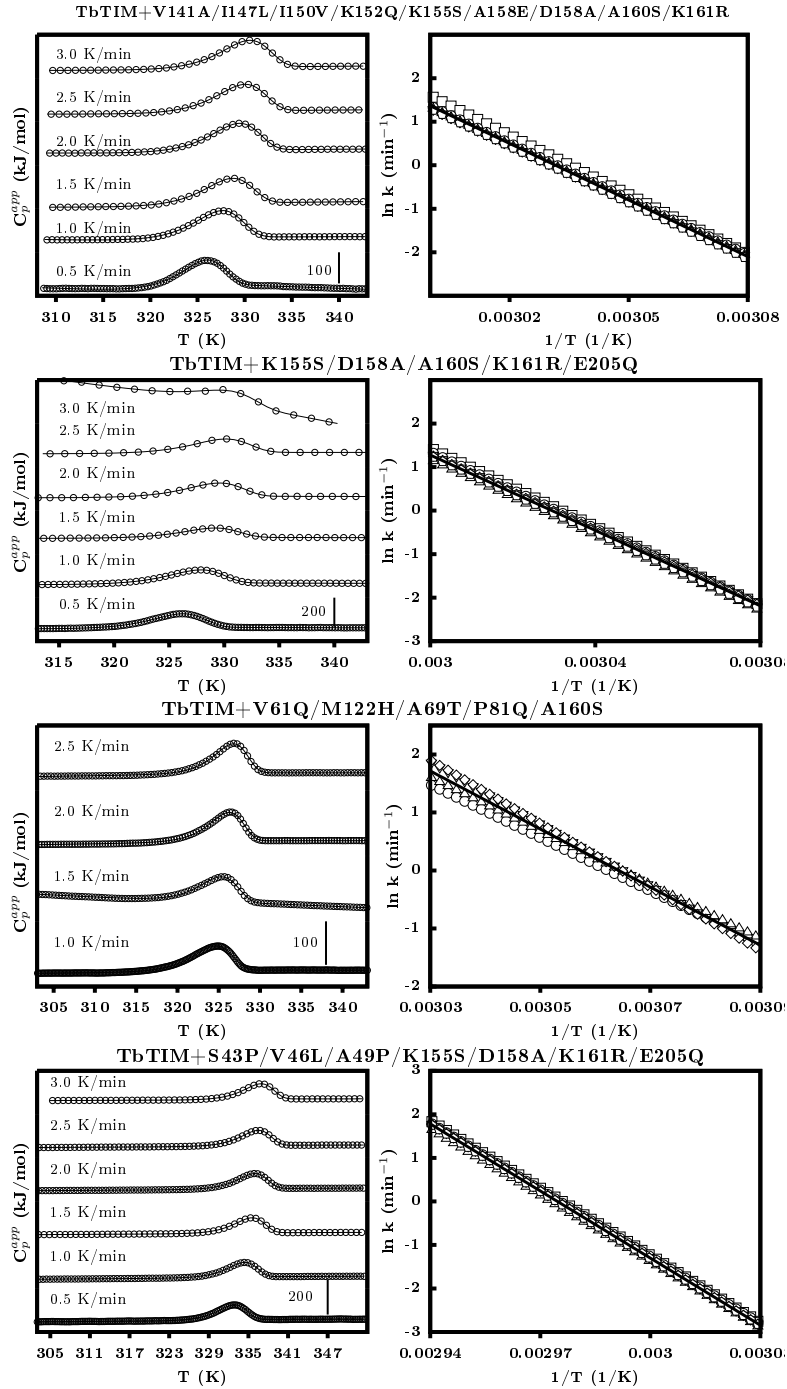


Figura B.2: Se muestran los termogramas de DSC de cuatro quimeras a las velocidades de barrido indicadas (izquierda). La concentración de proteína fue de 0.4 mg/mL en todos los casos. Los símbolos representan la capacidad calorífica aparente y las líneas continuas representan el ajuste al modelo de dos estados irreversible. Los perfiles fueron desplazados en el eje de las ordenadas por efectos de claridad. Se muestran los gráficos de Arrhenius (derecha) para las mismas proteínas. Los símbolos corresponden a las diferentes velocidades de barrido empleadas: 0.5, 1.0, 1.5, 2.0, 2.5 y 3.0 K/min. Las líneas continuas representan el mejor ajuste lineal de todos los puntos.

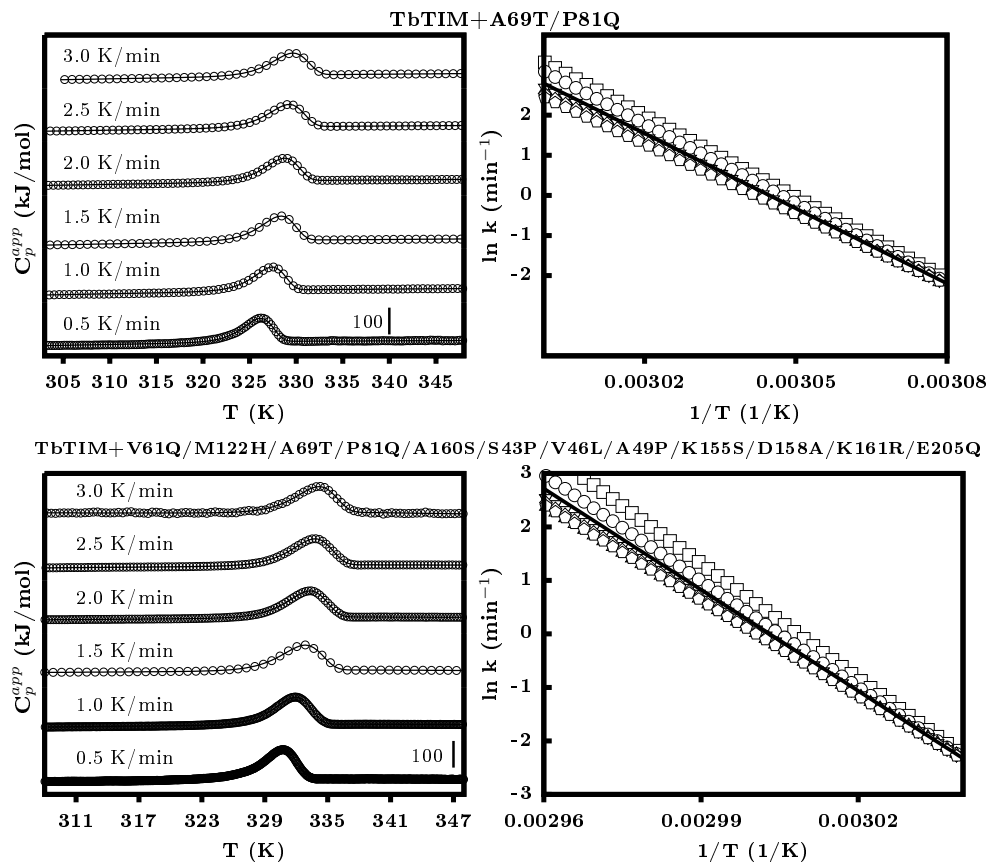


Figura B.3: Se muestran los termogramas de DSC de dos quimeras a las velocidades de barrido indicadas (izquierda). La concentración de proteína fue de 0.4 mg/mL en todos los casos. Los símbolos representan la capacidad calorífica aparente y las líneas continuas representan el ajuste al modelo de dos estados irreversible. Los perfiles fueron desplazados en el eje de las ordenadas por efectos de claridad. Se muestran los gráficos de Arrhenius (derecha) para las mismas proteínas. Los símbolos corresponden a las diferentes velocidades de barrido empleadas: 0.5, 1.0, 1.5, 2.0, 2.5 y 3.0 K/min. Las líneas continuas representan el mejor ajuste lineal de todos los puntos.

Apéndice C

Análisis general de las simulaciones a 298 K de TbTIM y TcTIM

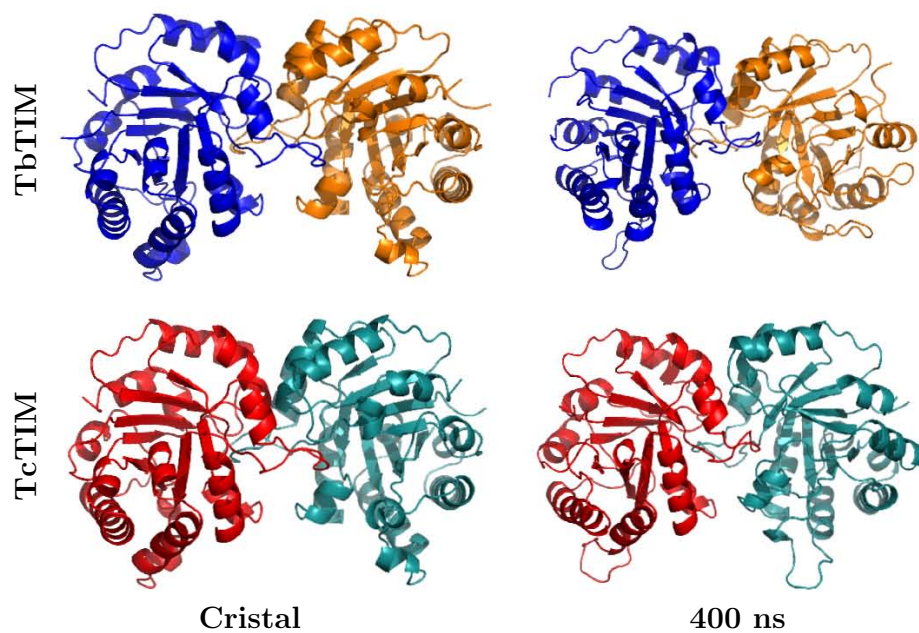


Figura C.1: Imágenes de las estructuras de TbTIM y TcTIM a 400 ns de simulación a 298 K, comparadas con las estructuras cristalinas

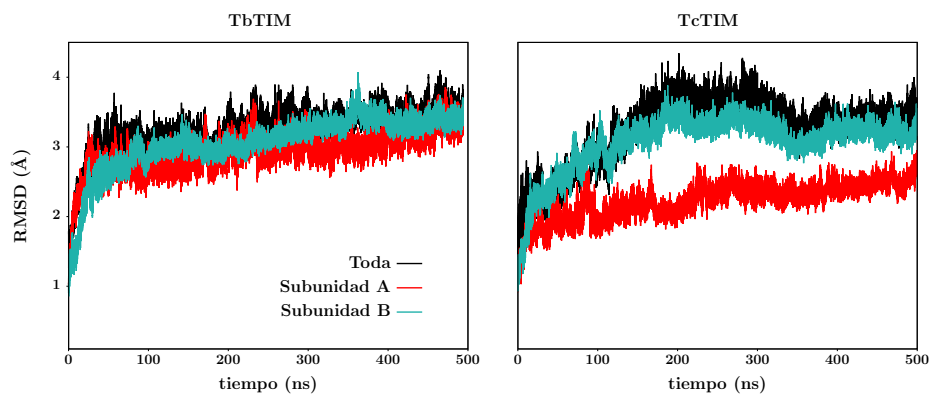


Figura C.2: Desviación cuadrática media de los C- α de las enzimas silvestres TbTIM y TcTIM durante las simulaciones a 298 K

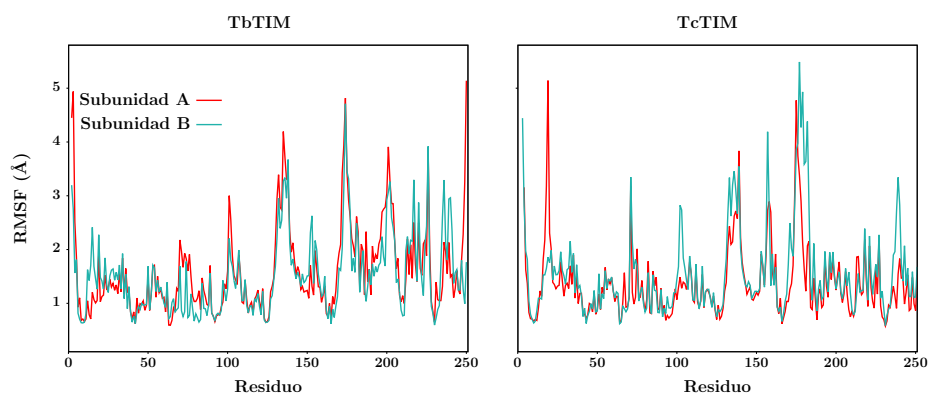


Figura C.3: Fluctuación cuadrática media de los residuos de las enzimas silvestres TbTIM y TcTIM durante las simulaciones a 298 K

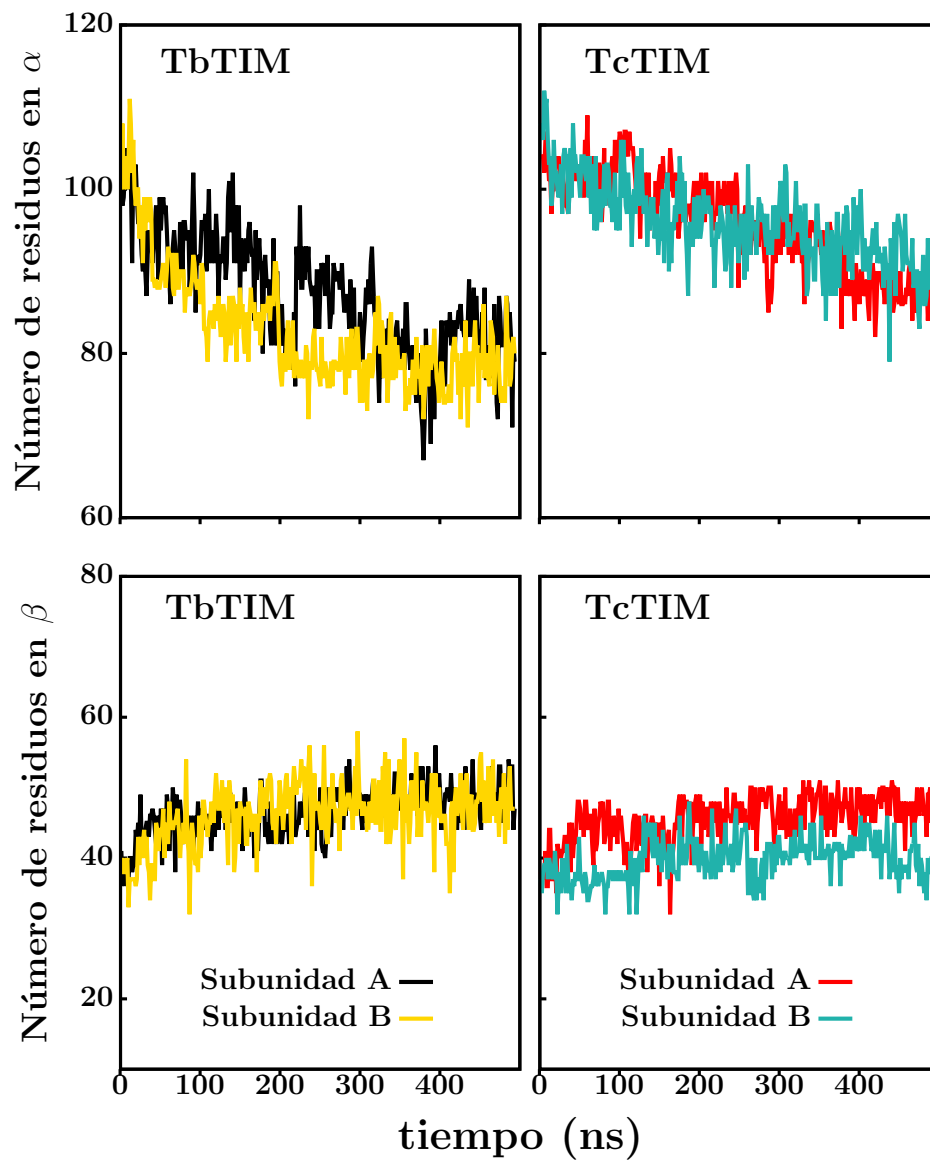


Figura C.4: Estructura secundaria de las enzimas silvestres TbTIM y TcTIM durante las simulaciones a 298 K

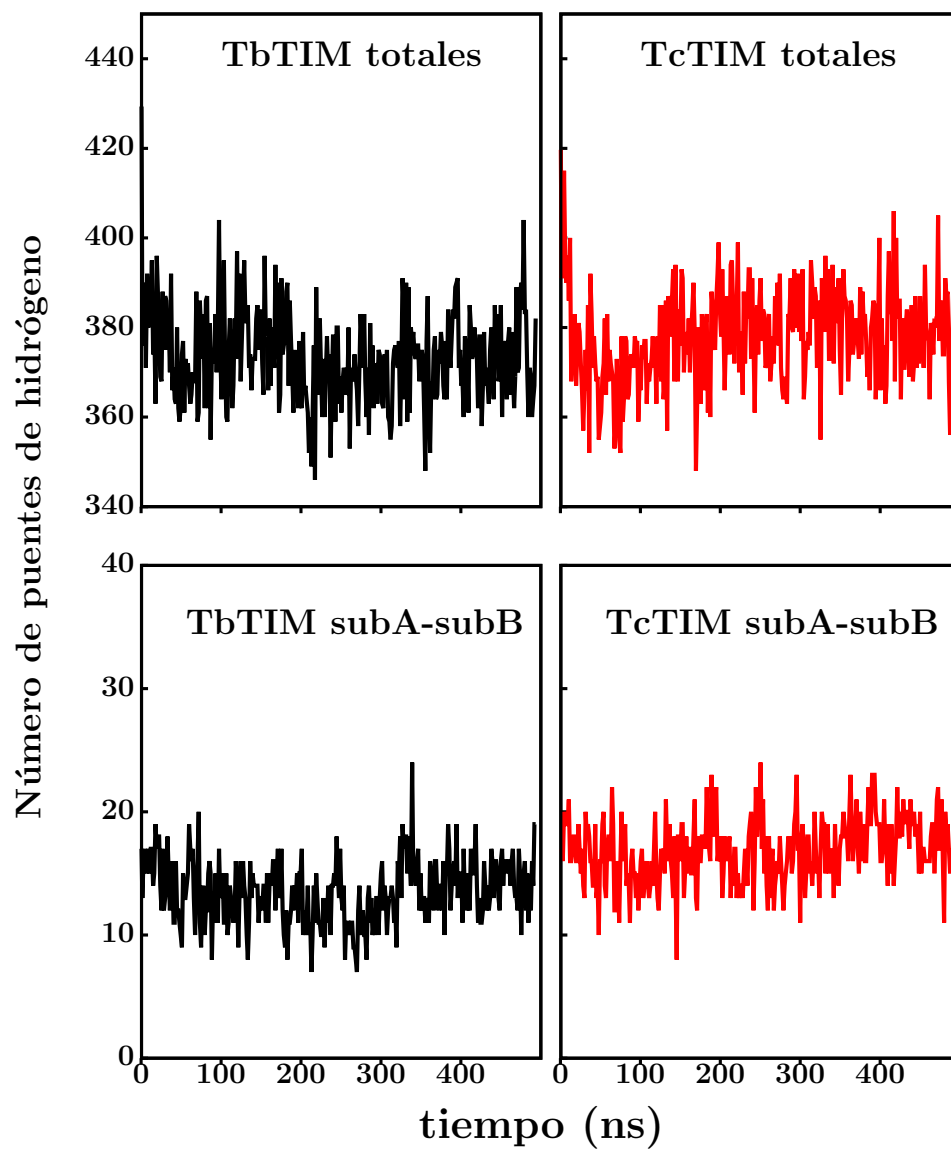


Figura C.5: Puentes de hidrógeno de las enzimas silvestres TbTIM y TcTIM durante las simulaciones a 298 K

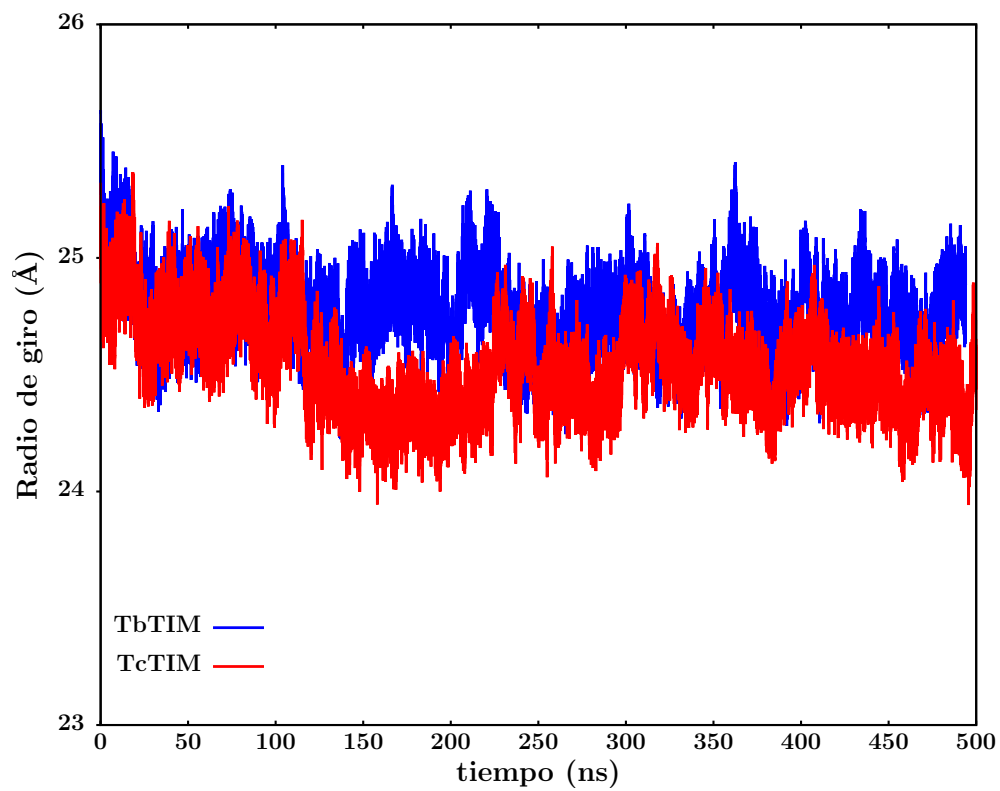


Figura C.6: Radio de giro de las enzimas silvestres TbTIM y TcTIM durante las simulaciones a 298 K

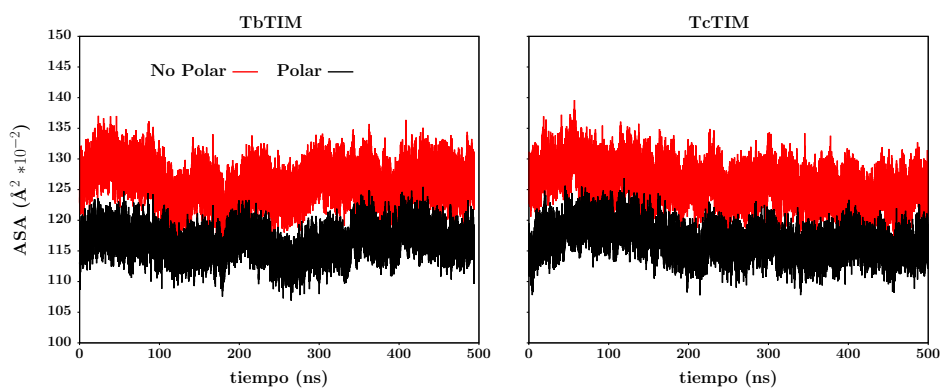


Figura C.7: Área accesible al solvente de las enzimas silvestres TbTIM y TcTIM durante las simulaciones a 298 K

Glosario

ASA Siglas de Accessible Solvent Area (Área accesible al solvente). Tamaño de la superficie de la proteína a la cual tiene acceso el modelo de solvente que se emplea en las simulaciones de dinámica molecular.

CS Siglas de Conformational Space (Espacio conformacional) Se encuentra conformado por los diferentes estados estructurales que puede adoptar una proteína.

DSC Siglas de Differential Scanning Calorimetry (Calorimetría diferencial de barrido). Técnica que permite medir la diferencia en capacidad calorífica entre los líquidos contenidos en dos celdas aisladas. Cuando una de las celdas contiene una muestra de proteína y la otra contiene un amortiguador, la traza de DSC obtenida será el resultado del proceso de desplegamiento de la proteína.

E_{act} Energía de activación. Es la barrera energética que separa el estado nativo del estado de transición (TS). Este parámetro es también una medida de la estabilidad cinética de una proteína, ya que entre más grande sea esta barrera, la probabilidad de cruzarla será menor y el tiempo de vida media del estado nativo será mayor.

Estabilidad cinética Es un concepto que está ligado al tiempo promedio en que una proteína mantiene su estructura y función bajo ciertas condiciones. La estabilidad cinética de una proteína determinará su tiempo de vida media, o sea, el tiempo en el que la proteína permanecerá plegada y funcional.

Flexibilidad de proteínas Está definida por los estados conformacionales que visita la proteína durante las fluctuaciones que sufre en el medio donde se encuentra.

Flexibilidad dinámica Se refiere a la rapidez con la que una proteína puede transitar de una conformación a otra.

Flexibilidad estática Se refiere a la diversidad estructural de conformaciones que se encuentran en equilibrio en el estado nativo de la proteína.

Flexibilidad térmica Es un parámetro acuñado en este trabajo que se define como la medida cuantitativa del incremento del espacio conformacional accesible a la proteína cuando se le suministra energía. Está definido por el valor de *TGD*.

MDS Siglas de Molecular Dynamics Simulations (Simulaciones de dinámica molecular). Herramienta computacional que nos permite estudiar la dinámica de un sistema confinado a una caja virtual de forma y tamaño definidos, bajo ciertas condiciones y durante cierto tiempo.

PMD Siglas de Progressive Moving Dispersion (Dispersión progresiva de movilidad). Se determinó como la distancia promedio entre cada estructura en un grupo de n estructuras y su posición media.

RMSD Siglas de Root Mean Square Deviation (Desviación cuadrática media). Es una medida promedio de la distancia entre los átomos que conforman dos estructuras.

TGD Siglas de Temperature Gradient of the Displacement (Dependencia con la temperatura de los desplazamientos). Es una cuantificación de la flexibilidad térmica de la proteína. Está definida por proteína, por región y por residuo (ver ecuación 2.8)

T_m Temperatura máxima de la transición en una traza de calorimetría diferencial de barrido.

W También llamada $T_{1/2}$, es el ancho de una traza de calorimetría diferencial de barrido a la mitad del área total del pico.