



# UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

## Maestría y Doctorado en Ciencias Bioquímicas

Mecanismos adaptativos en la evolución de la triosafosfato isomerasa

TESIS

QUE PARA OPTAR POR EL GRADO DE:  
Maestro en Ciencias

PRESENTA:  
Ricardo Muñiz Trejo

TUTOR PRINCIPAL

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

MIEMBROS DEL COMITÉ TUTOR

Dr. Arturo Carlos Il Becerra Bracho  
[Facultad de Ciencias](#)

Dr. Lorenzo Patrick Segovia Forcella  
[Instituto de Biotecnología](#)

Ciudad Universitaria, CD. MX. Septiembre, 2020



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

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

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

“La biología evolutiva es un campo notoriamente difuso  
de pesquisa científica”  
-Niles Eldredge, 2008, p. 10

## AGRADECIMIENTOS ACADÉMICOS

A la Universidad Nacional Autónoma de México (UNAM) y al Programa de Maestría y Doctorado en Ciencias Bioquímicas por su ofrecerme un espacio de alta calidad académica en el cuál he alimentado mi pasión por la investigación científica.

A mi tutor principal, el Dr. Daniel Alejandro Fernández Velasco, cuya espléndida asesoría ha sido invaluable para mi formación académica, profesional y personal desde hace casi un lustro.

A mi comité tutor, Dr. Arturo Carlos II Becerra Bracho y Dr. Lorenzo Patrick Segovia Forcella, por su brillante retroalimentación que fue moldeando este proyecto semestre a semestre.

A la M. en C. Blanca Ramos Cerrillo del Departamento de Ingeniería Celular y Biocatálisis del Instituto de Biotecnología de la UNAM por su ayuda para la secuenciación de los genes empleados en este proyecto. A la Dra. Gloria Hernández Alcántara del Laboratorio de Péptidos y Proteínas de la Facultad de Medicina de la UNAM, por la asistencia y capacitación brindada para la subclonación exitosa de estos mismos genes.

A la Dra. Jessica Díaz Salazar y al Dr. Miguel Antonio Costas Basín del Laboratorio de Biofisiología de la Facultad de Química de la UNAM, por el préstamo del equipo MicroCal VP-Capillary DSC y por el suministro de ácido fosfoglicohidroxámico para los experimentos de calorimetría de titulación isotérmica

A las técnicas académicas del Laboratorio de Fisiología e Ingeniería de Proteínas (LFQIP), Dra. María Isabel Velázquez López, M. en C. Laura Iliana Álvarez Añorve y Q. F. B. Eugenia Flores Robledo, por el apoyo técnico brindado durante la realización de este trabajo.

Al Dr. Sergio Romero Romero, por la asesoría y consejo ofrecidos continuamente durante la realización de este trabajo. Al M. en F. C. Alejandro Fábregas Tejeda por el espacio de discusión siempre abierto para temas de biología evolutiva.

A mi jurado de examen, Dra. Rosario Adelaida Muñoz Clares, Dra. Adela Rodríguez Romero, Dr. Enrique García Hernández, Dr. Alfredo Torres Larios y Dr. Diego González Halphen, por verter sus comentarios sobre el texto que sin lugar a duda ayudaron a nutrirlo y mejorarlo

Al Programa de Apoyo a los Estudios del Posgrado (PAEP) por el apoyo económico brindado para la asistencia al congreso *Evolution Evolving: Process, Mechanism and Theory* celebrado en Churchill College, Cambridge, Reino Unido del 1 al 4 de abril de 2019.

Al Consejo Nacional de Ciencia y Tecnología (Conacyt) por la beca recibida durante la maestría con número de CVU 926948. El desarrollo de este proyecto fue financiado por la Facultad de Medicina de la UNAM y por los proyectos PAPIIT IN219913 e IN220516, y Conacyt 254514.



## AGRADECIMIENTOS PERSONALES

A mi madre, Edith Trejo, porque tu apoyo y amor incondicional me han acompañado a donde quiera que vaya. Gracias por tu sentido del humor, pero, sobre todo, por transmitirme tu asertividad que me ha ayudado a abrirme paso por la vida.

A mi hermano, Eduardo Muñiz Trejo, porque con nadie más puedo compartir charlas que me hagan sentir tan humano y completo. Tu mente, perspicaz y curiosa, siempre me alimenta de nuevos saberes. No hay momento junto a ti en el que no crezca como persona.

A mi padre, Eduardo Muñiz, por el soporte incesante que me has brindado desde que tengo memoria. Gracias por tu constancia y entusiasmo.

A mi jefe, Alejandro Fernández, por abrirme las puertas del laboratorio cuando más lo necesitaba, por su amistad y por seguir siendo el ejemplo del tipo de científico que quiero llegar a ser: uno que no deja que su ego obnuble su trabajo ni su forma de tratar a las personas.

A mi (otro) jefe y figura materna académica, Sergio Romero, aunque estás del otro lado del charco nunca te he dejado de sentir cerca, gracias por ser tan espléndido amigo. Tus enseñanzas en la vida y en el laboratorio son la base con la que camino por este mundo de la ciencia.

A mi mejor amigo, Alejandro Fábregas, chico, qué sería de mi vida académica si no tuviera a alguien como tú con el cuál puedo discutir sobre los temas que me apasionan. Más importante aún, qué sería de mi vida en general si no te tuviera como amigo.

A mi hermanito, Mario Morales. Si con alguien puedo compartir los momentos difíciles es contigo, ya sea algún problema general o una pandemia que paralizó el mundo, sé bien que puedo contar contigo para escucharme y darme todo el apoyo.

A mi hermana, Fátima Ávila, tu sentido común y sensatez son unos consejeros infalibles. Ha sido muy grato para mí escucharte y hacerte parte de mi familia (ahora más que nunca). Los maratones de películas contigo, Alejandro y Mario los llevaré siempre conmigo.

A la Dra. Georgina Garza y el Dr. Alejandro Sosa, muchas gracias por su amistad, por las charlas de pasillo y haberse convertido en parte integral de mi desarrollo académico. Les quiero sobremanera.

A mi "alumna", Andrea Hernández, y lo pongo entre comillas porque yo en realidad me siento como tu par. Gracias por dejarme ser parte de tu vida académica y, por supuesto, por el apoyo en la realización de los experimentos durante este último año. Espero haber sido buena madre.

En estos dos años conocí muchas personas del programa de maestría, pero Javier Medina y Willy Delgado son, sin lugar a duda, a los más quiero y respeto. Gracias por todas las risas, los juegos, las horas de estudio y por ser quienes son.

Al LFQIP, que ha sido mi hogar por los últimos seis años. Tengo una deuda de por vida con los miembros de este laboratorio, los que están y los que se han ido, ¡mil gracias por todo! Ustedes se han convertido en parte de mi familia y yo les amo con todo mi ser.

## ÍNDICE

ÍNDICE DE FIGURAS Y TABLAS.....	7
LISTADO DE ABREVIATURAS.....	9
RESUMEN.....	10
INTRODUCCIÓN.....	11
<i>El papel de la evolución en el estudio de la bioquímica de proteínas</i> .....	11
<i>Retos del estudio evolutivo en la ciencia de proteínas</i> .....	11
<i>Las múltiples estrategias adaptativas para alcanzar la termoestabilidad en proteínas</i> .....	15
ANTECEDENTES.....	17
<i>Modelo de estudio: la triosafosfato isomerasa</i> .....	17
<i>La TIM de arqueas</i> .....	18
<i>Caracterización de la TIM del hipertermófilo Methanocaldococcus jannaschii</i> .....	19
<i>Selección de especies</i> .....	20
<i>La adaptación como concepto para organizar un trabajo de investigación</i> .....	22
HIPÓTESIS.....	23
OBJETIVOS.....	23
<i>Objetivo general</i> .....	23
<i>Preguntas particulares</i> .....	23
MÉTODOS.....	24
<i>Subclonación de genes</i> .....	24
<i>Expresión y purificación de las proteínas seleccionadas</i> .....	24
<i>Caracterización espectroscópica</i> .....	25
<i>Estructura secundaria</i> .....	25
<i>Estructura terciaria</i> .....	26
<i>Determinación de parámetros catalíticos</i> .....	27
<i>Determinación del estado oligomérico</i> .....	28
<i>Desnaturalización por temperatura</i> .....	29
<i>Modelo irreversible de dos estados</i> .....	30
<i>Modelo con transiciones independientes de más de dos estados</i> .....	32
<i>Desnaturalización por agentes caotrópicos</i> .....	32
<i>Dicroísmo circular</i> .....	33
<i>Anisotropía</i> .....	33
<i>Actividad catalítica</i> .....	33
<i>Procesamiento de los datos de desplegamiento químico</i> .....	34
<i>Unión a ligando</i> .....	35
<i>Análisis y visualización de datos</i> .....	37
RESULTADOS Y DISCUSIÓN.....	38
<i>Subclonación de los genes de las ArqTIM</i> .....	38
<i>Purificación y caracterización de las MesoTIM</i> .....	40
<i>Purificación y caracterización espectroscópica y funcional de las TermoTIM</i> .....	46
<i>Caracterización de estructuras secundaria y terciaria</i> .....	47

<i>Determinación del estado oligomérico</i> .....	48
<i>Obtención de parámetros catalíticos</i> .....	49
<i>Función, estabilidad y evolución de proteínas termófilas, parte I</i> .....	50
<b>Caracterización de la estabilidad y vías de plegamiento</b> .....	52
<i>Desnaturalización por temperatura</i> .....	52
<i>Denaturalización por agentes químicos</i> .....	63
<i>Función, estabilidad y evolución de proteínas termófilas, parte II</i> .....	74
<b>Caracterización de la unión a ligando</b> .....	77
Función, estabilidad y evolución de proteínas, parte III.....	81
<b>RESUMEN DE RESULTADOS</b> .....	83
<b>CONCLUSIONES</b> .....	85
<b>PERSPECTIVAS</b> .....	88
<b>REFERENCIAS</b> .....	89
<b>MATERIAL SUPLEMENTARIO</b> .....	101

## ÍNDICE DE FIGURAS Y TABLAS

Figura	Tabla	Título	Página
	1	Ejemplos de explicaciones adaptativas en la ciencia de proteínas.	13
1		Topología arquetípica de una triosafosfato isomerasa	17
2		Estructura cristalográfica de la TIM de <i>Methanocaldococcus jannaschii</i>	18
3		Árbol filogenético del orden <i>Methanococcales</i>	20
4		Temperaturas de crecimiento de los organismos seleccionados	21
	2	Datos de interés de los organismos y sus secuencias de la TIM para este trabajo	22
	3	Porcentaje de identidad entre las secuencias elegidas	22
5		Reacción acoplada para medir la actividad de una TIM	27
6		Elementos básicos de una endoterma para una proteína en solución	30
7		Elementos básicos de una isoterma de la unión a ligando	36
8		Digestiones de los plásmidos subclonados	38
	4	Disposición final de los genes utilizados en este proyecto	39
9		Purificación de las MesoTIM	41
10		Caracterización espectroscópica y funcional de las MesoTIM	42
	5	Parámetros espectroscópicos y funcionales de las MesoTIM	43
11		Purificación de las TermoTIM	46
12		Espectroscopía de las TermoTIM	47
	6	Parámetros de los espectros de DC e IF de las TermoTIM	47
13		Cromatografía de exclusión molecular analítica de las TermoTIM	48
	7	Determinación de las propiedades hidrodinámicas por A-SEC	49
14		Cinéticas enzimáticas de las TermoTIM a 25 °C	49
	8	Parámetros catalíticos de las TermoTIM a 25 °C	50
15		Desplegamiento térmico de MjaTIM-85 a distintas velocidades de barrido	53
16		Prueba de reversibilidad por temperatura de MfoTIM-75	54
17		Desplegamiento térmico de MfoTIM-75 a distintas velocidades de barrido	55
18		Métodos para calcular la energía de activación de desplegamiento	56
	9	Parámetros del desplegamiento térmico bajo control cinético	56
19		Desplegamiento térmico de MfoTIM-75 a distintas concentraciones de proteína	60
	10	Parámetros del desplegamiento térmico por un modelo de control termodinámico	61
20		Ajuste de MfoTIM-75 a modelo de desplegamiento de más de dos estados con transiciones independientes	62
21		Efecto de la urea sobre el plegamiento de MfoTIM-75	64
22		Curva de desplegamiento con GdnHCl de MjaTIM-85 con histidinas a 25 °C	65
23		Curvas de desnaturalización de MjaTIM-85 con histidinas a distintas concentraciones de proteína	66
24		Desplegamiento con GdnHCl de MjaTIM-85 a 25 °C	67
	11	Valores de $C_m$ de la desnaturalización de MjaTIM-85 sin histidinas por método empleado	67
25		Curvas de desnaturalización de MjaTIM-85 sin histidinas a dos concentraciones de proteína	68
26		Desplegamiento de MfoTIM-75 con GdnHCl a 25 °C	69
27		Curvas de desplegamiento con GdnHCl a distintas concentraciones de MfoTIM-75	71
28		Ajustes de la curva de desnaturalización de MfoTIM-75 a un modelo de dos estados	72
	12	Parámetros termodinámicos del desplegamiento químico de MfoTIM-75 a 25 °C	72
29		Modelo estructural de MfoTIM-75	73
	13	Cálculos del $\Delta ASA$ (Å) de desnaturalización de MfoTIM-75	73
30		ITC de unión a PGH de las MjaTIM-85 a dos temperaturas	77
	14	Firmas termodinámicas de unión a PGH de distintas TIM a 25 °C	78
31		Firmas termodinámicas de unión a PGH de las TermoTIM	78
32		Firma termodinámica de la MjaTIM-85 con sacarosa	81
33		Relación de las características biofísicas caracterizadas con la filogenia de la TIM de <i>Methanococcales</i>	85

<b>Figura</b>	<b>Tabla</b>	<b>Título</b>	<b>Página</b>
	S1	Lista de pares de residuos que interactúan en la superficie del tetramero	101
S1		Acercamiento a la superficie de tetramerización de MjaTIM-85	102
S2		Árbol filogenético de máxima verosimilitud de la TIM de arqueas	103
	S2	Código de colores del árbol filogenético de la TIM de arqueas	104
	S3	Resumen de características de las especies seleccionadas y secuencias solicitadas	105
S3		Espectros de deconvolución de las ArqTIM	106
	S4	Resultados de las diferentes aproximaciones para el cálculo de la energía de activación en kJ mol <sup>-1</sup>	107
S4		Efecto de los cationes monovalentes sobre la anisotropía de las tirosinas	107
S5		Estabilidad a la dilución de las TermoTIM	108
	S5	Valores de la constante de disociación (KD) observados para la dilución de las TermoTIM	108
S6		Isotermas de unión de MjaTIM-85 con PGH a varias temperaturas	111
	S6	Parámetros de las isotermas de unión de MjaTIM-85 bajo un modelo de unión 1:1	111
S7		Isotermas de unión de MfoTIM-75 con PGH a varias temperaturas	113
	S7	Parámetros de las isotermas de unión de MfoTIM-75 bajo un modelo de unión 1:1	114
S8		Isotermas de unión de MjaTIM-85 con PGH a varias temperaturas en 0.66 molal de sacarosa	116
	S8	Parámetros de las isotermas de unión de MfoTIM-75 con sacarosa bajo un modelo de unión 1:1	116

## LISTADO DE ABREVIATURAS

aa	Aminoácidos
ArqTIM	TIM del orden Methanococcales estudiadas en este trabajo
C <sub>m</sub>	Concentración media de desnaturalización
C <sub>p</sub>	Capacidad calorífica a presión constante. Sufijos: <sup>exc</sup> en exceso, <sup>ap</sup> aparente, <sub>u</sub> de unión.
C <sub>p</sub> TIM	TIM de <i>Clostridium perfringens</i>
DC	Dicroísmo circular
DHAP	Dihidroxiacetona fosfato
DO <sub>600nm</sub>	Densidad óptica a 600 nm
DSC	Calorimetría diferencial de barrido (siglas en inglés)
E <sub>A</sub>	Energía de activación
G	Energía libre de Gibbs. Sufijos: <sub>H2O</sub> en agua, <sub>u</sub> de unión
G3P	Gliceraldehído-3-fosfato
GdnHCl	Clorhidrato de guanidinio
H	Entalpía. Sufijos: <sub>vH</sub> van't Hoff, <sub>cal</sub> calorimétrico, <sub>u</sub> de unión
IF	Fluorescencia intrínseca
IPTG	Isopropil-β-D-1-tiogalactopiranosido
ITC	Calorimetría de titulación isotérmica (siglas en inglés)
K <sub>A</sub>	Constante de asociación
k <sub>cat</sub>	Número de recambio
K <sub>D</sub>	Constante de disociación
K <sub>eq</sub>	Constante de equilibrio
K <sub>M</sub>	Constante de Michaelis-Menten
LB	Caldo de cultivo de Luria Bertani
m	Dependencia del ΔG con la concentración de desnaturalizante
MaeTIM-46	TIM de <i>Methanococcus aeolicus</i> Nankai-3
MesoTIM	TIM de la familia mesófila Methanococcaceae
MfoTIM-75	TIM de <i>Methanoterris formicicus</i> Mc-S70
MjaTIM-85	TIM de <i>Methanocaldococcus jannaschii</i> DSM 2661
MmaTIM-38	TIM de <i>Methanococcus maripaludis</i> C7
MokTIM-60	TIM de <i>Methanothermococcus okinawensis</i> IH1
n	Estequiometría de la reacción en ITC
PGH	Ácido fosfoglicolohidroxámico
R	Constante de los gases ideales (8.314 J mol <sup>-1</sup> K <sup>-1</sup> )
S	Entropía. Sufijos: <sub>u</sub> de unión
SEC	Cromatografía de exclusión molecular (siglas en inglés)
SN	Selección natural
T	Temperatura
TbTIM	TIM de <i>Trypanosoma brucei</i>
TcTIM	TIM de <i>Trypanosoma cruzi</i>
TEA	Trietanolamina
TED	Amortiguador de TEA, EDTA y DTT (se indican las proporciones)
TermoTIM	TIM de la familia termófila Methanocaldococcaceae
TIM	Triosafofosfato isomerasa
T <sub>m</sub>	Temperatura media de desnaturalización. Sufijos: <sup>ap</sup> aparente
T <sub>opt</sub>	Temperatura óptima de crecimiento de un organismo
UV	Ultravioleta
v <sub>0</sub>	Velocidad inicial de una reacción
V <sub>máx</sub>	Velocidad máxima de una reacción.
α-GDH	Enzima α-glicerol-3-fosfato deshidrogenada
Δ	Cambio en
ε	Coefficiente de absortividad molar
θ, [θ]	Elipticidad, Elipticidad promedio por residuo
λ	Longitud de onda

## RESUMEN

En el estudio evolutivo de proteínas es común encontrar explicaciones de tipo adaptativo a distintas características biofísicas de una proteína que no siempre son corroboradas experimentalmente. El cómo cambian las proteínas cuando el organismo habita temperaturas elevadas ha sido sujeto a explicaciones adaptacionistas que no ofrecen claridad con respecto a los mecanismos causales por las que se presentan dichos cambios. Para llegar a una respuesta hacen falta estudios que contemplen analizar a detalle la naturaleza adaptativa de los cambios para determinar si verdaderamente tienen este estatus. En este trabajo se utiliza el método comparativo filogenético para determinar de manera heurística si las características biofísicas y bioquímicas de la triosafosfato isomerasa (TIM) de cinco arqueas del orden Methanococcales son adaptaciones a la vida a temperaturas elevadas. Debido a problemas en la expresión soluble y purificación de las secuencias de TIM de los mesófilos *Methanococcus maripaludis* y *Methanococcus aeolicus*, así como del termófilo *Methanothermococcus okinawensis*, la caracterización no pudo ser concluida exitosamente, pero datos preliminares sugieren que su eficiencia catalítica ronda los  $10^5 \text{ M}^{-1} \text{ s}^{-1}$ , valor inusualmente bajo para esta familia de enzimas. Por otra parte, las TIM de los termófilos *Methanotorrus formicicus* y *Methanocaldococcus jannaschii* son enzimas tetraméricas con una eficiencia catalítica superior a  $10^8 \text{ M}^{-1} \text{ s}^{-1}$  a 25 °C. Además, presentan una firma endotérmica de unión a ligando a temperaturas menores a 20 °C. Respectivamente, su termoestabilidad medida por la temperatura media de desnaturalización ( $T_m$ ) es de 91 y 111 °C. Ambas enzimas presentan irreversibilidad ante el desplegamiento térmico y son resistentes al desplegamiento químico por urea. También presentan diferencias notables en la energía de activación del desplegamiento térmico y en sus vías de desplegamiento. Desde el punto de vista adaptativo, la  $T_m$  y la firma termodinámica de unión a ligando parecen ser las únicas dos características que parecen ser resultado de selección natural. Al respecto de las diferencias en actividad, ni la temperatura de crecimiento del organismo, ni el metabolismo de estas arqueas metanogénicas parecen ser variables ambientales involucradas en su proceso evolutivo. Para el resto de las características descritas, se propone que la variación observada es consecuencia de la robustez de la actividad y, por tanto, se presentan de manera neutral ante la selección natural.

## INTRODUCCIÓN

### ***El papel de la evolución en el estudio de la bioquímica de proteínas***

Una de las preguntas que ha movido la curiosidad de generaciones de bioquímicas y bioquímicos desde hace más de medio siglo es *¿cómo es que la estructura nativa tridimensional de una proteína está determinada por las propiedades fisicoquímicas codificadas en la secuencia unidimensional de aminoácidos?* (Dill & MacCallum, 2012). Esta pregunta, también conocida como “el problema del plegamiento de las proteínas”, tiene sus raíces históricas en los trabajos del grupo de Christian Anfinsen (1961) sobre la capacidad de replegamiento espontáneo de la ribonucleasa bovina y se ha consolidado como uno de los campos de investigación más importantes para la ciencia de proteínas hasta nuestros días (Dill & MacCallum, 2012; Finkelstein, 2018; Fleming, 2017). De manera tradicional, la respuesta a dicha pregunta se puede buscar a través de la exploración experimental, mutación tras mutación, del espacio de secuencia de proteínas.<sup>1</sup> El problema de esta estrategia es que el espacio de secuencias es tan vasto que resultaría virtualmente imposible abarcarlo en su totalidad. Por ello, algunos grupos de trabajo se han armado de una perspectiva evolutiva para atender este problema. En este contexto, la enorme diversidad de estructuras y funciones de las proteínas actuales puede interpretarse como el resultado de un “estudio” masivo que se ha llevado a cabo de manera sostenida por millones de años y cuyo resultado se observa en los procesos de diversificación y optimización. Así, el estudio de proteínas nos ayuda a evaluar el papel que las fuerzas evolutivas (*i. e.* deriva génica, mutación, flujo génico, etc.) tienen sobre las propiedades físicas de estos complejos moleculares y responde cómo los cambios en la secuencia de aminoácidos moldearon una función o una propiedad específica (Harms & Thornton, 2013).

### ***Retos del estudio evolutivo en la ciencia de proteínas***

Como suele suceder en muchas otras áreas de la Biología, la evolución en la Bioquímica es tratada indistintamente como un equivalente de selección natural (SN). Gould & Lewontin en su artículo seminal de 1979, *The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme*, exponen esta tendencia que ha dominado la Biología desde mediados del siglo pasado, la cual consiste en interpretar a la SN como el agente causal más importante en toda forma, función o comportamiento orgánico, o lo que ellos denominaban el *programa adaptacionista*. Esencialmente, este pensamiento ignora deliberadamente o califica de poco importantes a otros procesos evolutivos en su capacidad de generar fenotipos complejos.

---

<sup>1</sup> Este año (2020) se cumplieron 50 años de la carta del prolífico biólogo evolutivo teórico, John Maynard Smith, enviada a la revista *Nature* titulada “Natural Selection and the Concept of Protein Space” en 1970, donde se acuñó la metáfora del “espacio de proteínas” (ver el homenaje que Ogbunugafor (2020) le rindió recientemente).



Es importante recordar que, al tratar con resultados históricos, la información que nos ofrecen los estudios evolutivos siempre estará incompleta y habrá cierto aspecto especulativo en las interpretaciones que se hacen al respecto. Sin embargo, limitarse a explicar las características de un organismo sólo como productos de la selección natural (*i. e.* adaptaciones) implica perderse de la información que otras fuerzas evolutivas ofrecen para trazar con mayor precisión el origen de la complejidad de los seres vivos (Lynch, 2007). Aplicado al problema del plegamiento, el *adaptacionismo* nos limita la comprensión acerca de cómo se desarrollan históricamente las propiedades fisicoquímicas que gobiernan a las proteínas.

Los estudios adaptacionistas, dicen Gould & Lewontin (1979), funcionan a través de “atomizar” las características de un organismo, es decir, tratan a los caracteres como partes individuales aisladas que son sujetas de *optimización* por SN para llevar a cabo una función concreta. Este programa trata a los seres vivos como una mera colección de objetos discretos y no como entidades integradas. Lewontin (1978) expresa de una manera muy elegante el problema de la atomización de las partes:

Is the leg a unit in evolution, so that the adaptive function of the leg can be inferred?  
If so, what about a part of the leg, say the foot, or a single toe, or one bone of a toe?  
(p. 217)

Un ejemplo clásico de “atomizar” el fenotipo se presenta en el desarrollo del mentón en humanos. Respecto a los de los simios, los cráneos humanos se caracterizan por tener un desarrollo neoténico, es decir, un cráneo de un humano adulto asemeja más en tamaño al de un simio recién nacido en referencia a su tamaño corporal. Así, el mentón, tan pronunciado en humanos e inexistente en simios, parece ser una estructura evolutiva con una función particular y algunas explicaciones adaptativas se han dado sobre su origen —que es producto de la selección sexual, por ejemplo—. No obstante, el mentón no se origina como un carácter independiente, sino que es resultado del desarrollo diferido de los huesos mandibulares. La mandíbula inferior consta de dos estructuras: la cresta ósea y la cresta alveolar. En humanos, ambas estructuras son neoténicas y se observa que se han encogido durante la filogenia, pero la cresta alveolar parece haberse encogido más rápido que la cresta dental, por lo que esta diferencia de tamaños provoca, en apariencia, el desarrollo de una nueva estructura: el mentón. En otras palabras, el mentón no es una “parte” del cráneo humano en el sentido evolutivo, sino una consecuencia de la distribución espacial de procesos del desarrollo, por lo que, al menos en cuanto origen respecta, no cumple ninguna función adaptativa (Lewontin, 1978).

De manera análoga, en los estudios de proteínas se suele descomponer al fenotipo en sus características biofísicas o en sus propiedades fisicoquímicas, “atomizándolo”, y su origen se interpreta

sólo en términos adaptativos. En la Tabla 1 se presenta una lista no exhaustiva de ejemplos en los que se ha apelado a una explicación adaptativa para dar razón de una característica biofísica particular de una proteína.

**Tabla 1. Ejemplos de explicaciones adaptativas en la ciencia de proteínas.**

Característica de la proteína	Explicación adaptativa	Referencia
La cinética de plegamiento	“Folding kinetics can be subject to natural selection” (p. 5).	Sikosek & Chan, 2014
La presencia de ciertos dominios.	“Given the extent of domain shuffling evolution, the limited repertoire of domain combinations present in genomes might imply that all combinations have undergone strong evolutionary selection” (p.322).	Han <i>et al.</i> , 2007
Topología de los paisajes de plegamiento y cooperatividad	“[...] the smooth energy landscapes and highly cooperative folding transitions observed for small naturally occurring proteins may also reflect the workings of natural selection” (p.573).	Scalley-Kim & Baker, 2004
Las dimensiones de un núcleo hidrofóbico en termófilos	“An increased hydrophobic core is a common adaptation feature in thermophilic proteins” (p. 159).	Brininger <i>et al.</i> , 2018
La energía de activación del desplegamiento térmico	“Given that natural selection has shaped these two enzymes with highly different energetic barriers for irreversible denaturation [...] (p. 168).	Quezada <i>et al.</i> , 2017
Estructura del estado de transición	“We speculate that the biological advantage of the generally observed diffuse transition state structure is to optimize folding cooperativity [...]” (p. 822).	Lindberg, Tångrot, & Oliveberg, 2002

Más allá de “atomizar” los fenotipos en unidades discretas, el problema principal con este tipo de estudios es que muchas veces caen en una *just-so story*. En biología evolutiva, el término *just-so story* se utiliza para describir una explicación evolutiva para la cual hay poca o nula evidencia. Una de sus características principales es que son explicaciones *ex post facto* que se eligen deliberadamente para ser siempre consistentes con las observaciones que se estén haciendo (Smith, 2016). En el caso de las *just-so stories* adaptacionistas, la presencia de un carácter determinado en un organismo se utiliza como la única evidencia para inferir que en la población ancestral debió haber variación heredable (no observada) y un régimen selectivo (no observado) que favoreció a esa variante del carácter en particular (Olson & Arroyo-Santos, 2015).

Para ilustrar este último punto, tomemos como ejemplo uno de los estudios presentados en la Tabla 1. En el trabajo de Lindberg *et al.* (2002), los autores analizan el efecto que permutaciones circulares de la proteína ribosomal S6 tienen sobre la estructura del estado de transición. Sin mutaciones, la S6 tiene un estado de transición con “estructura difusa típica” para una proteína con desplegamiento químico de dos estados, es decir, se conservan interacciones en sitios múltiples a lo largo de toda la cadena aminoacídica. Sin embargo, cuando la secuencia es permutada circularmente, los contactos del estado

de transición presentan una “nucleación polarizada atípica”, en otras palabras, las interacciones entre aminoácidos se conservan entre regiones particulares, mas no así para toda la cadena. Tanto la S6 nativa como las variantes permutadas alcanzan el estado nativo, pero su cinética de plegamiento es diferente. Como en otra proteína se observó que el estado de transición nucleado parece asociarse con agregación *in vitro*, los autores afirman que la estructura difusa debe ser una adaptación específica para optimizar el plegamiento. Esta aserción se da fuera de un contexto poblacional que dé nociones sobre el efecto de este fenotipo sobre la adecuación del organismo y asumiendo que las permutaciones reflejan de algún modo la variación que existió en las formas ancestrales de esta proteína; sin embargo, ya se le impregnó una carga histórica implícita a este carácter al afirmar que es producto de una ventaja adaptativa.

Con el ejemplo anterior no se pretende decir que estos estudios no sean valiosos, ni que la SN no pueda operar a ese nivel. Lo que se pretende advertir es que generar evidencia de un proceso adaptativo no es tarea sencilla y que se requieren pruebas explícitas para ello. De manera general, existen tres métodos básicos que se utilizan en la biología evolutiva para demostrar que un carácter se originó por SN: 1) el *método comparativo filogenético*, el cual analiza la convergencia de un fenotipo de manera cuantitativa a través de métodos estadísticos, 2) el *método poblacional*, que estudia la adecuación (*fitness*) de manera directa a través de las diferencias asociadas con la variación heredable entre los caracteres de una especie,<sup>2</sup> y 3) el *método de optimalidad*, en el cual se hacen predicciones sobre la configuración de máxima adecuación de un carácter dados ciertos criterios biofísicos establecidos por el grupo de investigación a través de un modelo específico (para descripciones más detalladas sobre estos métodos, véase Olson & Arroyo-Santos, 2015; Olson & Rosell, 2015). Cabe mencionar que lo ideal debería ser probar las hipótesis adaptativas por los tres métodos, ya que todas tienen sus flaquezas, aunque en la práctica esto no siempre es posible.

En la literatura sobre evolución molecular hay numerosos ejemplos donde se emplean algunos de estos métodos para corroborar la acción de la SN sobre ciertos rasgos de una proteína. Un buen ejemplo es el trabajo realizado por Bershtein y colaboradores (2006) en el que se correlaciona la adecuación de *E. coli* con un umbral en la estabilidad termodinámica de la  $\beta$ -lactamasa TEM-1 a partir de una serie de mutantes que compiten en presencia de ampicilina. Sin embargo, el propósito de este trabajo es exhibir que la falta de una perspectiva evolutiva plural ha limitado algunas de las

---

<sup>2</sup> Como muchos otros conceptos en biología evolutiva, la noción de adecuación se ha discutido ampliamente entre biólogos teóricos y filósofos de la ciencia (ver Rosenberg & Bouchard (2020) para una discusión extendida del concepto). Para propósitos de este trabajo, la adecuación se define como la propensión mensurable de un individuo de dejar descendencia y que surge de sus diferencias biológicas intrínsecas con otros individuos de la misma población (Wagner, 2010).

explicaciones que se hacen sobre el origen evolutivo de algunas características biofísicas de las proteínas.

### ***Las múltiples estrategias adaptativas para alcanzar la termoestabilidad en proteínas***

Específicamente nos interesa el problema de la adquisición de la termoestabilidad en proteínas de organismos termófilos e hipertermófilos.<sup>3</sup> En la literatura se ha descrito ampliamente que las proteínas de organismos termófilos tienen una estructura tridimensional muy similar a las de sus homólogos mesófilos, pero detentan una estabilidad térmica elevada. Su termoestabilidad ha sido sujeta a múltiples estudios teóricos y experimentales, los cuales indican que existen una multiplicidad de estrategias “adaptativas” por las cuales las proteínas termófilas adquieran una termoestabilidad tan elevada (Jaenicke & Sterner, 2013; Sadeghi *et al*, 2006). Esta propiedad se ha atribuido a cambios en la composición de aminoácidos de la secuencia y en la longitud de estructuras secundarias (asas y hélices); a la reducción en el número y volumen de cavidades; a la relación entre el área polar/no polar expuesta y oculta (v. gr. cambios en el estado oligomérico) y al incremento en el número de puentes de hidrógeno y salinos; entre otras cosas (Razvi & Scholtz, 2006; Sadeghi *et al.*, 2006; Kumar & Nussinov, 2001; Szilágyi & Závodszy, 2000). Curiosamente, ninguna de estas estrategias se presenta de manera consistente entre las distintas familias de proteínas e incluso llegan a ser diferentes entre los miembros de una misma familia (Szilágyi & Závodszy, 2000). Por ello, pese a que existe un conocimiento fisicoquímico muy sólido al respecto de *cómo* incrementar la termoestabilidad de las proteínas, no es del todo claro el *porqué* unas estrategias se presentan en tal o cual familia de proteínas y en otras no.

Para complementar este último punto, es importante recordar la distinción que el biólogo evolutivo Ernst Mayr (1961) hizo sobre las causas próximas y las causas últimas en la biología. *Grosso modo*, las causas próximas son aquellas que se relacionan con la operación e interacción entre los elementos estructurales de un organismo y responden al “¿cómo?” (¿cómo es que funciona? ¿cómo es que opera?). Las causas últimas, por otra parte, son aquellas que atienden la historia de una estructura o de una función y responden al “¿por qué?” (¿por qué es así? ¿por qué se presenta de esta manera?). Desde esta perspectiva, la termoestabilidad se ha descrito ampliamente desde sus causas próximas, pero no desde sus causas últimas. La razón de esto, sostenemos en este trabajo, es que muchas de las estrategias para incrementar la termoestabilidad se presentan como adaptativas, pero no se corroboran experimentalmente, lo cual ha limitado la comprensión acerca de sus raíces históricas y deja espacios abiertos para la interpretación de la causalidad en el origen de la termoestabilidad.

---

<sup>3</sup> Según su temperatura óptima de crecimiento, los organismos se pueden clasificar en *hipertermófilos* (>80°C), *termófilos* (60-80 °C), *mesófilos* (15-60 °C) y *psicrófilos* (<15 °C) (Rothschild & Mancinelli, 2001).

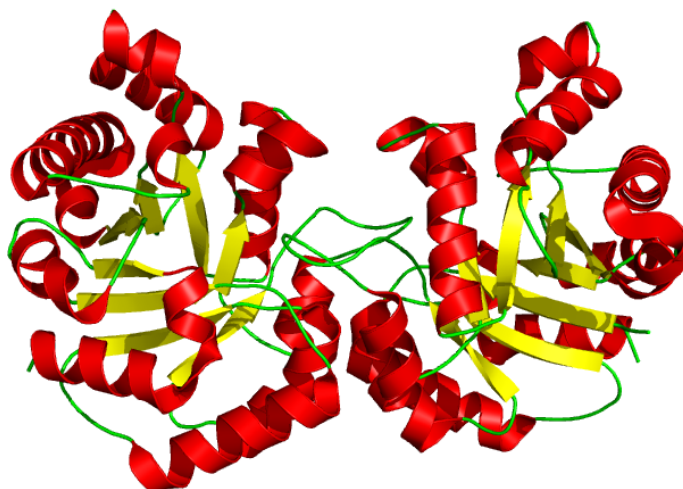
Además, creemos que dirigir experimentos para entender el origen evolutivo de la termoestabilidad nos puede ayudar reconocer también si el origen de otras características biofísicas es adaptativo o no, y nos puede dar información sobre si es válido interpretarlas como caracteres independientes sujetos a fuerzas evolutivas.

Bajo esta premisa, este proyecto busca implementar una perspectiva contemporánea de la teoría evolutiva aplicada al estudio de la bioquímica de proteínas con el propósito de explorar las relaciones entre distintas características biofísicas y bioquímicas de una proteína, como lo son la termoestabilidad, la estabilidad conformacional, la actividad, etc., con la temperatura de crecimiento del organismo como variable ambiental. Lo anterior se realizará por el método comparativo filogenético, el cual nos permite determinar dónde surgieron las diferencias entre las características durante la filogenia. Para este propósito, se decidió emplear a la triosafosfato isomerasa de organismos del dominio Archaea como modelo de estudio.

## ANTECEDENTES

### **Modelo de estudio: la triosafosfato isomerasa.**

La triosafosfato isomerasa (también llamada TIM, EC 5.3.1.1) es una enzima que participa en la glucólisis, gluconeogénesis, síntesis de triglicéridos y en el ciclo de Calvin. Cataliza la interconversión reversible del D-gliceraldehído-3-fosfato (G3P) a la dihidroxiacetona fosfato (DHAP) y es considerada, de manera histórica, una enzima “catalíticamente perfecta” (Albery & Knowles, 1976; Peimbert *et al.*, 2008). En la naturaleza, la TIM siempre se encuentra como oligómero, principalmente como *homodímero*, conformado por monómeros de 250 residuos de longitud en promedio (Figura 1). Por su participación en la vía glicolítica, esta enzima se encuentra conservada en la gran mayoría de los seres vivos, aunque los porcentajes de identidad entre las secuencias de distintas especies pueden ser tan elevados como 98.4% entre la TIM de humano y la de conejo, o tan bajos como 12% entre la TIM del parásito *Plasmodium falciparum* y la de la arquea *Pyrococcus furiosus* (Peimbert *et al.*, 2008).



**Figura 1. Topología arquetípica de una triosafosfato isomerasa.** Estructura tridimensional de una triosafosfato isomerasa dimérica. Hélices  $\alpha$  en rojo, hebras  $\beta$  en amarillo, asas en verde (código PDB: 4Y90).

En 2011, Silverman y colaboradores (2001) mutaron 182 posiciones estructurales de los 248 residuos de la TIM de levadura por uno de siete residuos (FVLAKEQ) de un “alfabeto simplificado” de proteínas.<sup>4</sup> De estas mutantes, 142 exhibían una actividad catalítica similar a la nativa. Además de las diferencias presentes entre las secuencias de TIM naturales, este experimento soporta la idea de que la TIM es una proteína mutacionalmente robusta, es decir, que el fenotipo de la molécula no resiste sin cambios importantes las mutaciones en la secuencia. Debido a su origen evolutivo común, a su robustez

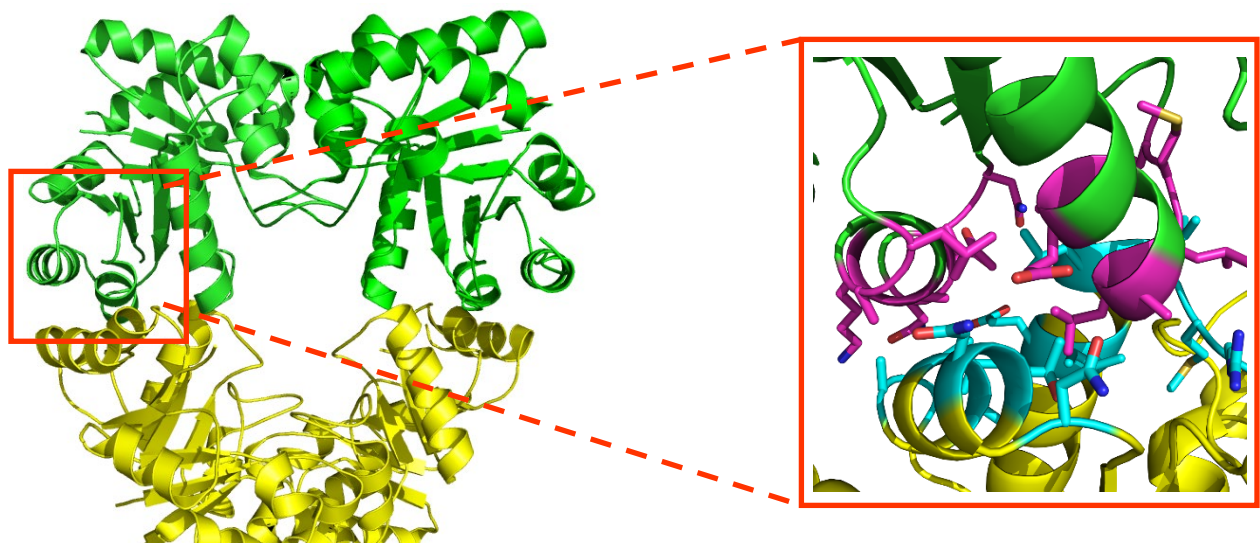
---

<sup>4</sup> Los autores retoman la noción de un “alfabeto simplificado” de una revisión de Plaxco *et al.* (1998) donde se discute el diseño de proteínas estables y topológicamente complejas construidas con el menor número de aminoácidos diferentes posible.

mutacional y a su ubiquidad en los tres dominios de la vida, la TIM representa una enzima de gran interés evolutivo.

### **La TIM de arqueas**

Por su ubiquidad, la TIM es buen modelo para estudiar los mecanismos moleculares por los cuales las proteínas se adaptan a temperaturas elevadas. Por ello, estudiar las TIM del dominio Archaea pueden ser muy informativo, ya que es en este clado donde se encuentran los organismos que soportan las temperaturas más elevadas (Battistuzzi & Hedges, 2009). Las TIM provenientes de arqueas presentan la particularidad de ser secuencias más cortas por casi 20 aminoácidos; estas diferencias se observan como acortamientos de las hélices  $\alpha 6$  y  $\alpha 7$  de las TIM de bacterias y eucariontes (Bräsen *et al.*, 2014). Esta característica hace que las TIM de arqueas sean más compactas en estructura y se separen en un grupo monofilético bien definido cuando se determina su filogenia (Ronimus & Morgan, 2003). Por otro lado, los estudios de caracterización de TIM de arqueas son relativamente escasos, se tiene información sobre ocho especies en total, aunque no se profundiza en la caracterización en muchos de ellos. La arquea mesófila *Methanobacterium bryantii* (Kohlhoff, Dahm, & Hensel, 1996) y la psicrófila *Methanococcoides burtonii* (Dhaunta *et al.*, 2013) tienen TIMs diméricas; sin embargo, la TIM de la mayoría de las arqueas termófilas o hipertermófilas se asocia en *homotetrámeros* a partir de la dimerización de dímeros, como la TIM de *Methanocaldococcus jannaschii* (Figura 2).



**Figura 2. Estructura cristalográfica de la TIM de *Methanocaldococcus jannaschii*.** Los tetrámeros de la TIM de arqueas se asocian formando un dímero de dímeros, mostrados en la figura mediante distintos colores. Se hace un acercamiento a la superficie de tetramerización, donde la asociación entre subunidades está dada por puentes de hidrógeno y enlaces de van der Waals (Gayathri *et al.* 2007, código PDB: 2H6R). Ver detalles en el material suplementario A.

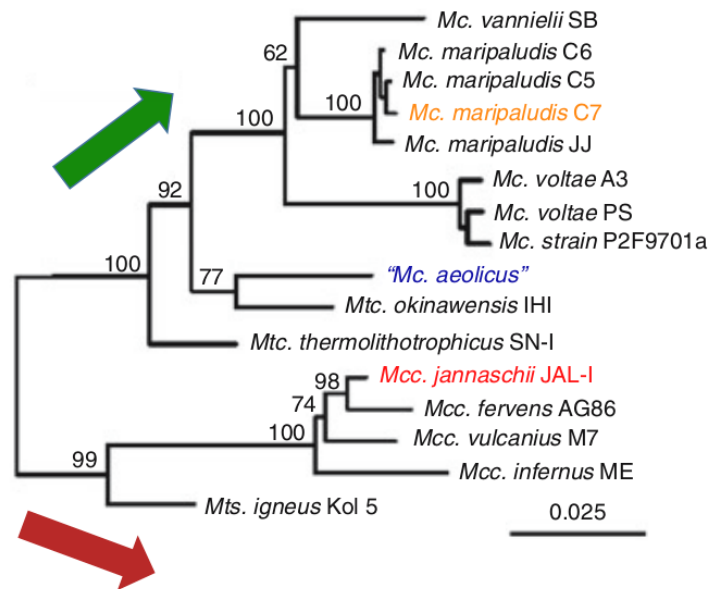
Este cambio en el estado de asociación se ha sugerido como el mecanismo *adaptativo* por el cual las TIM de arqueas se mantienen estables a temperaturas elevadas (Walden *et al.*, 2001, 2004). Esta

hipótesis se formuló sin realizar pruebas explícitas sobre su veracidad. Evidencia posterior demostró que la TIM de *Thermoproteus tenax* no es un homotetrámero estricto, sino que se encuentra en un equilibrio dímero inactivo-tetrámero activo, siendo ambas formas estables a altas temperaturas (Walden *et al.*, 2004). Por otro lado, la TIM de *Thermoplasma acidophilum* es un dímero obligado cuya estabilidad parece estar dada por una estructura compacta con asas reducidas (Park *et al.*, 2016). Por estas razones, la hipótesis de la tetramerización como único mecanismo adaptativo de las TIM de arqueas para alcanzar la termoestabilidad se descartó y no se tiene una explicación histórica del origen de este rasgo hasta la fecha.

### **Caracterización de la TIM del hipertermófilo *Methanocaldococcus jannaschii***

En el trabajo realizado en la licenciatura (Muñiz Trejo, 2017), decidimos explorar con mayor detalle la evolución del estado oligomérico de la TIM en arqueas. Con este objetivo se eligió al orden Methanococcales como grupo de estudio para ceñir el trabajo a un contexto filogenéticamente relevante (durante ese trabajo generamos un árbol filogenético de la TIM de arqueas, el cual se actualizó para el proyecto presente, ver Figura S2). Methanococcales comprende a un grupo de arqueas marinas, anaerobias obligadas y productoras de metano que requieren sales de mar para su óptimo crecimiento y que utilizan dióxido de carbono como aceptor de electrones (Whitman & Jeanthon, 2006). La razón principal por la que se eligió este taxón para trabajar es que está dividido en dos familias: Methanocaldococcaceae, compuesta por géneros de arqueas hipertermófilas; y Methanococcaceae, la cual tiene asignados géneros tanto mesófilos como termófilos moderados (Whitman & Jeanthon, 2006; Figura 3).





**Figura 3. Árbol filogenético del orden Methanococcales.** La flecha verde representa al linaje mesófilo de Methanococcaceae y la flecha roja al linaje hipertermófilo de Methanocaldococcaceae. Las especies indicadas con color fueron las seleccionadas para el trabajo de licenciatura (Modificado de Whitman & Jeanthon, 2006).

Además, como se indicó antes, ya se conoce la estructura cristalográfica de la TIM tetramérica de la arquea hipertermófila *Methanocaldococcus jannaschii* (Gayathri *et al.*, 2007, ver Figura 2), perteneciente de este grupo, por lo que ya había algunas referencias estructurales de las cuales podíamos partir para la caracterización.

En ese estudio previo, decidimos trabajar con las secuencias de los mesófilos *Methanococcus aeolicus* Nankai-3 y *Methanococcus maripaludis* C7, y del hipertermófilo *Methanocaldococcus jannaschii* DSM 2661. Sin embargo, las dos TIM mesófilas no pudieron expresarse de manera soluble, por lo que sólo pudimos trabajar con la última. Esta enzima sí se expresó de manera soluble y se obtuvieron datos sobre su actividad catalítica, su estado oligomérico y su desnaturalización térmica y química, pero no fue posible caracterizarla sin etiqueta de histidinas, por lo que los datos no son completamente comparables para un estudio de tipo evolutivo.

### **Selección de especies**

Para la realización de este trabajo, como se verá más adelante, se encontraron condiciones para la expresión soluble de las dos TIM mesófilas, por lo que se decidió seguir con la caracterización de estas enzimas del orden Methanococcales (en la Tabla S3 se compilan algunas de las características de las especies seleccionadas, así como las secuencias de aminoácidos de las TIM elegidas). Además de

esas secuencias, se propuso ampliar el número de proteínas por caracterizar para abarcar los cuatro géneros de este clado. La lista final de las especies seleccionadas es la siguiente:<sup>5</sup>

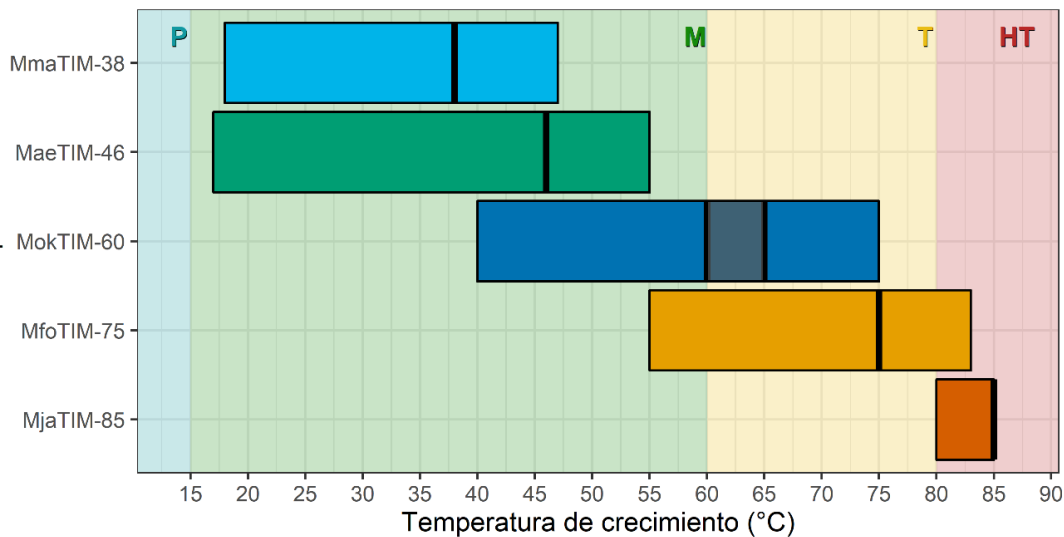
De la familia Methanocaldococcaceae (o **TermoTIM**)

- *Methanocaldococcus jannaschii* DSM 2661 (**MjaTIM-85**)
- *Methanotorrus formicicus* Mc-S-70 (**MfoTIM-75**)

De la familia Methanococcaceae (o **MesoTIM**)

- *Methanothermococcus okinawensis* IH1 (**MokTIM-60**)
- *Methanococcus aeolicus* Nankai-3 (**MaeTIM-46**)
- *Methanococcus maripaludis* C7 (**MmaTIM-38**).

Entre estos cinco organismos se puede abarcar un amplio intervalo de temperaturas que van desde los ~17 a los 85°C (Figura 4). En las Tabla 2 y Tabla 3 se resumen algunos datos de interés de las secuencias de las TIM de las arqueas seleccionadas (colectivamente denominadas **ArqTIM** en este trabajo).



**Figura 4. Temperaturas de crecimiento de los organismos seleccionados.** En barras de color se muestra el intervalo en el que crecen las arqueas seleccionadas y con líneas negras se indica la temperatura óptima de crecimiento ( $T_{opt}$ ) de cada una. En el fondo se marca la clasificación de los organismos según su temperatura óptima de crecimiento. P. Psicrófilo, M. Mesófilo, T. Termófilo y HT. Hipertermófilo.

<sup>5</sup> Las abreviaturas indicadas obtienen la primera letra del género y las siguientes dos letras del epíteto específico de la especie a la que pertenecen. El número indicado al final corresponde a la temperatura óptima de crecimiento del organismo para facilitar la referencia al tipo de hábitat.

**Tabla 2. Datos de interés de los organismos y sus secuencias de la TIM para este trabajo.**

Característica	MmaTIM-38	MaeTIM-46	MokTIM-60	MfoTIM-75	MjaTIM-85
T <sub>opt</sub> de la especie (°C)	38 (Mesófilo)	46 (Mesófilo)	60-65 (Termófilo)	75 (Termófilo)	85 (Hipertermófilo)
Intervalo de crecimiento de la especie (°C)	18-47	17-55	40-75	55-83	80-85
No. de aa	217	226	222	221	219
pl teórico	5.21	5.08	5.6	5.64	5.48
Masa molecular del monómero (kDa)	22.88	24	23.48	23.5	23.32
No. de Trp	0	0	0	0	0
No. de Tyr	3	6	4	4	4
No. de Phe	1	3	2	1	2
No. de Cys	4	4	5	3	5
No. de Pro	9	8	9	9	8
No. de His	3	2	5	4	4

**Tabla 3. Porcentaje de identidad entre las secuencias elegidas.**

<b>MmaTIM-38</b>	---				
<b>MaeTIM-46</b>	64.1	---			
<b>MokTIM-60</b>	73.9	74.3	---		
<b>MfoTIM-75</b>	74.8	64.6	77.9	---	
<b>MjaTIM-85</b>	74.3	64.1	75.2	80.5	---
	<b>MmaTIM-38</b>	<b>MaeTIM-46</b>	<b>MokTIM-60</b>	<b>MfoTIM-75</b>	<b>MjaTIM-85</b>

Algo destacable de estas secuencias es la ausencia de triptófanos. Esta característica parece ser sinapomórfica de las TIM de Methanococcales, pues secuencias de TIM de otras arqueas sí cuentan con este residuo. Aunque se espera por la relación filogenética tan estrecha entre las ArqTIM, es también notable el porcentaje de identidad elevado, el cual se encuentra entre el 64 y el 80 %.

### **La adaptación como concepto para organizar un trabajo de investigación**

La crítica que se hace al asumir que cualquier característica biológica es una adaptación se relaciona, como ya se mencionó, con que no se suele dar una corroboración experimental de que sea así y que se toma a la selección como el agente causal más importante en la evolución de los seres vivos. Algunos autores han denominado a esta práctica como *adaptacionismo empírico* (Godfrey-Smith, 2001; Lewens, 2009). No obstante, la noción de adaptación también se puede utilizar como una guía heurística para proponer un proyecto de investigación. Es decir, se plantea que una característica X es una adaptación para una variable ambiental T de modo que se pueda organizar una investigación que nos permita determinar si X es, efectivamente, una adaptación para T. Aún si no lo es, esta aproximación nos permite recopilar evidencia de cuál podría ser el verdadero origen evolutivo de un carácter (o de si X es una adaptación para otra variable ambiental) y proponer nuevos experimentos.

Lo anterior se hace sin afirmar que la selección natural tenga una mayor importancia causal que otras fuerzas evolutivas. A esta perspectiva se le conoce como *adaptacionismo metodológico (ibid.)*, y es desde esta aproximación con la que se definen los objetivos y la hipótesis de este trabajo.

## HIPÓTESIS

Características biofísicas de las ArqTIM, como su estabilidad conformacional, termoestabilidad, estado oligomérico y actividad, son adaptaciones que permitieron la vida a altas temperaturas.

## OBJETIVOS

### ***Objetivo general***

Determinar si las características biofísicas de la TIM de las arqueas Methanococcales (v. gr. estabilidad conformacional, termoestabilidad, estado oligomérico, etc.) se seleccionaron de acuerdo con la temperatura en la que estos microorganismos habitan.

### ***Preguntas particulares***

- ¿Cuál es variación existente entre las características biofísicas de las ArqTIM?
- ¿Por qué se presenta esta variación?
- ¿Puede la SN ser el mecanismo causal en la evolución de las características biofísicas de estas proteínas? ¿De cuáles?
- ¿Qué otras fuerzas evolutivas pueden operar sobre éstas?
- ¿Qué otras variables ambientales pueden estar jugando un papel en la evolución de las ArqTIM?

## MÉTODOS

### ***Subclonación de genes***

Para la producción heteróloga de las ArqTIM, se adquirieron de la compañía GenScript construcciones con un sitio de restricción para *NdeI* en el extremo 5' y un sitio para *XhoI* en el extremo 3'. Los genes se subclonaron en un vector pET22b(+) para regular la expresión de los polipéptidos con la adición de isopropil-β-D-1-tiogalactopiranosido (IPTG). En estas construcciones nuevas no se incluyó etiqueta de afinidad alguna para la proteína. Los insertos se purificaron por obtención de banda en un gel de agarosa al 1% con el kit de extracción de banda de ZymoResearch después de digerir por 12 horas con las enzimas de restricción correspondientes de la marca New England BioLabs Inc. La reacción de ligación se realizó en una proporción 3:1 de inserto-vector por 12 horas a 22 °C y 2 horas a 16°C con polimerasa T4 de la marca Thermo Fisher Scientific.

Todos los plásmidos utilizados para la sobre-expresión fueron secuenciados para corroborar la integridad de los genes. Esto se realizó a partir del aislamiento del plásmido superenrollado con el sistema de extracción de plásmido Miniprep de Biomiga Inc. El cebador de secuenciación utilizado fue el de la polimerasa T7: 5'- TAA TAC GAC TCA CTA TAG GG-3'.

### ***Expresión y purificación de las proteínas seleccionadas***

Los plásmidos de MjaTIM-85 y MfoTIM-75 fueron transformados en células quimiocompetentes de *E. coli*-Gold (DE3) por choque térmico. Para la expresión, las células se cultivaron en medio LB con ampicilina 100 µg mL<sup>-1</sup> a 37 °C y agitación a 220 rpm hasta alcanzar una densidad óptica a 600 nm (DO<sub>600nm</sub>) de 0.8-1.0, donde se indujo la expresión con 1 mM de IPTG. Después de 16 horas, las células fueron cosechadas y se rompieron por sonicación en amortiguador de lisis (fosfatos de sodio 35 mM, NaCl 300 mM, pH 8.0). Los lisados celulares se centrifugaron a 15,000 rpm por 15 minutos para recuperar el sobrenadante. Los sobrenadantes se incubaron a 70 °C por 20 minutos para precipitar por calor algunas proteínas endógenas de *E. coli*, las cuales se separaron de la muestra por centrifugación. Posteriormente se realizó una precipitación de las proteínas de interés con sulfato de amonio, MjaTIM-85 precipita entre el 50 a 80% de saturación, mientras que MfoTIM-75 lo hace entre el 60 y 90%. Las proteínas se pasaron por una columna de intercambio aniónico HiTrap Q en amortiguador de trietanolamina 10 mM, EDTA 1 mM, DTT 1 mM (TED) pH 7.0. Para poder purificar a MjaTIM-85 fue necesario un paso cromatográfico adicional por la columna de intercambio aniónico en TED 10:1:1, pH 8.5 para concluir su purificación.

Para las MesoTIM se probaron una serie de métodos para optimizar su expresión y purificación. Aquí sólo se describe el método por el cual se obtuvo el mejor resultado de purificación. Los plásmidos de MokTIM-60, MaeTIM-46 y MmaTIM-38 fueron insertados por choque térmico en células quimiocompetentes de *E. coli*-Gold (DE3) que ya contaban con el plásmido pG-TF2 del sistema de coexpresión con chaperonas de TakaraBio Inc, el cual permite inducir la expresión de las chaperonas trigger factor, GroEL y GroES con la adición de tetraciclina en el medio. Las células se cultivaron en medio LB con ampicilina  $100 \mu\text{g mL}^{-1}$ , cloranfenicol  $35 \mu\text{g mL}^{-1}$  y tetraciclina  $10 \text{ ng mL}^{-1}$  a  $37^\circ\text{C}$  y agitación a 220 rpm hasta alcanzar una  $\text{DO}_{600\text{nm}}$  de 0.4-0.6, donde se indujo la expresión con 1 mM de IPTG. La inducción se dejó bajo las mismas condiciones por 16 horas antes de cosechar las células y romper por presión en una prensa de French en el amortiguador de lisis antes descrito. Los lisados celulares se centrifugaron a 15000 rpm por 15 minutos para recuperar el sobrenadante. Los sobrenadantes se precipitaron con sulfato de amonio (30 a 60% de saturación), para después pasar el contenido en una columna de intercambio aniónico HiTrap Q, primero en amortiguador TED 10:1:1, pH 7.0 y posteriormente en amortiguador TED 10:1:1, pH 8.5.

Todas las columnas de cromatografía se acoplaron a un sistema Äkta FPLC (GE®). Para todos los casos, los cambios de amortiguador se llevaron a cabo en membranas de diálisis con corte de 10 kDa con agitación constante a  $4^\circ\text{C}$ . Para cada diálisis se realizaron en total tres cambios del amortiguador objetivo, con 3-4 horas de diferencia entre dos de ellos y una noche de diferencia en el restante, sin un orden en particular. En todos los casos se procuró mantener una relación 1:100 entre la muestra y el volumen total de la diálisis. Las proteínas se concentraron por ultracentrifugación en membranas con corte de 10 kDa. La cuantificación de las muestras se hizo por el método del ácido bicinónico (Walker, 1996).

### **Caracterización espectroscópica**

#### *Estructura secundaria*

El dicroísmo circular (DC) es una técnica para la caracterización estructural de moléculas quirales que nos puede dar indicios sobre el estado de la estructura secundaria y terciaria de las proteínas. El fenómeno de DC se refiere a la absorción diferencial de luz polarizada circularmente de dos componentes de la misma amplitud: uno viaja rotando en favor de las manecillas del reloj y otro en contra. Esta diferencia en la absorción se manifiesta en la lectura de un espectro de luz polarizada elípticamente, razón por la cual los datos producidos suelen expresarse en grados de elipticidad ( $\theta$ ). Para nuestros experimentos se obtuvieron espectros de DC en el intervalo del ultravioleta (UV) lejano ( $\lambda < 240 \text{ nm}$ ). En proteínas, este intervalo de interés mide la absorción dada por el enlace peptídico, por lo que los datos obtenidos dan información sobre el tipo de estructura secundaria que presenta la proteína (Greenfield, 2007; Kelly, Jess, & Price, 2005). Los ensayos de DC se realizaron en un

espectropolarímetro Chirascan (Applied Photophysics ©) acoplado a un sistema Peltier (PolyScience ®) para el control de la temperatura. En todos los casos se siguió la señal de DC en el espectro de 180 a 260 nm, midiendo cada nanómetro, con 0.4 mg mL<sup>-1</sup> de la proteína en amortiguador de fosfato de sodio 10mM, pH 8.0 a 25°C en una celda de cuarzo de 1 mm de ancho.

Para la comparación entre resultados, los datos obtenidos fueron normalizados por su masa molecular (*MM*) y el número de residuos (*N*) para obtener la elipticidad promedio por residuo  $[\theta]$  según las ecuaciones 1 y 2 (Kelly et al., 2005):

$$[\theta] = \frac{PPR \times \theta_i}{10 \times d \times c} \quad (1)$$

$$PPR = \frac{MM}{(N-1)} \quad (2)$$

donde *PPR* es el peso promedio por residuo,  $\theta_i$  son los grados de elipticidad a cada longitud de onda, *d* es el paso de luz en centímetros y *c* es la concentración de proteína en g mL<sup>-1</sup>.

Los datos normalizados se deconvolucionaron por un método de selección variable con el programa CDSSTR (Sreerama, Venyaminov, & Woody, 2000; Sreerama & Woody, 2000) a través del servidor en línea DichroWeb (Whitmore & Wallace, 2004, 2008; <http://dichroweb.cryst.bbk.ac.uk/html/home.shtml>).

### *Estructura terciaria*

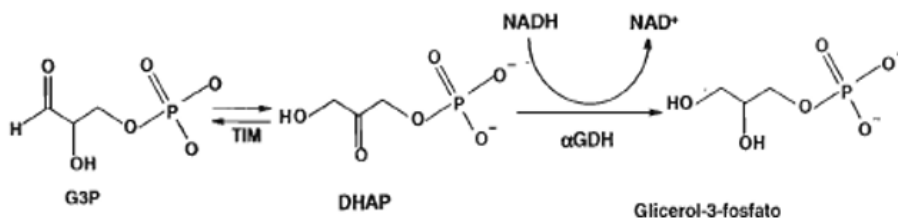
La fluorescencia intrínseca (IF) de una proteína proviene de los aminoácidos aromáticos que poseen en su secuencia: triptófano, tirosina y fenilalanina. Debido a que estos residuos, en la mayoría de las ocasiones, se encuentran total o parcialmente ocultos —particularmente el triptófano— en el núcleo hidrofóbico o en las superficies de interacción entre subunidades de una proteína, sus espectros de emisión presentan intensidades y máximos de emisión diferentes a los que tienen cuando se encuentran expuestos al solvente. Por esta razón, la IF puede utilizarse como una medida del estado terciario de las proteínas, ya que podemos determinar si los residuos se encuentran empaquetados al interior de una proteína dadas sus características espectrales (Royer, 2006). Los espectros de IF se realizaron con un espectrofluorómetro PC1 (ISS©) acoplado a un sistema Peltier (PolyScience®) para mantener control sobre la temperatura de la celda de muestra. Ya que las ArqTIMs no presentan residuos de triptófano en su secuencia, todas las mediciones se realizaron con una longitud de onda de excitación de 278 nm —el valor recomendado para excitar tirosinas— y se colectaron espectros de emisión de 295 a 420 nm. Los ensayos se hicieron con 0.4 mg mL<sup>-1</sup> de la proteína en amortiguador de

fosfatos de sodio 10 mM, pH 8.0 a 25°C. Como se verá más adelante, los espectros de fluorescencia se utilizaron para corroborar el efecto de contaminantes en las preparaciones enzimáticas.

### **Determinación de parámetros catalíticos**

La espectroscopía de absorción es uno de los métodos más comunes para realizar ensayos de actividad enzimática. Esta técnica es particularmente útil cuando alguno de los reactivos involucrados en la reacción presenta color en el espectro de la luz visible o UV, y su transformación implica un cambio espectral de coloración. En un ensayo enzimático por absorción se utiliza un espectrofotómetro para medir de manera continua el cambio en absorbancia de la sustancia colorida en un intervalo de tiempo. Estos experimentos nos permiten registrar el progreso de una reacción enzimática y obtener un gráfico de producción de producto por unidad de tiempo. De este gráfico se puede obtener la velocidad inicial de la reacción ( $v_0$ ), la cual se considera como el intervalo de tiempo en el cual la velocidad de la reacción permanece constante y así determinar los parámetros catalíticos de la enzima en condiciones ideales (Purich, 2010b).

En los ensayos enzimáticos de las TIM evaluadas en este estudio se analizó la conversión del gliceraldehído-3-fosfato (G3P) a dihidroxiacetona fosfato (DHAP). Estos dos reactivos no presentan color, por lo que para determinar los parámetros catalíticos se realizaron ensayos acoplados con la enzima  $\alpha$ -glicerol-3-fosfato deshidrogenasa ( $\alpha$ -GDH, Roche), la cual reduce el DHAP a glicerol-3-fosfato a través de la oxidación de una molécula de NADH (Gómez-Puyou *et al.*, 1995, Figura 5). El NADH tiene un máximo de absorbancia a 340 nm y un coeficiente de absorción molar ( $\epsilon$ ) de 6220  $M^{-1} cm^{-1}$ .



**Figura 5. Reacción acoplada para medir la actividad de una TIM.** Tomada de Romero Romero, 2011.

A partir de dividir el valor de las  $v_0$  obtenidas entre la cantidad de proteína en la celda de reacción y el  $\epsilon$  del NADH se puede calcular la actividad específica en  $\mu$ moles de DHAP formados por minuto por miligramo de proteína. Este procedimiento se repitió con distintas concentraciones iniciales de G3P para determinar la dependencia de la actividad específica con la concentración de sustrato [S]. Los datos obtenidos se ajustaron a una ecuación no lineal de Michaelis-Menten (ecuación 3):



$$v_0 = \frac{V_{m\acute{a}x} \times [S]}{K_M + [S]} \quad (3)$$

donde  $V_{m\acute{a}x}$  es la velocidad l\imite que puede alcanzar una reacci\on catalizada por una enzima y  $K_M$  es la concentraci\on de substrato a la cual se observa la mitad del valor de la  $V_{m\acute{a}x}$ . Al dividir la  $V_{m\acute{a}x}$  entre la cantidad de enzima total en el ensayo se puede obtener el n\umero de recambio ( $k_{cat}$ ), el cual indica el n\umero de mol\eculas de substrato convertidas en producto por mol\ecula de enzima por unidad de tiempo. Asimismo, se puede obtener la eficiencia catal\itica de la reacci\on ( $k_{cat} / K_M$ ), la cual es una constante que nos indica la especificidad que tiene una enzima por el substrato (Purich, 2010a). Para calcular la eficiencia catal\itica es necesario hacer una modificaci\on del valor de la  $K_M$  obtenida, esto se debe a que, en soluci\on acuosa, el G3P existe como diol y como aldeh\ido en una relaci\on 29:1, pero la TIM s\olo puede catalizar la forma aldeh\ido de la mol\ecula (Trentham, McMurray & Pogson, 1969).

Todos los ensayos se realizaron en amortiguador TED 100:10:1 mM, pH 7.4. Cada reacci\on se coloc\o en una celda de pl\astico con  $20 \mu\text{g mL}^{-1}$  de  $\alpha$ -GDH, 0.2 mM de NADH y  $5 \text{ ng mL}^{-1}$  de la preparaci\on enzim\atica en un volumen total de 1 mL. Las concentraciones de G3P evaluadas fueron de 0.025 a 3–4 mM. La actividad se monitore\o con un espectrofot\ometro Beckman DU7500 termorregulado (Beckman Coulter \textcircled{c}) a 25 \text{ }^\circ\text{C}.

### **Determinaci\on del estado oligom\erico**

Para determinar el estado oligom\erico hay una gran disponibilidad de m\etodos que nos dan informaci\on sobre el estado de autoasociaci\on de las prote\inas de inter\es. En este trabajo se emple\o la cromatograf\ia de exclusi\on anal\itica (SEC) para este fin.

La SEC es un m\etodo que se basa en la migraci\on de las prote\inas a trav\es de una matriz de esferas hidratadas hechas de pol\imeros entrecruzados como el dextrano, la agarosa o la poli\acril\amida. La muestra de prote\inas al interior de la columna se reparte en dos espacios: el que corresponde al solvente que rodea a las esferas, y el que corresponde al solvente que se encuentra en el interior de \estas. Mientras m\as peque\na sea una prote\ina, la fracci\on que se va a repartir entre los intersticios de las esferas ser\a mayor, por lo que se necesitar\a un mayor volumen de amortiguador para que dicha prote\ina salga de la columna. La retenci\on de una prote\ina al interior de una columna es, en principio, inversamente proporcional a su tama\no en soluci\on, o m\as espec\ificamente, a su radio de Stokes, esto es, el radio de una esfera s\olida con el mismo coeficiente de difusi\on traslacional. Para prote\inas globulares, este m\etodo da estimaciones del radio de Stokes con las cuales se puede determinar la

masa molecular por medio de una calibración con proteínas estándar bien caracterizadas (Gell, Grant, & Mackay, 2012; Hong, Koza, & Bouvier, 2012).

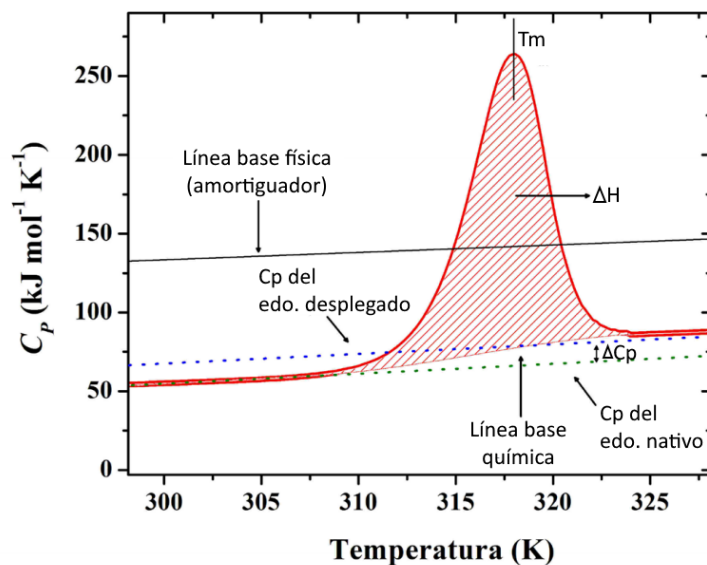
Idealmente, la separación de proteínas por SEC sólo depende de los efectos estéricos, pero en ocasiones la resina y las proteínas pueden interactuar de manera no específica, si esto ocurre el tiempo de retención puede verse alterado (GE Healthcare, s/f). Para evitar dichas interacciones, las muestras de proteínas se incubaron en un amortiguador de fosfatos de sodio 50 mM, NaCl 200 mM, pH 8.0. Además, las muestras se incubaron por al menos 24 horas a 25 °C para asegurar que los cambios en el estado oligomérico, si los hubiera, se encontraran al equilibrio. Para cada proteína se ensayaron muestras a 0.5, 0.1 y 0.05 mg mL<sup>-1</sup>.

Los análisis de SEC se realizaron en una columna de exclusión molecular analítica Superdex 200 10/300 GL (GE ©) acoplada a un sistema FPLC (GE©). La calibración de la masa molecular y el radio de Stokes se realizó con los estándares para filtración en gel de BioRad© y se incluyó una calibración con las TIM dimérica de *Clostridium perfringens* (CpTIM) y de *Methanocaldococcus jannaschii* (MjaTIM-85) con etiqueta de histidinas. El análisis consistió en obtener una regresión lineal del gráfico de *log masa molecular* contra *volumen de retención* de los puntos de calibración. De lo anterior se interpolaron los valores de masa molecular tomando en consideración el volumen de retención de las muestras estudiadas y se determinó el número de monómeros que deben considerarse por partícula.

### **Desnaturalización por temperatura**

La desnaturalización por temperatura se evaluó por calorimetría diferencial de barrido (DSC, por sus siglas en inglés). Esta técnica es un método calorimétrico empleado para estudiar los procesos inducidos por cambios en la temperatura. *Grosso modo*, un dispositivo de DSC contiene dos celdas idénticas al interior de una chaqueta adiabática. Durante un experimento, una celda se llena con solución amortiguadora y otra con la solución muestra (que puede ser amortiguador si se está determinando el valor de referencia), para posteriormente comenzar a calentar ambas celdas de manera simultánea y a la misma velocidad (llamada velocidad de barrido). Mientras no existan cambios en la muestra que emitan o absorban calor (como la desnaturalización de una proteína), la diferencia de temperatura ( $\Delta T$ ) entre ambas celdas debería ser casi cero. En un evento de desnaturalización, la diferencia incrementa, por lo que el dispositivo tendrá que regular su suministro de energía para volver a tener una  $\Delta T$  cercana a cero. A partir de esas diferencias en el suministro se puede calcular la capacidad calorífica a presión constante ( $C_p$ ) del proceso observado. (Johnson, 2013; Romero-Romero, Fernández-Velasco, & Costas, 2018). Los experimentos de DSC generan gráficos de  $C_p$  contra  $T$  llamados trazas calorimétricas o endotermas (ver Figura 6). Estos ensayos se realizaron en

colaboración con el Dr. Miguel Costas Basín y la Dra. Jessica Díaz Salazar del Laboratorio de Biofísicoquímica de la Facultad de Química, con el calorímetro capilar VP-Capillary DSC (Microcal ©). Todas las trazas calorimétricas se obtuvieron en amortiguador de fosfatos de sodio 10 mM, pH 8.0. La concentración de proteína, el intervalo y la velocidad de barrido se establecieron según el tipo de experimento y la proteína evaluada.



**Figura 6. Elementos básicos de una endoterma para una proteína en solución.** Cambio en entalpía,  $\Delta H$ . Temperatura media de desnaturalización,  $T_m$  (Modificada de Romero-Romero *et al.*, 2018).

Las trazas de desnaturalización térmica obtenidas por DSC se corrigieron por el tiempo de respuesta del calorímetro y se normalizaron con respecto a la concentración de proteína *considerando la masa molecular del monómero*. También se les restó la línea base física (*i. e.* línea blanco amortiguador-amortiguador) para obtener la capacidad calorífica en exceso ( $C_p^{exc}$ ).

#### Modelo irreversible de dos estados

Para el análisis de los datos bajo control cinético se hizo un ajuste de los datos a un modelo irreversible de dos estados (Sanchez-Ruiz, 1992):



donde  $N$  es el estado nativo que, en un solo paso, alcanza un estado final ( $F$ ) de cual es incapaz de regresar. Esta conversión cinética está descrita por una constante cinética de primer orden dependiente de la temperatura ( $k$ ). Las trazas calorimétricas de este modelo están descritas por las ecuaciones 4 y 5:

$$C_p^{ap} = \frac{\Delta H \times E_A}{R \times T_m^2} e^x \times e^{-e^x} \quad (4)$$

$$x = \frac{E_A}{R \times T_m^2} (T - T_m) \quad (5)$$

donde  $T_m$  es la temperatura que corresponde a la mitad de la transición,  $T$  es la temperatura en cada punto de la endoterma,  $E_A$  es la energía de activación del proceso,  $\Delta H$  es la entalpía de desplegamiento,  $R$  es la constante de los gases ideales ( $8.314 \text{ J mol}^{-1} \text{ K}^{-1}$ ) y  $C_p^{ap}$  es la capacidad calorífica aparente corregida por la línea base química.

Los cuatro métodos para calcular la  $E_A$  que se describen a continuación se pueden encontrar en Costas *et al.*, 2009:

- 1) Se ajustaron de manera individual las trazas a diferentes velocidades de barrido según el modelo irreversible de dos estados (ecuaciones 2 y 3) y se obtuvo un valor promedio de  $E_A$ .
- 2) De la dependencia entre la velocidad de barrido ( $v$ ) sobre la  $T_m$  (ecuación 6):

$$\ln\left(\frac{v}{T_m^2}\right) = \text{constante} - \frac{E_A}{R \times T_m} \quad (6)$$

Así, un gráfico de  $\ln(v/T_m^2)$  contra  $1/T_m$  el valor de la  $E_A$  se obtiene a partir de la pendiente.

- 3) Dado que la velocidad de la conversión entre N y F está dada por una constante de velocidad de primer orden ( $k$ ), el proceso se puede analizar de acuerdo con la ecuación de Arrhenius (7):

$$k(T) = \exp\left[-\frac{E_A}{R}\left(\frac{1}{T} - \frac{1}{T^*}\right)\right] \quad (7)$$

donde  $T^*$  es la temperatura en la cual  $k = 1 \text{ min}^{-1}$ . Con el gráfico  $1/T$  contra  $\ln k$  para todas las velocidades de barrido, se multiplica el valor de la pendiente por  $R$  y se obtiene el valor de la  $E_A$ .

- 4) Se ha observado que existe una relación inversa entre el valor de la  $E_A$  y el ancho de la transición calorimétrica descrita por la ecuación 8:

$$W = 2.45 \frac{R \times T_m^2}{E_A} \quad (8)$$

donde  $W$  es el ancho de la transición calorimétrica a la mitad de la altura máxima del pico.

#### *Modelo con transiciones independientes de más de dos estados*

Para el análisis de los datos bajo control termodinámico se restó la base química para eliminar el efecto del  $\Delta C_p$  y se implementó un modelo de desplegamiento con transiciones independientes de más de dos estados (MicroCal, 1998, ecuaciones 9 y 10):

$$C_p(T) = \frac{K_A(T) \times \Delta H_{vH} \times \Delta H_{cal}}{(1 + K_A(T))^2 \times RT^2} + \dots \quad (9)$$

$$K_A(T) = \exp \left[ \frac{-\Delta H_{vH}}{R \times T} \times \left( 1 - \frac{T}{T_m} \right) \right] \quad (10)$$

donde  $C_p$  es la capacidad calorífica en cada punto,  $K_A$  es la constante de equilibrio en cada punto,  $\Delta H_{vH}$  es la entalpía de van't Hoff,  $\Delta H_{cal}$  es el valor calorimétrico de la entalpía. Este modelo permite analizar los datos cuando hay más de una transición. El  $\Delta H_{vH}$  corresponde al cambio de entalpía por unidad cooperativa de plegamiento que participa en la reacción de desnaturalización según la ecuación (11) de van't Hoff:

$$\left( \frac{\delta \ln K_A}{\delta T} \right) = \frac{\Delta H_{vH}}{R \times T^2} \quad (11)$$

#### **Desnaturalización por agentes caotrópicos**

Para establecer el tiempo de equilibrio en la desnaturalización con clorhidrato de guanidinio (GdnHCl) se incubó la proteína a 25°C en concentraciones diferentes del desnaturalizante en un intervalo de 0 a 7 M, midiendo la señal de DC a 222 nm a las 12, 36 y 48 horas de incubación.

Para todos los experimentos descritos en esta sección, las proteínas fueron incubadas a 25 °C en amortiguador de fosfatos de sodio 10 mM, pH 8.0 con concentraciones ascendentes del agente desnaturalizante realizadas a partir de la dilución de una solución madre. Para corroborar la concentración final del desnaturalizante en la solución madre, así como en cada una de las diluciones, se determinó el índice de refracción obtenido con un refractómetro Abbe NAR-3T (Atago U.S.A, Inc.), manteniendo una temperatura controlada de 25 °C con un baño recirculador de agua. Dichos valores

fueron ingresados en la calculadora en línea *Denaturant Concentration Calculator*© (disponible en <http://sosnick.uchicago.edu/gdmcl.html>) del grupo Sosnick de la Universidad de Chicago (2017).

#### *Dicroísmo circular*

Los ensayos de dicroísmo circular se realizaron en un espectropolarímetro Chirascan™ (Applied Photophysics©) acoplado a un sistema Peltier (PolyScience®) para controlar la temperatura (25 °C). Los espectros se obtuvieron haciendo barridos por triplicado de 190 a 260 nm de longitud de onda con medidas cada 1 nm. Para construir la curva de desnaturalización a distintas concentraciones de desnaturalizante, el monocromador del equipo se fijó en 222 nm y se obtuvo el valor promedio de 200 mediciones para cada punto de la gráfica.

#### *Anisotropía*

La anisotropía de fluorescencia (simplemente referida como *anisotropía* durante este texto) es un fenómeno en el que una partícula al ser excitada con luz polarizada genera un espectro de emisión que también se encuentra polarizado. Esto se debe a que, en una solución homogénea, los fluoróforos de una muestra están aleatoriamente orientados, por lo que aquellos que en el momento de la excitación estén en eje con el vector eléctrico serán excitados preferencialmente. Las medidas de anisotropía revelan el desplazamiento angular promedio por el que pasa un fluoróforo entre el evento de excitación y su subsecuente emisión en forma de fotón. El desplazamiento angular depende de la tasa de rotación difusional de la partícula, la forma y tamaño de la molécula en rotación y de la viscosidad del solvente. Así, mientras menor sea la tasa de difusión rotacional, mayor será el valor de anisotropía (Lakowicz, 2006).

Las medidas de anisotropía se realizaron con un espectrofluorómetro PC1 (ISS©) acoplado a un sistema Peltier (PolyScience®) para mantener control sobre la temperatura de la celda de muestra (25 °C). Debido a la ausencia de triptófanos en las secuencias peptídicas elegidas, la longitud de onda de excitación utilizada fue de 278 nm y la emisión se colectó a 303 nm, que corresponde al máximo de emisión en fluorescencia. Para cada proteína se determinó el factor G, el cual constituye un factor de corrección de la sensibilidad del sistema de detección de luz polarizada del equipo (Lakowicz, 2006). Los datos presentados se obtuvieron del promedio de cien mediciones del valor de anisotropía para cada punto de las curvas de desnaturalización.

#### *Actividad catalítica*

Los ensayos de actividad se llevaron bajo las mismas condiciones especificadas anteriormente, salvo que todos los experimentos se hicieron con 4 mM de G3P. La cantidad total de proteína utilizada para cada proteína se especifica en las figuras correspondientes.

### Procesamiento de los datos de desplegamiento químico

Los datos que fueron irreversibles al desplegamiento por GdnHCl se ajustaron a la ecuación de Boltzmann (12) o doble Boltzmann (13), según el caso, para determinar de manera cuantitativa el valor de la concentración media de desnaturalización ( $C_m$ ) para cada transición.

$$y = \frac{A_1 - A_2}{1 + e^{(x-x_0)/dx}} + A_2 \quad (12)$$

La ecuación 12 describe una curva sigmoideal donde  $A_1$  es el valor inicial de  $y$  antes de la transición,  $A_2$  es el valor final de  $y$  después de la transición,  $dx$  es una constante que describe la pendiente de la transición, y  $x_0$  es el punto de inflexión de la curva (OriginLab, s/f-a), el cual se considera como el valor de la  $C_m$  en este caso.

$$y = y_0 + A \left[ \frac{p}{1 + e^{\frac{x-x_{01}}{k_1}}} + \frac{1-p}{1 + e^{\frac{x-x_{02}}{k_2}}} \right] \quad (13)$$

La ecuación 13 describe la suma de dos curvas sigmoideales, por lo que comprende dos transiciones consecutivas.  $A$  es la diferencia que existe entre los valores inicial y final de  $y$ ,  $y_0$  es el valor inicial de  $y$ ,  $p$  es el valor de la curva a la mitad de  $A$ ,  $k_1$  y  $x_{01}$  son la pendiente y el punto de inflexión de la primera transición, y  $k_2$  y  $x_{02}$  son la pendiente y el punto de inflexión de la segunda transición, respectivamente (OriginLab, s/f-b). En este caso,  $x_{01}$  y  $x_{02}$  se consideran como los valores de la  $C_m$  para cada transición.

En los casos donde sí hubo reversibilidad, se puede determinar el valor de su estabilidad conformacional. Debido a que la desnaturalización se manifiesta en presencia de un desnaturalizante, es necesario hacer una estimación de su energía libre de Gibbs como si estuvieran en 0 M de desnaturalizante ( $\Delta G_{H_2O}$ ). Esta estimación se hizo con tres aproximaciones diferentes:

- 1) Los datos crudos fueron analizados bajo la ecuación no lineal de desplegamiento de dos estados (14) de Santoro & Bolen (1988):

$$y = \frac{(y_N + m_N[M]) + (y_D + m_D[M])e^{-\frac{\Delta G_{H_2O} - m[M]}{R \times T}}}{1 + e^{-\frac{\Delta G_{H_2O} - m[M]}{R \times T}}} \quad (14)$$

donde  $y$  es la señal observada según el método,  $y_N$  y  $y_D$  son las ordenadas al origen y  $m_N$  y  $m_D$  son las pendientes de las líneas base de la pre y postransición,  $[M]$  es la concentración de desnaturalizante y  $m$  es un valor que indica la dependencia del  $\Delta G$  con la concentración de desnaturalizante.

- 2) Los datos fueron normalizados para obtener el valor del cambio fraccional ( $f_D$ ) con la ecuación 15:

$$f_D = \frac{(y_N + m_N[M]) - y}{(y_D + m_D[M]) - (y_N + m_N[M])} \quad (15)$$

y posteriormente fueron ajustados a la ecuación anterior (14).

- 3) A partir de los valores de  $f_D$  se determinó la constante de equilibrio ( $Keq$ ) (ecuación 16) y el valor de estabilidad para cada punto de la curva de desnaturalización ( $\Delta G_D$ ) (ecuación 17).

$$Keq = \frac{f_D}{1 - f_D} \quad (16)$$

$$\Delta G_D = -R \times T \times \ln Keq \quad (17)$$

Con los valores resultantes se aplicó una extrapolación lineal de los datos de la transición según la ecuación (18) de Greene & Pace (1974):

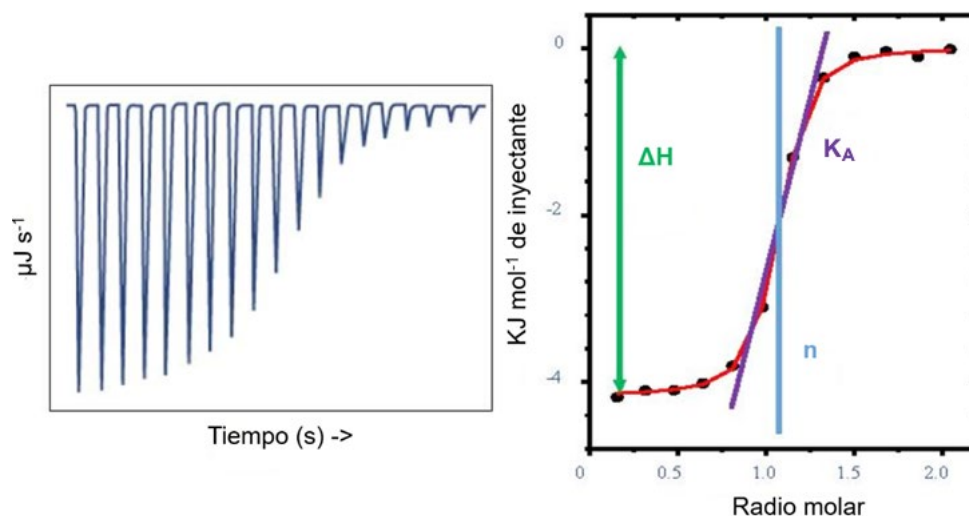
$$\Delta G_D = m[M] + \Delta G_{H_2O} \quad (18)$$

### **Unión a ligando**

Una de las técnicas que se utiliza para caracterizar biofísicamente la interacción de proteínas con otras moléculas es la calorimetría de titulación isotérmica (ITC, por sus siglas en inglés). Similar a la DSC, los dispositivos de ITC tienen dos celdas idénticas embebidas en una chaqueta adiabática y mantienen ambas celdas con una  $\Delta T$  de casi cero. Esta técnica mide las interacciones entre dos partículas al detectar el calor absorbido o emitido por una reacción química que altera el  $\Delta T$  de las celdas. En otras palabras, mide el cambio en la entalpía de unión ( $\Delta H_U$ ) bajo condiciones térmicas muy controladas. De manera general, la ITC consiste en evaluar el progreso de la transferencia de calor durante la titulación de una solución que contiene la proteína de interés —localizada en una celda de muestra— con alícuotas de la molécula que se le une —ubicada en la jeringa del calorímetro— en una serie de inyecciones de poco volumen. De un experimento de ITC (Figura 7) se pueden obtener directamente



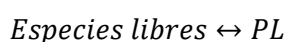
valores numéricos del  $\Delta H_U$ , la constante de asociación ( $K_A$ ) y la estequiometría ( $n$ ) de la reacción (Baranauskiene, Kuo, Chen, & Matulis, 2019; Brautigam, Zhao, Vargas, Keller, & Schuck, 2016; Doyle, 1997; Holdgate & Ward, 2005).



**Figura 7. Elementos básicos de una isoterma de la unión a ligando.** En el panel de la izquierda se muestran los datos crudos de un experimento de titulación. En el de la derecha la curva isoterma que se obtiene de un experimento ajustado a un modelo de unión 1:1 (rojo) (Modificado de Frasca, 2016)

Estos experimentos se realizaron con la ayuda de un equipo Affinity ITC LV (TA Instruments®) y se midió la unión de las TIM estudiadas con el ácido fosfoglicolohidroxámico (PGH), el cual es un inhibidor fosforilado análogo del estado de transición de esta familia de proteínas. Para cada experimento se desgasificaron las soluciones que fueron introducidas en el dispositivo. El amortiguador de trabajo de estos ensayos fue trietanolamina (TEA) 100 mM, pH 7.4. Para cada protocolo se dejó equilibrando al equipo en la temperatura de medición por lo menos media hora después que no se observaban cambios en la señal térmica del equipo. Los experimentos consistieron en 30 inyecciones de 1  $\mu$ L de PGH 5.1 mM en una solución 0.3 mM de monómero de proteína, además de una inyección inicial sin PGH. Entre cada inyección se dio un intervalo de 300 segundos con una agitación constante a 125 rpm. Posteriormente se hicieron experimentos con la proteína en TEA 100 mM, sacarosa 0.66 molal, pH 7.4 siguiendo el mismo diseño experimental.

Después de la colección de datos, el procesamiento de los mismos se realizó en la versión web del programa *AFFINImiter ITC* de la compañía S4SD (Dumas, 2019; Piñeiro *et al.*, 2019; <https://www.affinimeter.com/site/>) para calcular el  $\Delta H_U$ , la  $K_A$  y la  $n$ . El análisis se llevó a cabo tomando en cuenta un modelo de unión 1:1 entre la proteína (P) y el PGH (L):



El resto de los parámetros asociados a la unión se calcularon con relaciones termodinámicas sencillas. El inverso de la  $K_A$  nos da la constante de disociación ( $K_D$ ). Con el valor de la  $K_A$  podemos obtener un valor de la energía libre de Gibbs de unión ( $\Delta G_U$ ) del proceso con la ecuación 19:

$$\Delta G_U = -R \times T \times \ln K_A \quad (19)$$

A través de la ecuación (20) de Gibbs-Helmholtz se puede determinar la contribución de la entropía de unión ( $\Delta S_U$ ):

$$\Delta G_U = \Delta H_U - T \Delta S_U \quad (20)$$

Finalmente, con los datos termodinámicos obtenidos de estos experimentos a distintas temperaturas se puede obtener el valor del cambio en la capacidad calorífica de unión ( $\Delta C_{p_U}$ ) al graficar el cambio del  $\Delta H$  con la temperatura bajo la siguiente relación (ecuación 21):

$$\Delta C_{p_U} = \frac{\partial \Delta H_U}{\partial T} \quad (21)$$

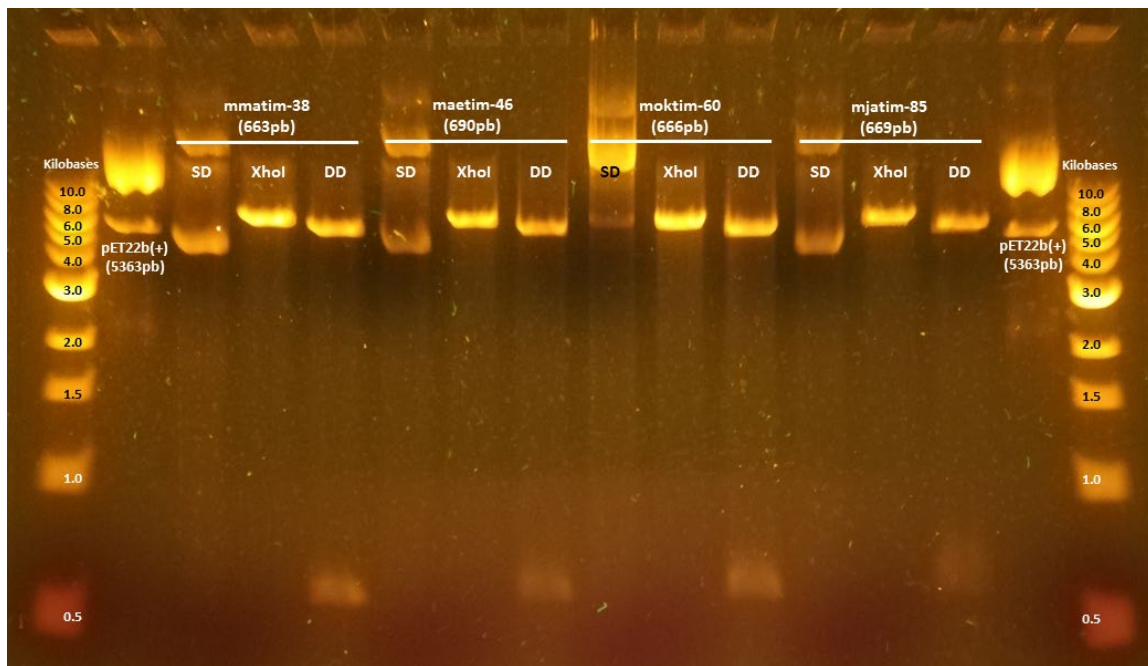
### **Análisis y visualización de datos**

Salvo por los datos de ITC, como se indicó, todos los ajustes se realizaron con el programa Origin v. 7.0383 (OriginLab, s/f-c) y las gráficas se generaron con los paquetes ggplot2 v. 3.3.2 (Wickham, 2016) y ggpubr v. 0.4.0 (Kassambara, 2020) del entorno de programación R v. 4.0.1 (R Core Team, 2020).

## RESULTADOS Y DISCUSIÓN

### **Subclonación de los genes de las ArqTIM**

Previo a este trabajo se había identificado que a MjaTIM-85 no se le puede remover la etiqueta de histidinas con trombina, ya que el sitio de corte parece estar escondido (Muñiz Trejo, 2017). Además, en pruebas en columnas de afinidad HisTrap, la asociación de la etiqueta era muy débil (no había unión a la columna con [imidazol] >5 mM), por lo que los resultados de la purificación no eran satisfactorios sin hacer una precipitación por calor previa de las proteínas de *E. coli*. Asimismo, en este trabajo se encontró que para MmaTIM-38 y MaeTIM-46 ocurre un fenómeno similar (datos no mostrados), por lo que se decidió expresar a todas las ArqTIM sin etiquetas de afinidad. Para ello, con la oportuna capacitación de la Dra. Gloria Hernández Alcántara de la Facultad de Medicina de la UNAM, los genes se subclonaron exitosamente en un vector de sobreexpresión pET22b(+). Por la posición de los sitios de restricción, esta construcción no incluye etiquetas de histidinas alguna. Este vector confiere resistencia a ampicilina y permite la expresión por inducción con el IPTG. Los resultados de la subclonación se muestran en la Figura 8.



**Figura 8. Digestiones de los plásmidos subclonados.** Se muestran los resultados de la subclonación en pET22b(+). Para cada gen se observa el plásmido sin digerir (SD), la versión linearizada con *XhoI* y la doble digestión del inserto (DD). En los extremos se puso marcador de tamaño y vector sin inserto linearizado para la comparación.

Posterior a la subclonación, los genes fueron secuenciados para corroborar que no hubiera mutaciones en la secuencia debido al proceso de intercambio entre vectores. El trabajo de secuenciación fue realizado en el Instituto de Biotecnología de la UNAM por la M. en C. Blanca Ramos Cerrillo y el Dr.

Lorenzo Segovia Forcella. Estos resultados mostraron que no hubo cambios en ninguna de las secuencias de nucleótidos trabajadas. En la Tabla 4 se resumen la disposición final en la que quedaron los insertos.

**Tabla 4. Disposición final de los genes utilizados en este proyecto.**

Inserto	Vector de origen	Vector final	Uso de codones
<i>mmatim-38</i>	pUC57	pET22b(+)	Optimizada para <i>E. coli</i>
<i>maetim-46</i>	pET28b(+)		Secuencia nativa
<i>moktim-60</i>	pET28b(+)		Optimizada para <i>E. coli</i>
<i>mfotim-75</i>	pET22b(+)*		Secuencia nativa
<i>mjatim-85</i>	pET28b(+)		Secuencia nativa

\*No fue necesaria la subclonación.

### **Purificación y caracterización de las MesoTIM**

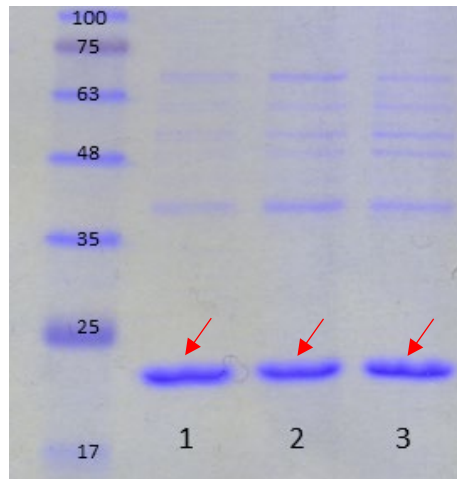
En el trabajo de licenciatura se abordaron diferentes estrategias metodológicas para la sobre-expresión soluble de MaeTIM-46 y MmaTIM-38, como la expresión en diferentes cepas de *E. coli*, la expresión a diferentes temperaturas y la purificación en condiciones desnaturalizantes; sin embargo, en ninguna de las condiciones experimentales probadas hubo expresión de proteínas en forma soluble (*c.fr.* Muñiz Trejo, 2017). Durante el desarrollo de este proyecto se encontró el mismo problema de solubilidad para MokTIM-60, por lo que parece que las TIM de la familia mesófila Methanococcaceae tienen alguna característica común que las hace poco solubles cuando son expresadas heterológicamente. Para solventar este problema, se implementaron un par de estrategias nuevas.

Uno de los métodos más comunes para mejorar la solubilidad de las proteínas que se van a cuerpos de inclusión es la expresión a bajas temperaturas. Esto se basa en el principio de que la tasa de producción de una proteína disminuye conforme baja la temperatura, lo que se refleja positivamente en la calidad de la proteína producida (Schein, 1989; Vera, González-Montalbán, Arís, & Villaverde, 2007). No obstante, esta estrategia no es completamente infalible. Una de las razones es que las chaperonas endógenas de *E. coli* pierden hasta el 70% de su actividad en bajas temperaturas. Para subsanar esta deficiencia, se probó la expresión a 10 °C en células *E. coli* ArticExpress RIL (Agilent Technologies), las cuales incluyen un plásmido con las chaperonas Cpn10 y Cpn60 del psicrófilo *Oleispira antarctica*. Estas chaperonas tienen un alto porcentaje de identidad con el sistema GroEL/ES de *E. coli* y su actividad es óptima entre 4 y 12 °C (Agilent Technologies, 2015). Desafortunadamente, no se detectó expresión soluble con la implementación de esta cepa (datos no mostrados).

Una segunda estrategia fue la co-expresión de chaperonas moleculares utilizando el conjunto de plásmidos de chaperonas de Takara Bio Inc. (2017). Este sistema incluye cinco plásmidos con diferentes combinaciones de chaperonas, de las cuales sólo el plásmido pG-Tf2 transformado en células de *E.coli* BL21-Gold (DE3) presentó resultados positivos en cuanto a solubilidad. pG-Tf2 expresa a las chaperonas citoplasmáticas GroEL/ES y a la chaperona asociada a ribosoma *trigger factor* (TF) en presencia de tetraciclina. Se ha demostrado *in vivo* que la sobre-expresión del TF incrementa la capacidad de GroEL de asociarse a proteínas desplegadas (Kandror, Sherman, Rhode, & Goldberg, 1995), lo que podría explicar por qué sólo en esta condición se obtuvieron versiones solubles de las MesoTIM.

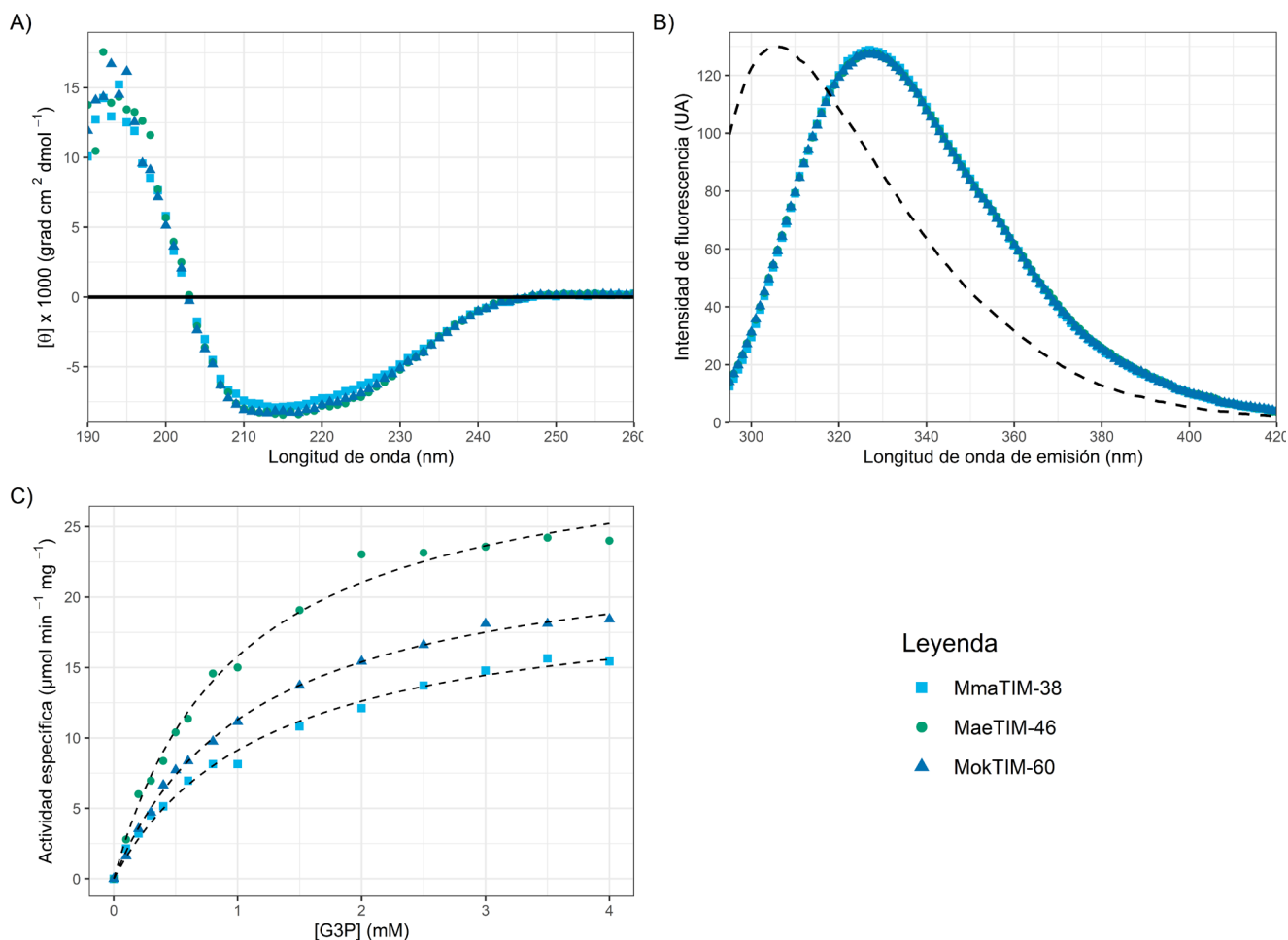
Para la producción de las proteínas se llevaron a cabo varias pruebas con diferentes condiciones de sobre-expresión que implicaron variar la temperatura de crecimiento e inducción (25, 30 y 37 °C), la densidad óptica para la inducción (0.2 a 1), la concentración de IPTG (0.1 a 1 mM), la concentración de tetraciclina (2 a 10 ng mL<sup>-1</sup>) y el tiempo de inducción (4 a 16 horas); a pesar de ello, no se pudo

obtener una producción total mayor a los 3 ó 4 mg por litro de cultivo. Aunado a la baja producción, la purificación representó un reto aún mayor, ya no que no pudimos obtener muestras de alta pureza para estas proteínas (ver Figura 9).



**Figura 9. Purificación de las MesoTIM.** Gel de SDS-PAGE de las proteínas seleccionadas. En el primer carril las bandas corresponden a los marcadores de masa molecular. 1) MmaTIM-38. 2) MaeTIM-46. 3) MokTIM-60. Las flechas rojas indican las proteínas de interés. Las condiciones de purificación para estas muestras particulares se describieron en los métodos.

Se probaron diferentes estrategias para eliminar los contaminantes. Por cromatografía de exclusión se utilizaron columnas Superdex-75 y Superdex-200 preparativas, pero en ningún caso se observó que la pureza de la muestra incrementara. Dado que el punto isoeléctrico de estas proteínas es de alrededor de 5 (ver Tabla 2), se ensayó por cromatografía de intercambio aniónico la interacción con una columna HiTrapQ HP a pH 7.0 y 8.5. Aunque en estos casos sí se redujo el número de contaminantes, no se lograron obtener proteínas con una pureza mayor a la observada en la Figura 9. Pese a ello, se procedió a analizar algunas de sus propiedades espectroscópicas y funcionales para corroborar si las proteínas expresadas se encontraban en condiciones para ser caracterizadas y si se podía extraer información pese a la presencia de contaminantes.



**Figura 10. Caracterización espectroscópica y funcional de las MesoTIM.** Todos los experimentos se realizaron a 25 °C. A) Espectros de dicroísmo circular en el UV lejano. B) Espectros de fluorescencia intrínseca, longitud de onda excitación 278 nm. En líneas discontinuas un espectro típico de una proteína sin triptófanos. C) Actividad catalítica de las proteínas. En líneas discontinuas se observan curvas estimadas de la ecuación de Michaelis-Menten. Los experimentos de actividad se llevaron a cabo con 150 ng totales de proteína.

Para todos los casos se analizó la estructura secundaria por dicroísmo circular en UV lejano y la estructura terciaria por fluorescencia intrínseca de las tirosinas. En el caso del dicroísmo circular, los espectros obtenidos (Figura 10A) resultaron ser congruentes con lo esperado para una proteína con topología de barril ( $\beta/\alpha$ )<sub>8</sub>, pero los estimados de porcentaje de estructura no correspondían a lo esperado para una TIM (Tabla 5). De acuerdo con el análisis de deconvolución (Figura S3), las MesoTIM tienen un ~20% de hélices  $\alpha$  y un ~30% de hojas  $\beta$ . Estos resultados son contrastantes con lo que se había observado previamente para MjaTIM-85 con etiqueta de histidinas, donde encontramos un ~30% de hélices  $\alpha$  y un 20% de hojas  $\beta$ . Pese a la diferencia, esta observación por sí sola no necesariamente implica que haya un problema mayor con el plegamiento de estas proteínas, pero junto a los datos de fluorescencia algunos problemas por la presencia de contaminantes se hacen presentes.

El escenario más común para proteínas sin triptófanos es que los espectros de fluorescencia tengan un máximo de emisión alrededor de 303 nm debido a la contribución de las tirosinas (ejemplificado con líneas discontinuas en la figura). Sin embargo, como se observa en la Figura 10B y en la Tabla 5, los espectros de las MesoTIM se parecen más a los de proteínas con triptófanos, cuyo máximo de emisión está corrido hacia ~327 nm.

**Tabla 5. Parámetros espectroscópicos y funcionales de las MesoTIM.** Los parámetros se muestran por método.

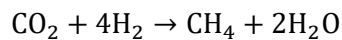
Parámetro	MmaTIM-38	MaeTIM-46	MokTIM-60	MjaTIM-85 con HisTag*
<i>Deconvolución de los espectros de dicroísmo</i>				
Hélices $\alpha$	20%	22%	22%	33.7%
Hojas $\beta$	30%	29%	30%	19.7%
Estructura no repetitiva	49%	48%	48%	46.6%
<i>Espectros de fluorescencia</i>				
Intensidad máx (UA)	128.8	127.6	127.4	442.7
$\lambda$ máx (nm)	327	326	328	302
<i>Parámetros catalíticos</i>				
$V_{\text{máx}}$ ( $\mu\text{mol min}^{-1} \text{mg}^{-1}$ )	20.41	31.47	24.15	2115.2
$K_M$ (mM)	1.23	0.99	1.14	0.19
$k_{\text{cat}}$ ( $\text{min}^{-1}$ )	$4.67 \times 10^2$	$7.55 \times 10^2$	$5.67 \times 10^2$	$5.36 \times 10^4$
Eficiencia catalítica ( $\text{M}^{-1} \text{s}^{-1}$ )	$1.84 \times 10^5$	$3.69 \times 10^5$	$2.4 \times 10^5$	$1.4 \times 10^8$

Una posible explicación de este fenómeno es la formación de tirosinatos. Los tirosinatos son la forma ionizada de la porción p-hidroxifenil de los residuos de tirosina, cuya longitud máxima de emisión ( $\lambda$ ) es mayor a los 310 nm. La presencia de éstos se debe a la transferencia de un protón del anillo fenólico a un aceptor cercano, usualmente la cadena lateral de otro aminoácido vecino. Es posible determinar experimentalmente la formación de tirosinatos si se capturan espectros de fluorescencia con las proteínas inmersas un medio ácido ( $\text{pH} < 1$ ), lo cual ayuda a mantener las tirosinas protonadas ( $\text{pKa}$  de las tirosinas = 10.07) y a que los espectros recuperen su  $\lambda$  hacia los 303 nm (Prendergast, Hampton, & Jones, 1984). Esta prueba fue realizada en las tres MesoTIM, pero no se observó ningún corrimiento total ni parcial de los espectros de emisión (datos no mostrados). Lo anterior demuestra que, sin lugar a duda, hay triptófanos están contribuyendo a la señal espectral, probablemente derivados de los contaminantes presentes, los cuales tienen un rendimiento cuántico mucho mayor al de las tirosinas. Lo anterior, junto a la discordancia en los datos de dicroísmo, es indicativo de que los contaminantes proteicos están deformando los espectros y no nos permiten obtener una imagen clara al respecto del plegamiento de estas proteínas, por lo que se requiere mejorar los protocolos de purificación.



Se obtuvieron también medidas de la actividad catalítica de las MesoTIM (Figura 10C) y los parámetros catalíticos de estas TIM están muy por debajo de lo esperado para una proteína de este tipo (Tabla 5). En la literatura se encuentran descritas varias TIM que no se consideran “catalíticamente perfectas” como las de *Helicobacter pylori* (Chu, Lai, Huang, & Sun, 2008), *Oryctolagus cuniculus* (Krietsch, Pentchev, Klingenburg, Hofstätter, & Bücher, 1970) y *Plasmodium falciparum* (Joubert, Neitz, & Louw, 2001), pero su eficiencia catalítica nunca es menor a los  $10^7 \text{ M}^{-1} \text{ s}^{-1}$ . Las MesoTIM, en contraparte, están tres órdenes de magnitud más abajo del límite de difusión ( $10^8 \text{ M}^{-1} \text{ s}^{-1}$ ).

Aunque es un comportamiento atípico, el que las MesoTIM no tengan una actividad elevada puede ser consecuencia de un proceso biológico, por lo que en este punto se vale la pena hacer una pequeña consideración acerca del metabolismo de los Methanococcales. En estos organismos, la obtención de energía se da principalmente por metanogénesis a través del ciclo de Wolfe (Thauer, 2012):



En cuanto a su metabolismo del carbono, datos bioinformáticos indican que estas arqueas carecen de los transportadores de hexosas y pentosas requeridos para incorporar carbohidratos simples al medio intracelular. Los Methanococcales, en cambio, utilizan la gluconeogénesis para almacenar el carbono que obtienen del  $\text{CO}_2$  y del acetato en forma de glucógeno.<sup>6</sup> Cuando las condiciones son adversas, la degradación del glucógeno y la glucólisis se convierten en la fuente de energía celular (Goyal, Zhou, & Karimi, 2016). De acuerdo con algunas hipótesis, el que una enzima en particular mantenga cierto nivel de actividad puede deberse a que existe una presión selectiva que mantiene la actividad por encima de cierto umbral donde no se afecta la adecuación de la especie, por lo que si dicha presión selectiva se relaja, la enzima puede perder eficiencia catalítica (Newton, Arcus, & Patrick, 2015). De esta manera, se podría especular que el metabolismo del carbono de los Methanococcales tuvo repercusiones diferentes en la evolución de las enzimas glucolíticas de estos organismos y que, por tanto, no sería del todo extraño encontrar que la actividad de la TIM sea distinta —menor en este caso—.<sup>7</sup> No obstante, como se verá en el siguiente capítulo, la actividad baja parece corresponder

---

<sup>6</sup> Tanto el  $\text{CO}_2$  como el acetato se pueden utilizar para sintetizar acetil-CoA. Está última molécula puede ser convertida en piruvato para iniciar la gluconeogénesis. En el caso del  $\text{CO}_2$ , la reacción es catalizada por el complejo CO deshidrogenasa/acetil-CoA sintasa (EC 1.2.7.4), el cual primero reduce esta molécula a CO y luego la une a un grupo metilo y a la CoA (Shieh & Whitman, 1988). Por otro lado, el acetato es incorporado del medio extracelular y catalizado por la acetato-CoA ligasa (EC 6.2.1.1) con el consumo de una molécula de ATP (Shieh & Whitman, 1987).

<sup>7</sup> Al respecto, en el caso de otras enzimas relacionadas con la glucólisis/gluconeogénesis en *Methanococcus maripaludis* se han descrito diferencias en la actividad que favorecen la gluconeogénesis. Por ejemplo, se sabe que la actividad de la fructosa-1,6-bisfosfatasa (EC 3.1.3.11) de este organismo es de casi el doble que la actividad de la 6-fosfofructocinasa (EC 2.7.1.11). Además, enzimas con actividad reversible como la fosfoglucomutasa (EC 5.4.2.8), la glucosa-6-fosfato isomerasa

únicamente a las MesoTIM, no a Methanococcales en su conjunto, por lo que el metabolismo de estos organismos no parece influir en ello. De igual manera, el que MokTIM-60, de origen termófilo, tenga baja actividad indica que la temperatura de crecimiento del organismo no está relacionada directamente con este cambio. De este modo surgen dos posibilidades, la primera es que la actividad disminuida tenga un origen no adaptativo y la segunda es que estas proteínas no estén plegadas correctamente. No se puede descartar, hasta este punto, que las chaperonas estén formando agregados solubles porque la presencia de contaminantes no nos permite inferir el estado del plegamiento de las MesoTIM.

En cuanto a la purificación, es necesario encontrar estrategias para incrementar la cantidad y asegurar la calidad de estas proteínas. Para ello, se sugiere la adición de “chaperonas químicas” que ayudan a la actividad de GroEL/ES como glicerol, prolina y trehalosa, las cuales se pueden agregar al medio de cultivo para incrementar la cantidad de proteína y la calidad de las mismas (Diamant, Eliahu, Rosenthal, & Goloubinoff, 2001). Por otro lado, la implementación de otros métodos de purificación, como la interacción hidrofóbica, podría ayudar a incrementar la pureza final de las muestras.

Una vez establecido un protocolo de purificación efectivo, será necesario discernir si la actividad baja es de origen biológico o es un problema de expresión heteróloga. Primero con resultados actualizados de dicroísmo y fluorescencia, y segundo, se puede cuantificar la calidad de una muestra de proteína a través de la calorimetría de titulación isotérmica (ITC). Esta técnica permite cuantificar la concentración de enzima activa —adecuadamente plegada— en una muestra pura a partir de la estequiometría observada en estos ensayos, ya que ésta indica la proporción de proteína capaz de unir al ligando (Baranauskiene *et al.*, 2019; Oliveira & Domingues, 2018). Así, la ITC se puede emplear como un método auxiliar para establecer si hay problemas con el plegamiento, concretamente, con la conformación del sitio activo.

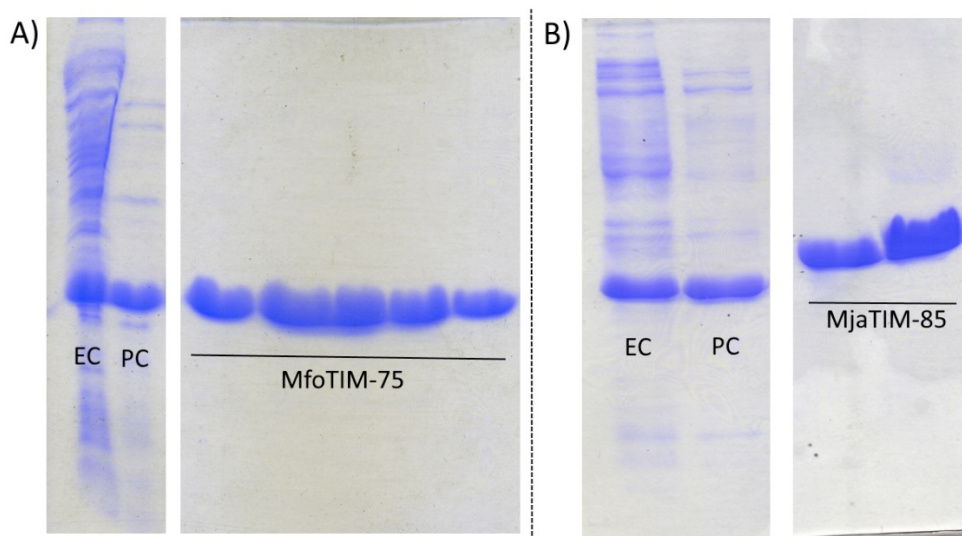
Debido a que no se pudo concluir con la caracterización de las MesoTIM por los problemas expuestos, el objetivo general de este trabajo no se pudo completar cabalmente. Sin embargo, la discusión de las siguientes secciones se centrará en cómo los resultados obtenidos al caracterizar a las TermoTIM nos permiten discutir las ideas actuales sobre la evolución de proteínas y cómo éstas nos sirven para proponer futuras líneas de investigación.

---

(EC 5.3.1.9) y la fructosa-1,6-bisfosfato aldolasa (EC 4.1.2.13) tienen mayor actividad en la dirección anabólica (Yu, Ladapo, & Whitman, 1994).

### ***Purificación y caracterización espectroscópica y funcional de las TermoTIM***

Para MfoTIM-75 y MjaTIM-85, ambas sin etiqueta de histidinas, se establecieron protocolos de purificación de los cuales se obtuvieron muestras de alta pureza (Figura 11). Estas proteínas fueron solubles sin necesidad de co-expresar chaperonas, por lo que sólo hubo que ajustar las condiciones de expresión en cuanto a la densidad óptima de inducción y la concentración de IPTG (ver Métodos). Ambas proteínas provienen de organismos termófilos, por lo que es posible dar un tratamiento de precipitación por calor para eliminar algunos de los contaminantes presentes. Finalmente, dada la ausencia de etiquetas de afinidad, se implementó una precipitación por sulfato de amonio como paso inicial de purificación y para concentrar las proteínas antes del paso cromatográfico. MfoTIM-75 precipita entre 60 y 90% de saturación de amonio, mientras que MjaTIM-85 lo hace entre 50 y 80%.

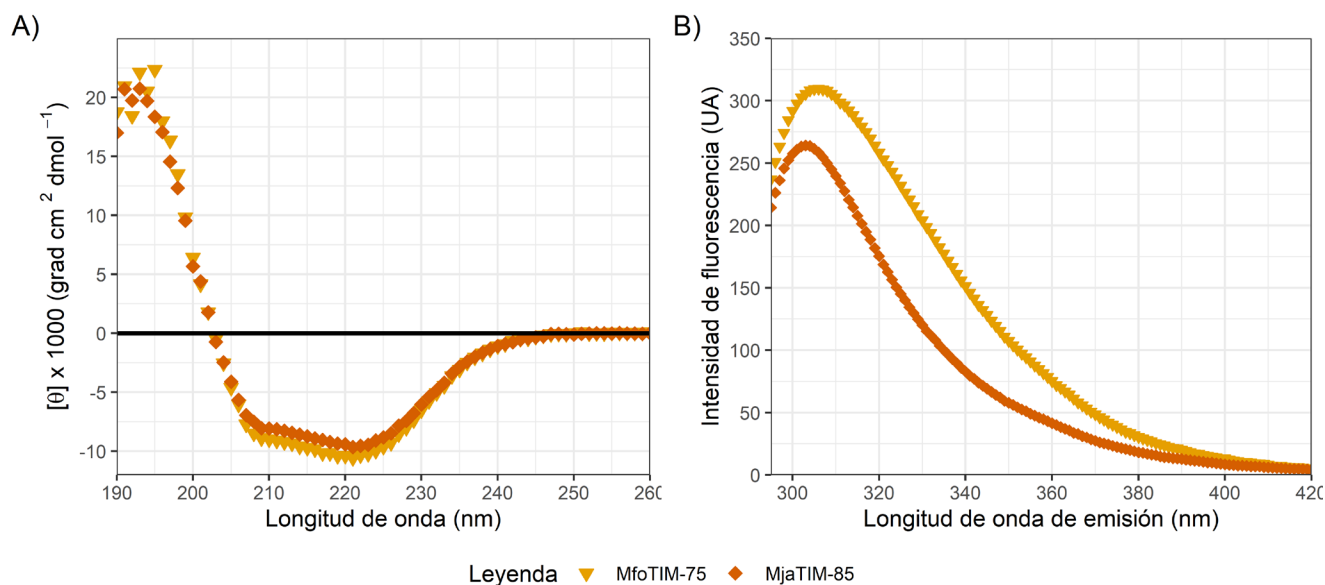


**Figura 11. Purificación de las TermoTIM.** Geles SDS-PAGE de los pasos de purificación A) MfoTIM-75 después de HiTrap Q a pH7. B) MjaTIM-85 después de HiTrap Q a pH 8.5. EC. Extracto crudo. PC. Precipitación por calor.

Ambas TermoTIM se separaron en una columna de intercambio aniónico HiTrap Q HP a pH 7.0. MjaTIM-85 necesitó un paso adicional a pH 8.5 para complementar la purificación. Esto probablemente debido a que su precipitación con sulfato de amonio lleva más impurezas en el intervalo de saturación. En ambos casos, la purificación nos dio en promedio una producción total de proteína pura por litro de medio de cultivo entre 20 y 25 mg totales. Con estas condiciones se pudo proseguir con la caracterización estructural y funcional de ambas proteínas.

### Caracterización de estructuras secundaria y terciaria

Se realizó una caracterización espectroscópica para determinar el estado del plegamiento de las TermoTIM. Se evaluaron la estructura secundaria y terciaria a 25 °C (Figura 12) y se analizaron ambos espectros para obtener los parámetros correspondientes (Tabla 6, revisar Figura S3 para las deconvoluciones correspondientes).



**Figura 12. Espectroscopía de las TermoTIM.** Los espectros se obtuvieron con 0.4 mg mL<sup>-1</sup> de proteína correspondiente en amortiguador de fosfatos de sodio 10 mM, pH 8.0 a 25°C. A) Elipticidad promedio por residuo en el DC UV lejano. B) Intensidad de fluorescencia de las tirosinas al excitar con una longitud de onda de 278 nm.

**Tabla 6. Parámetros de los espectros de DC e IF de las TermoTIM.** Entre paréntesis se indica el porcentaje de estructura secundaria descrito para la estructura cristalográfica de MjaTIM-85 (PDB: 2H6R).

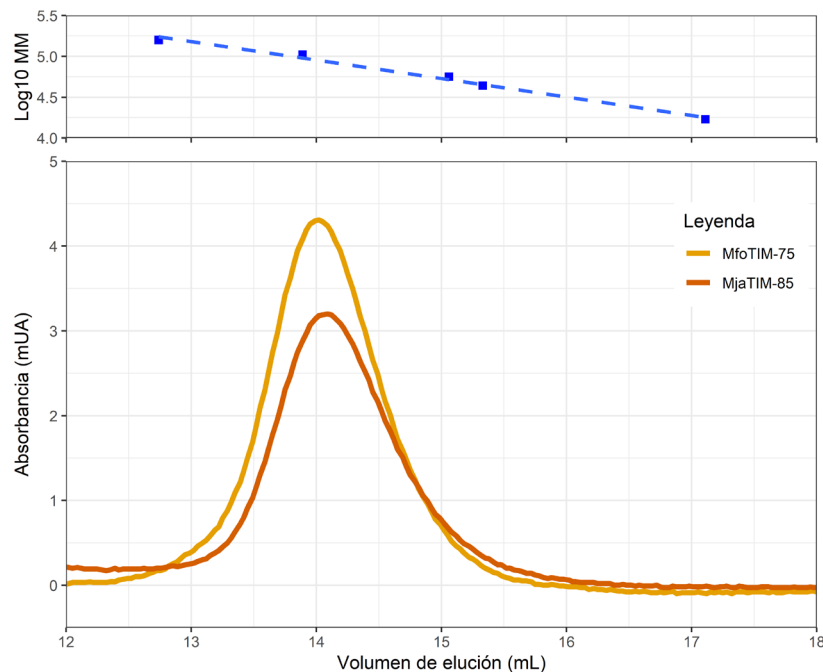
Estructura secundaria	MfoTIM-75	MjaTIM-85
<i>Deconvolución de los espectros de dicroísmo</i>		
Hélices $\alpha$	33%	38% (36)
Hojas $\beta$	22%	18% (17)
Estructura no repetitiva	44%	45% (47)
<i>Espectros de fluorescencia</i>		
Intensidad máx (UA)	309.8	264.1
$\lambda$ máx (nm)	303	302

El análisis de deconvolución de los espectros de DC son congruentes con una proteína  $\alpha/\beta$  como la TIM, con mínimos esperados entre 222 y 208 nm, así como un máximo cercano a 195 nm. Además, el análisis de deconvolución arroja que los porcentajes de estructura secundarias sí corresponden con lo

observado en la la estructura cristalográfica de la MjaTIM-85 (PDB: 2H6R). En cuanto a su estructura terciaria, los espectros de fluorescencia poseen el máximo esperado en 303 nm, típico para proteínas carentes de triptófanos. Esto nos indica que ambas enzimas están bien empaquetadas y tienen una topología congruente con otras TIM.

### **Determinación del estado oligomérico**

Se estimó el estado oligomérico utilizando cromatografía de exclusión molecular con una columna Superdex 200 10/300 GL. La calibración se hizo utilizando los estándares de filtración en gel de Bio-Rad (s.f.): mioglobina (17 KDa), ovoalbúmina (44 KDa) y  $\gamma$ -globulina (158 KDa), así como muestras de la MjaTIM con etiqueta de histidinas (105 KDa; Muñiz Trejo, 2017) y la TIM de *Clostridium perfringens* (CpTIM; 56 KDa; Romero Romero, 2017) a manera de tener en la calibración una TIM tetramérica y una dimérica, respectivamente (ver Figura 13).



**Figura 13. Cromatografía de exclusión molecular analítica de las ThermoTIM.** En el panel inferior se observa la absorbancia de muestras a  $0.5 \text{ mg mL}^{-1}$  de cada proteína en amortiguador de fosfato de sodio 50mM, NaCl 200mM, pH 8.0. En el panel superior se muestra la curva de calibración. Estándares de calibración, de izquierda a derecha:  $\gamma$ -globulina, MjaTIM con HisTag, CpTIM, ovoalbúmina y mioglobina. En línea discontinua el ajuste lineal de la calibración,  $R^2=0.991$ .

Para ambas proteínas se hicieron cromatogramas a diferentes concentraciones: 0.5, 0.1 y  $0.05 \text{ mg mL}^{-1}$ , con el fin de determinar si había algún efecto de la concentración en el estado oligomérico (aunque sólo se muestran los datos de  $0.5 \text{ mg mL}^{-1}$  en la Figura 13 para simplificar). El volumen de elución no se modificó en el intervalo de concentraciones evaluadas ni surgieron nuevos picos, por lo que no

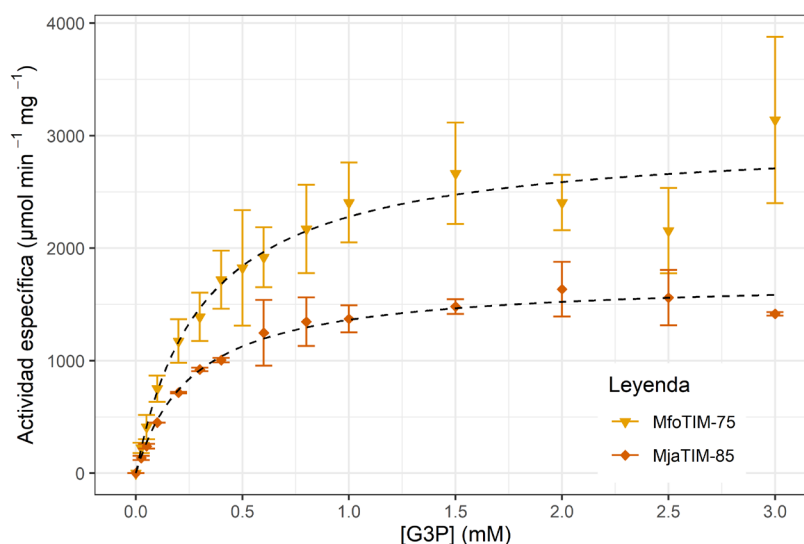
parece haber cambios en el estado de asociación. Cabe mencionar que, debido a la ausencia de triptófanos, no se ensayaron concentraciones menores de proteína porque la señal no se distingue del ruido del detector de absorbancia del equipo. A partir de la curva de calibración se determinaron las propiedades hidrodinámicas de las TermoTIM, como su masa molecular experimental y su radio de Stokes (Tabla 7). Estos resultados nos indican que las TermoTIM son tetraméricas y que su estado oligomérico no es dependiente de concentración en el intervalo evaluado.

**Tabla 7. Determinación de las propiedades hidrodinámicas por A-SEC.**

TIM	Masa molecular teórico del monómero (kDa)	Volumen de elución (mL)	Radio de Stokes (Å)	Masa molecular experimental (kDa)	Estado oligomérico
<i>Mfo</i> TIM-75	23.50	13.92	40.95	89.18	Tetramero (3.8)
<i>Mja</i> TIM-85	23.32	13.92	40.72	87.82	Tetramero (3.8)

### Obtención de parámetros catalíticos

Finalmente se caracterizaron las propiedades catalíticas de estas proteínas. Para ello, se evaluaron cinéticas a 25 °C con diferentes concentraciones de G3P, como se observa en la Figura 14. La razón de elegir esta temperatura fue que así se podían comparar con datos de otras TIM y a que no nos era posible realizar ensayos en la temperatura óptima de crecimiento de los organismos debido a que la enzima acoplada, la  $\alpha$ -glicerol-3-fosfato deshidrogenasa, se inactiva por encima de los 65 °C (Merck, 2020). Los datos se ajustaron a la ecuación de Michaelis-Menten y se obtuvieron los parámetros catalíticos como se aprecia en la Tabla 8.



**Figura 14. Cinéticas enzimáticas de las TermoTIM a 25 °C.** Se midió la actividad catalítica a distintas concentraciones de G3P. Los ensayos se realizaron en amortiguador TED 100:10:1, pH 7.4. Concentración de proteína por punto: 5 ng mL<sup>-1</sup>. Cada punto representa el promedio de tres lotes de purificación diferentes y las barras de error corresponden la desviación estándar.

**Tabla 8. Parámetros catalíticos de las TermoTIM a 25 °C.** Se incluyen en la tabla para la comparación los parámetros de las TIM del hipertermófilo *Pyrococcus furiosus* (PfuTIM, Sharma & Guptasarma, 2015), del mesófilo *Clostridium perfringens* (CpTIM, Romero Romero, 2017) y del psicrófilo *Vibrio marinus* (VmTIM, Alvarez *et al.*, 1998, los datos se obtuvieron a 10 °C)

Parámetro	MfoTIM-75	MjaTIM-85	PfuTIM	CpTIM	VmTIM
$V_{\text{máx}}$ ( $\mu\text{mol min}^{-1} \text{mg}^{-1}$ )	2990 $\pm$ 159	1723 $\pm$ 47	61940	13930	--
$K_M$ (mM)	0.21 $\pm$ 0.06	0.26 $\pm$ 0.03	1.22	0.72	1.9
$k_{\text{cat}}$ ( $\text{min}^{-1}$ )	(7.0 $\pm$ 0.4) $\times 10^4$	(3.67 $\pm$ 0.1) $\times 10^4$	9.29 $\times 10^7$	3.81 $\times 10^5$	4.2 $\times 10^5$
Eficiencia catalítica ( $\text{M}^{-1} \text{s}^{-1}$ )	(1.6 $\pm$ 0.3) $\times 10^8$	(7.5 $\pm$ 0.2) $\times 10^7$	1.27 $\times 10^9$	2.24 $\times 10^8$	1.32 $\times 10^7$

Estos datos nos demuestran que ambas enzimas tienen actividad de triosafosfato isomerasa y que su eficiencia catalítica está cerca o sobre el límite de difusión ( $10^8 \text{M}^{-1} \text{s}^{-1}$ ), lo cual las hace equiparables con casi cualquier otra TIM mesófila reportada a 25 °C.

### ***Función, estabilidad y evolución de proteínas termófilas, parte I***

El que las TermoTIM tengan actividad elevada a temperatura ambiente es contrastante con lo que usualmente se piensa sobre la relación entre estabilidad, función y temperatura. De manera general, se considera que las enzimas termófilas presentan una mayor rigidez que sus homólogas mesófilas cuando éstas se incuban a una temperatura baja (*i.e.* 20-37 °C). Esta rigidez excesiva se manifiesta a través de la pérdida de actividad catalítica por debajo de la temperatura fisiológica del organismo de procedencia (Vieille & Zeikus, 2001). Una proteína muy estable (o muy rígida en este contexto) no será capaz de realizar su función en una tasa metabólica relevante, ya que no puede realizar los rearrreglos estructurales necesarios. En contraparte, la catálisis será impedida si la proteína es muy flexible, ya que no será capaz de mantener su estructura nativa por mucho tiempo, ni de establecer la unión correcta con el ligando. Este razonamiento implica que existe un balance entre flexibilidad y estabilidad, y este balance implica un estrecho vínculo entre estructura y actividad que se considera regularmente como una adaptación a la temperatura de los organismos ectotermos, la cual consiste en cambiar sutilmente las interacciones que modulan la flexibilidad y la rigidez (Dong, Liao, Meng, & Somero, 2018; Fields, 2001; Fields, Dong, Meng, & Somero, 2015; Somero, 1995).

Las TermoTIM no cumplen esta característica, pues su actividad no está acotada a un intervalo pequeño por debajo de su valor de termoestabilidad, sino son activas en un intervalo muy amplio de temperatura. Además de nuestros datos, se ha reportado que las TIM de *Pyrococcus furiosus* y *Thermococcus onnurineus* tienen una eficiencia catalítica sobre el límite de difusión ( $10^9 \text{M}^{-1} \text{s}^{-1}$ ) a 25 °C (Sharma & Guptasarma, 2015), por lo que la eficiencia catalítica elevada no parece ser una característica única de las TIM estudiadas y, por tanto, se puede considerar como una característica antigua en las TIM del

dominio Archaea, lo mismo que el estado oligomérico. Además, la versión con histidinas de la MjaTIM-85 que se estudió en la licenciatura también poseía una eficiencia catalítica sobre el límite de difusión (Muñiz Trejo, 2017). Al respecto, para ampliar esta discusión sobre la rigidez y la flexibilidad es necesario primero explorar la estabilidad de las TermoTIM, de lo cual ahondaremos en el capítulo siguiente.



### ***Caracterización de la estabilidad y vías de plegamiento***

Para que una proteína pueda desempeñar su función biológica es necesario que ésta sea estable a las condiciones de su entorno, llámese ello pH, temperatura o fuerza iónica. La estabilidad de una proteína se puede evaluar ya sea desde una perspectiva termodinámica o desde una perspectiva cinética. Desde el punto de vista termodinámico, la función está garantizada por el equilibrio químico que existe entre el estado nativo y el estado desplegado. En otras palabras, nos dice que la forma nativa de una proteína es la forma favorecida de manera espontánea (léase  $\Delta G < 0$ ) en ciertas condiciones ambientales y, por tanto, su estabilidad puede ser descrita bajo los preceptos de la termodinámica clásica. Cuando hablamos de una estabilidad cinética, nos referimos a que una proteína mantiene el estado nativo bajo ciertas condiciones por un determinado periodo de tiempo, presumiblemente el suficiente para llevar a cabo su función en una escala fisiológicamente relevante para el organismo. Esto se debe a que hay una barrera de energía libre (la energía de activación,  $E_A$ ) que separa al estado nativo de sus formas no funcionales. El poseer un determinado tipo de estabilidad no elimina la posibilidad de tener la otra de manera categórica, de hecho, existen proteínas que detentan ambas y otras no (Romero-Romero *et al.*, 2018; Sanchez-Ruiz, 2010).

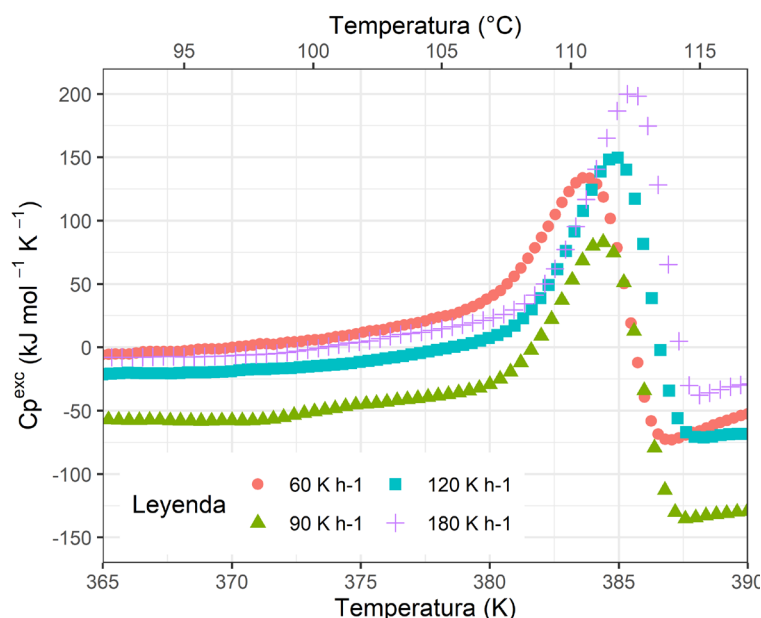
En esta sección, los experimentos realizados buscan determinar precisamente cuál es la naturaleza de la estabilidad de las TermoTIM. Para ello, se analizó el desplegamiento provocado por temperatura, el cual también nos permite evaluar la termoestabilidad descrita como el valor la temperatura media de desnaturalización ( $T_m$ ), y por agentes químicos como la urea y el clorhidrato de guanidinio (GdnHCl). Asimismo, estos métodos nos permiten proponer posibles vías de plegamiento que siguen las enzimas.

#### ***Desnaturalización por temperatura***

Experimentalmente, es posible trabajar con la desnaturalización térmica seguida por dicroísmo circular (DC) a distintas longitudes de onda. Sin embargo, la temperatura media de desnaturalización de ambas proteínas, como se verá a continuación, es demasiado elevada para los límites técnicos del espectropolarímetro utilizado en este trabajo, por lo que todos los experimentos de este tipo se realizaron únicamente por calorimetría diferencial de barrido (DSC).

#### ***Análisis de los datos bajo control cinético***

En el trabajo previo de licenciatura (Muñiz Trejo, 2017), se había encontrado que MjaTIM-85 con etiqueta de histidinas poseía un desplegamiento térmico irreversible controlado cinéticamente que podía ser analizado bajo un modelo irreversible de dos estados y también presentaba agregación posterior a la transición. La versión sin etiqueta de histidinas presenta un comportamiento similar. Los termogramas de desnaturalización correspondientes se muestran en la Figura 15.



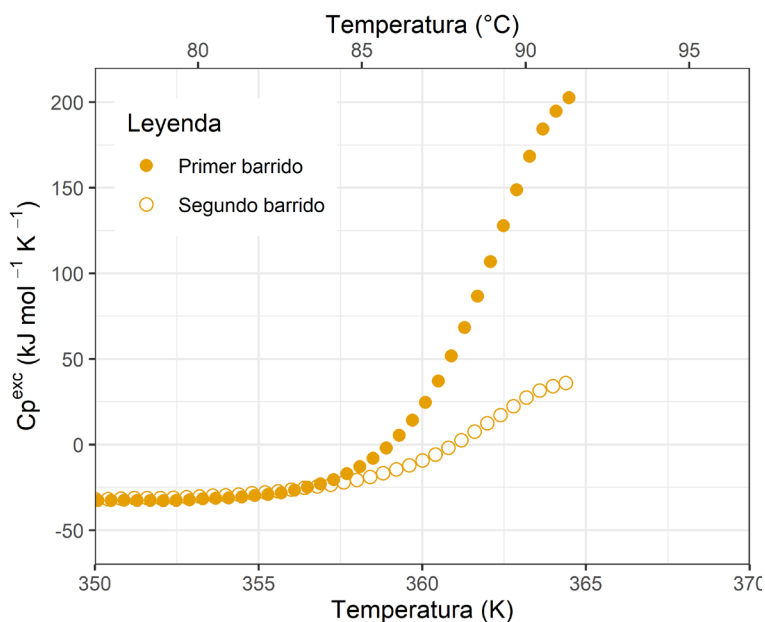
**Figura 15. Desplegamiento térmico de MjaTIM-85 a distintas velocidades de barrido.** Se muestra la región de las transiciones de las trazas calorimétricas con  $0.4 \text{ mg mL}^{-1}$  de la proteína en amortiguador de fosfatos  $10 \text{ mM}$ ,  $\text{pH } 8.0$  evaluadas a diferentes velocidades de barrido.

La  $T_m$  obtenida para esta proteína fue de  $\sim 112 \text{ }^\circ\text{C}$  (mismo valor que con histidinas, ver Tabla 9). En estos ensayos se muestran los resultados de los experimentos a diferentes velocidades de barrido, con los cuales se encontró que existe cierta dependencia de la  $T_m$  con la velocidad de barrido. La diferencia entre la  $T_m$  más elevada y la más baja fue de  $1.7 \text{ }^\circ\text{C}$ . Además, si se observa la Figura, es muy ostensible que la línea base de la postransición —que se interpreta como la capacidad calorífica ( $C_p$ ) del estado desnaturalizado— se encuentra por debajo de la línea base de la pretransición —que se asume como la  $C_p$  del estado nativo.

Regularmente, cuando se rompe el núcleo hidrofóbico de una proteína durante la desnaturalización, la exposición de los grupos no polares hacia el solvente resulta en la estructuración del agua alrededor de dichos grupos, lo cual incrementa la  $C_p$  del estado desplegado con respecto a la  $C_p$  del estado nativo, por lo cual se espera que el  $\Delta C_p$  de desplegamiento sea siempre positivo. Sin embargo, la magnitud del  $\Delta C_p$  observado puede variar respecto a lo esperado si el estado desplegado mantiene estructura residual, es decir, si mantiene grupos no polares escondidos, o bien, si la muestra agrega tan pronto se despliega. En este caso, debido al  $\Delta C_p$  negativo, podemos asumir que hay agregación (Jelesarov & Bosshard, 1999). Además, por el intervalo de temperatura que se maneja, podemos conjeturar que los residuos de aminoácidos, como asparagina, glutamina y cisteína, y algunos enlaces peptídicos podrían pasar por modificaciones químicas irreversibles (Vieille & Zeikus, 2001). Por estas razones no se presenta reversibilidad alguna en el proceso de desplegamiento de esta proteína. A

partir de estos datos se hizo un análisis de desplegamiento bajo un modelo irreversible de dos estados que se resume en la Figura 18 y Tabla 9.

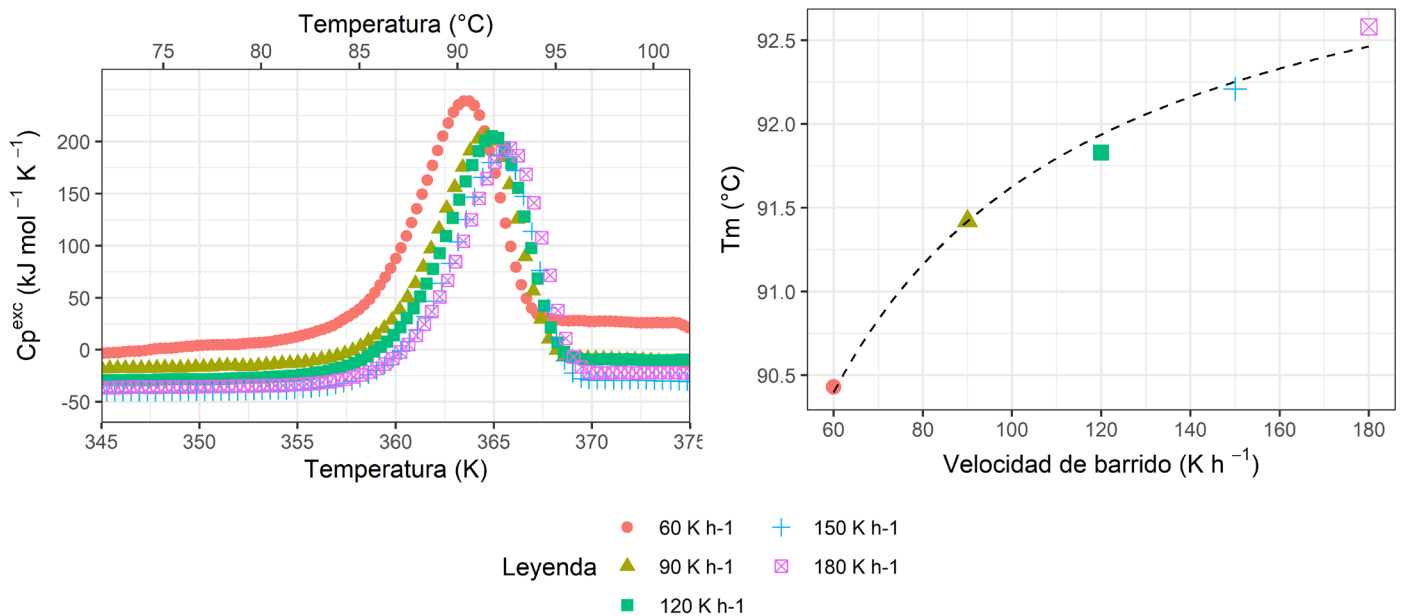
En cuanto al análisis del desplegamiento térmico de MfoTIM-75, se determinó que la  $T_m$  tiene un valor de  $\sim 92^\circ\text{C}$  (ver Tabla 9). Con esta información se realizaron diversos experimentos de reversibilidad (ver Figura 16).



**Figura 16. Prueba de reversibilidad por temperatura de MfoTIM-75.** Se evaluó la reversibilidad de esta proteína al tomar dos trazas consecutivas de la misma muestra después de enfriar. Velocidad de barrido  $90\text{ K h}^{-1}$  con  $0.4\text{ mg mL}^{-1}$  de proteína en amortiguador de fosfato de sodio  $10\text{ mM}$ ,  $\text{pH } 8.0$ .

Para probar la reversibilidad por temperatura se apeló al principio de repetibilidad, el cual consiste en hacer un barrido hasta llegar a cierta temperatura desnaturizante, enfriar la muestra y volver a calentar a la misma temperatura. Si la muestra es completamente reversible, se espera obtener la misma traza calorimétrica en el primer barrido que en el segundo (Ibarra-Molero & Sanchez-Ruiz, 2006). Para ello se realizaron tres experimentos en los cuales la temperatura final fue diferente: 1) al llegar a la postransición, 2) al llegar a la mitad de la altura del pico después de la  $T_m$  y 3) al llegar a la  $T_m$  (en la Figura sólo se muestra esta última condición). En todos los escenarios se observó una disminución en la señal con respecto a la primera traza. La mayor recuperación se presenta al llegar a la  $T_m$ , con cerca del 30% del área inicial, lo que indica que hay una pequeña porción de la muestra que sí se repliega, pero esto no se considera suficiente para admitir reversibilidad “operacional” en el proceso de desnaturización térmica. Por lo tanto, para propósitos prácticos, se considera como un

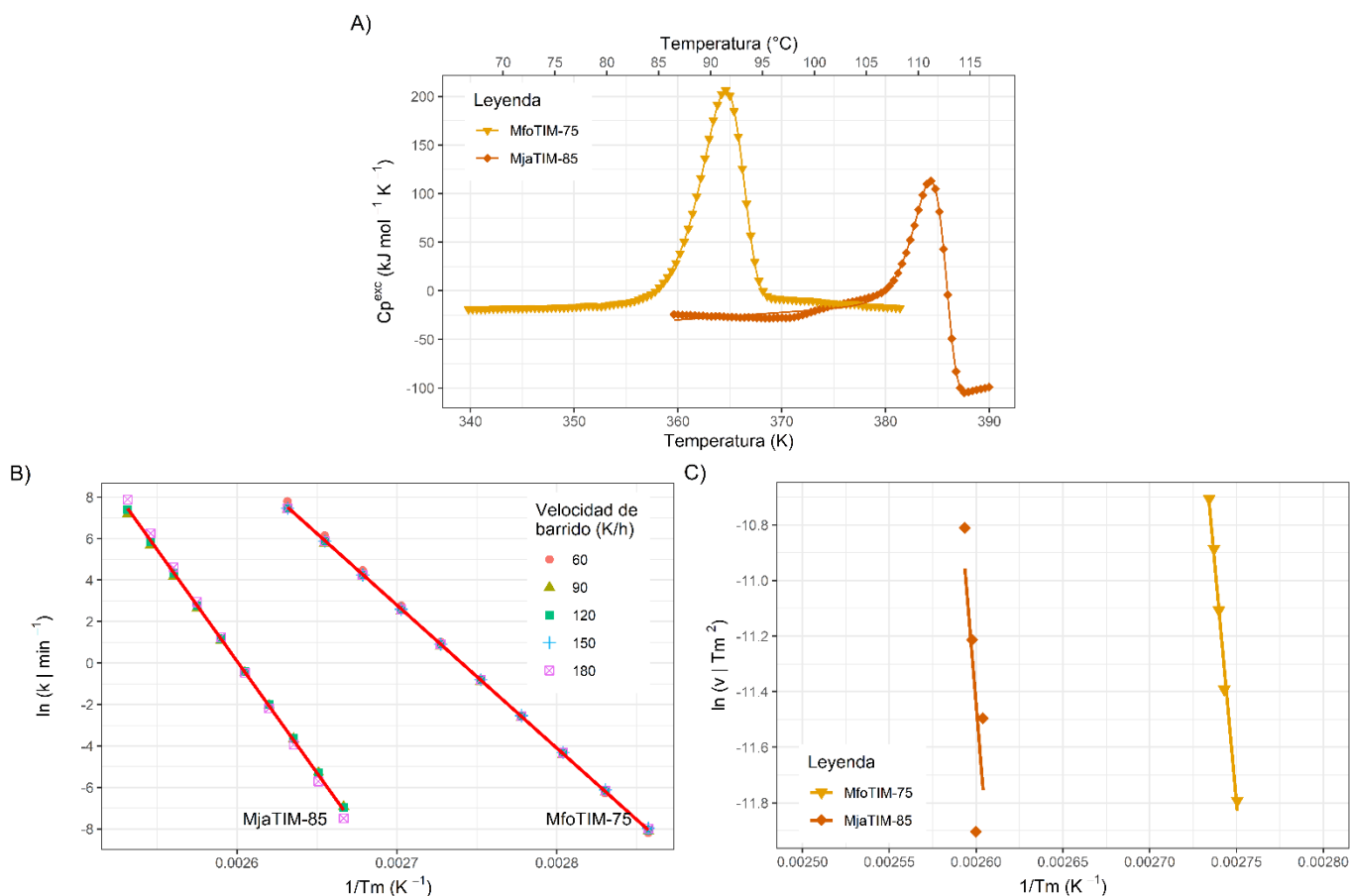
proceso irreversible y se sugiere que la transición tiene un componente cinético importante (Ibarra-Molero & Sanchez-Ruiz, 2006).<sup>8</sup>



**Figura 17. Despliegamiento térmico de MfoTIM-75 a distintas velocidades de barrido.** En el panel izquierdo se muestra cada endoterma. La concentración de proteína por experimento fue  $0.4 \text{ mg mL}^{-1}$ . En el panel derecho se muestra la dependencia de la  $T_m$  con la velocidad de barrido. Los puntos se ajustaron a una curva asintótica sólo con fines ilustrativos.

Esto último se corroboró en experimentos a diferentes velocidades de barrido (Figura 17A). En estos datos se observa que la  $T_m$  es dependiente de esta variable, con una diferencia de  $2.15 \text{ }^{\circ}\text{C}$  entre la  $T_m$  más elevada y la más baja. Con esta evidencia, podemos hacer el mismo análisis de dos estados irreversible, cuyos resultados se encuentran resumidos en la Tabla 9 y la Figura 18. Aunque en este caso cabe hacer la distinción que, a diferencia del caso anterior, la  $C_p$  del estado desplegado está por encima del estado nativo, por lo que podríamos asumir que la irreversibilidad está dada principalmente por alteraciones irreversibles de la proteína, probablemente por modificaciones químicas favorecidas a altas temperaturas (v. gr. la desaminación de asparaginas y glutaminas), o por agregación de la muestra.

<sup>8</sup> La transición se considera operacionalmente reversible si el tamaño del pico del segundo barrido alcanza entre el 85 y 95% del tamaño del primero. Ver referencia en texto.



**Figura 18. Métodos para calcular la energía de activación de desplegamiento.** A) Ajustes de las trazas calorimétricas al modelo irreversible de dos estados. B) Gráfico de Arrhenius para cada velocidad de barrido empleada. C) Dependencia de la  $T_m$  según la velocidad de barrido. Las líneas rojas corresponden a los ajustes respectivos.

**Tabla 9. Parámetros del desplegamiento térmico bajo control cinético.**

TIM	$T_m^{ap}$ <sup>1</sup> (°C)	$\Delta H_b$ <sup>1</sup> ( $\text{kJ mol}^{-1}$ )	Energía de activación ( $\text{kJ mol}^{-1}$ ) <sup>2</sup>	$t_{1/2}$ a 25 °C (años)	$t_{1/2}$ a 75 °C (horas)	$t_{1/2}$ a 85 °C (horas)
<b>MfoTIM-75</b>	$91.7 \pm 0.8$	$1142.5 \pm 38.7$	$575.1 \pm 5.3$	$3.27 \times 10^{12}$	103.6	0.4
<b>MjaTIM-85</b>	$111.6 \pm 0.6$	$717.5 \pm 66.5$	$829.6 \pm 131.7$	$2.24 \times 10^{29}$	$5.8 \times 10^{10}$	$1.02 \times 10^7$
<b>MjaTIM-85 con histidinas</b> <sup>3</sup>	$111.8 \pm 0.1$	$739.46 \pm 3.4$	$971.6 \pm 19.5$	ND	ND	ND

<sup>1</sup> Se hizo el promedio de los valores obtenidos para cada velocidad de barrido. Se presentan los errores de los estimados.

<sup>2</sup> El cálculo de la energía de activación del desplegamiento representa el promedio de las cuatro estrategias utilizadas para medir este parámetro, según como se establece en la sección de Métodos. Los valores específicos para cada método se encuentran en la Tabla S4.

<sup>3</sup> Datos de Muñiz Trejo, 2017

En la Tabla 9 se resumen los datos obtenidos del ajuste de las trazas de las TermoTIM al modelo irreversible de dos estados (Sanchez-Ruiz, 1992):



A partir de este análisis se puede calcular la energía de activación ( $E_A$ ) del proceso de desplegamiento, la entalpía de desnaturalización ( $\Delta H_D$ ) y la  $T_m$  aparente ( $T_m^{ap}$ ). La conversión entre N y F está asociada a una constante de velocidad de primer orden ( $k$ ), por lo que se puede construir un gráfico de Arrhenius para el desplegamiento de cada proteína (Figura 18B).

Los valores de  $k$  observados corresponden sólo al intervalo de temperatura que corresponde a la transición, pero se pueden extrapolar a cualquier otra temperatura por una regresión lineal. Se puede calcular, además, la vida media ( $t_{1/2}$ ) de la proteína para cada valor de  $k$ , por lo que en la Tabla 9 se indican algunos valores de  $t_{1/2}$  que se consideraron relevantes. Los datos observados en la transición se ajustan a una línea recta, lo que sugiere que la aplicación del modelo irreversible de dos estados representa adecuadamente la naturaleza de los datos. Sin embargo, fuera del intervalo de transición observada no podemos asegurar la  $k$  mantenga una dependencia lineal con la temperatura. Por esta razón, el único valor confiable de  $t_{1/2}$  es el de MfoTIM-75 a 85 °C, pues está dentro de su transición. El resto de los valores se presentan sólo como referencia.

Como se puede observar en la Tabla 9, la presencia de la etiqueta de histidinas no afecta en modo alguno la  $T_m^{ap}$  de la MjaTIM-85 y tampoco refleja cambios importantes en el  $\Delta H_D$ . Donde sí encontramos diferencias notables es en la comparación entre la MfoTIM-75 y MjaTIM-85. Las diferencias en el  $\Delta H_D$  se pueden relacionar con el fenómeno de agregación que se observa en las trazas calorimétricas de la MjaTIM-85 que se mencionó anteriormente o a que el empaquetamiento de la estructura nativa es diferente. Dada la formación de agregados durante la desnaturalización, se puede esperar una menor entalpía de desplegamiento por la cantidad de enlaces que se mantienen durante este proceso, lo cual especulamos es debido a la presencia de estructura residual. En MfoTIM-75 no se observa agregación (aunque sí hay alteraciones irreversibles), por lo que podemos esperar una entalpía mayor, que refleja una mayor cantidad de enlaces rotos y una menor estructura residual remanente durante el desplegamiento térmico.

Cabe resaltar que, aunque ambas proteínas son tetraméricas, la diferencia entre sus valores de  $T_m$  es de ~20 °C. Resulta evidente entonces que, en este caso, la diferencia en termoestabilidad no se debe al estado oligomérico, sino a algún otro factor estructural. Al respecto, hace falta obtener y cotejar

información cristalográfica para determinar cuáles son las diferencias que pueden ayudar a explicar dicha diferencia. Esta diferencia también vale ser contextualizada a nivel del organismo. Recordemos que los intervalos de crecimiento de los arqueas respectivas sólo se traslapan en el intervalo de los 80 a 83°C (ver Figura 4). En este sentido, la  $T_m$  de la MfoTIM-75 (91.7°C) parecería lo suficientemente elevada para permanecer plegada en el intervalo de temperatura de *Methanoterris formicicus* y de *Methanocaldococcus jannaschii*. No obstante, la transición comienza poco antes de los 85 °C, por lo que su estabilidad cinética a esa temperatura podría no ser suficiente para mantener la proteína funcionando por un tiempo fisiológicamente relevante ( $t_{1/2}$  estimada de MfoTIM-75 a 85 °C es de 24.5 minutos) en el intervalo de temperatura al que vive *Methanocaldococcus jannaschii*. Debido a ello, y al menos para estas dos proteínas, sí parece haber una relación entre la  $T_m$  y la temperatura óptima de crecimiento que podemos asociar con un proceso adaptativo.<sup>9</sup> Esta conclusión se ajusta adecuadamente para los datos obtenidos *in vitro*, pero hay que tener en mente que en condiciones celulares las interacciones proteína-proteína derivadas del hacinamiento molecular intracelular suelen incrementar la  $T_m$  *in vivo* (Despa, Orgill, & Lee, 2005; Mateus et al., 2018).

Para comprender los datos de  $E_A$  (Tabla 9) es necesario hablar primero de cuáles son las raíces estructurales de este parámetro. La  $E_A$ , que en este caso constituye la barrera energética que se tiene que sortear para llegar del estado nativo al estado desplegado, se ha asociado con la presencia de barreras de solvatación-desolvatación. Las barreras de solvatación se definen como aquellas asociadas al reacomodo de una cantidad finita de agua asociada al rearreglo conformacional de una proteína. Durante el proceso de desplegamiento (o de plegamiento, según sea el caso) existe una asincronía entre la ruptura de las interacciones internas de la proteína con la penetración del agua en los sitios que acaban de liberarse. El estado de transición, en este sentido, se piensa como un estado caracterizado por una red de contactos internos parcialmente rotos, pero no satisfechos todavía por moléculas de agua circundantes. De acuerdo con diversos estudios computacionales y experimentales, la magnitud de esta red insatisfecha por agua es la principal contribución a la  $E_A$  (Rodríguez-Larrea, Minning, Borchert, & Sanchez-Ruiz, 2006; Sanchez-Ruiz, 2010).

Para nuestros datos, esto nos indica que la presencia de la etiqueta de histidinas no está modificando las barreras de solvatación del estado de transición de manera radical con respecto a la MjaTIM-85 nativa, ya que la diferencia entre ambos valores de activación es pequeña. No obstante, la situación cambia radicalmente cuando comparamos la barrera energética entre la MfoTIM-75 y la MjaTIM-85. Aquí la diferencia de casi 250 kJ mol<sup>-1</sup> nos sugiere que hay una modificación importante entre las redes

---

<sup>9</sup> Para un trabajo que explora la asociación adaptativa de la  $T_m$  con la temperatura óptima de crecimiento desde una perspectiva comparativa, véase Hart *et al.*, 2014. Esta referencia será discutida con mayor detalle en la sección de reconocimiento molecular (*vid. infra*).

de interacciones parcialmente rotas en el estado de transición. Es importante contextualizar que entre ambas proteínas existe un 80% de identidad, por lo que los cambios en magnitud se deben a una cantidad relativamente pequeña de cambios entre las secuencias.

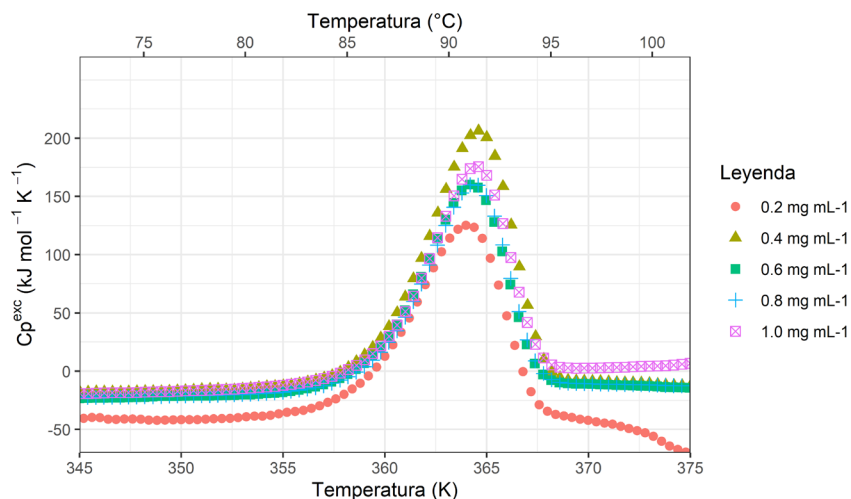
Este caso recuerda parcialmente a lo encontrado por Costas y colaboradores (2009) entre las TIM eucariontes de *Trypanosoma cruzi* (TcTIM) y de *Trypanosoma brucei* (TbTIM), cuyas energías de activación tienen un valor de 809 y 397 kJ mol<sup>-1</sup>, respectivamente, y comparten 74% de identidad. Estos autores encontraron que la exposición al solvente, provocada por urea, durante la desnaturalización no es suficiente para explicar la variación entre los valores de E<sub>A</sub> de estas TIM y concluyen que las diferencias se deben a las contribuciones de las barreras de solvatación en el proceso. Posteriormente, este grupo de trabajo definió cuantitativamente las diferencias a partir de un nueva medida que acuñaron: la flexibilidad térmica (Quezada *et al.*, 2017). Esta medida nace de la observación de que no toda la proteína se vuelve flexible al mismo tiempo cuando se incrementa la temperatura, sino que ciertas regiones se hacen más móviles que otras, por lo que el espacio conformacional que se puede visitar incrementa de manera preferencial. Así, la flexibilidad térmica describe la relación entre la energía de activación con los cambios cooperativos en la movilidad de estas regiones específicas dentro del gradiente de temperatura. Son estos cambios asincrónicos los que modifican la red de interacciones rotas en el estado de transición y que conllevan a una mayor energía de activación y, por tanto, a una mayor estabilidad cinética. Con este antecedente, podemos especular que la flexibilidad térmica de MjaTIM-85 es superior a la de MfoTIM-75. Estudios de desnaturalización térmica con urea nos pueden ofrecer más información al respecto de la accesibilidad del solvente en el estado de transición y nos ayudarían a concretar el análisis de las diferencias estructurales entre ambas proteínas.

#### *Determinación de la vía de plegamiento por desplegado térmico*

Según el principio de Le Châtelier “si se presenta una *perturbación* externa sobre un sistema en equilibrio, el sistema se ajustará de tal manera que se cancele parcialmente dicha perturbación en la medida que el sistema alcanza una nueva posición de equilibrio”(Chang, 2010, p. 638). Con perturbación, este principio puede referirse a cambios en la temperatura, la presión, la concentración, etc. En el análisis de proteínas oligoméricas, si hay cambios en el estado oligomérico durante el proceso de desplegado, se espera que el equilibrio químico sea diferente si el experimento se repite con diferentes concentraciones de proteína. Esto se ve reflejado en el desplazamiento de la curva de desnaturalización hacia una concentración menor del desnaturalizante, al disminuir la concentración de proteína. Si una curva de desnaturalización presenta múltiples transiciones, sólo se desplazarán aquellas que impliquen un cambio en el estado oligomérico (Gardner, Monroe, Kihara, &

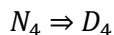


Park, 2016). Debido a que la agregación puede estar enmascarando los efectos de disociación de subunidades, no llevamos a cabo estos experimentos para MjaTIM-85, pero sí para MfoTIM-75.



**Figura 19. Despliegamiento térmico de MfoTIM-75 a distintas concentraciones de proteína.** Todos los experimentos se realizaron con una velocidad de barrido de  $90 \text{ K h}^{-1}$ .

En la Figura 19 se observan los resultados evaluados en el intervalo de concentración de 0.2 a 1  $\text{mg mL}^{-1}$ . Se observa que no hay una dependencia clara de la  $T_m$  con la concentración de proteína. La diferencia entre las  $T_m$  más alta y baja es de apenas  $0.53 \text{ }^\circ\text{C}$ , por lo que no podemos concluir que exista un fenómeno de disociación vinculado al despliegamiento. No obstante, hay que matizar que la desnaturalización no es reversible, por lo que el significado de esta falta de dependencia con la concentración necesita ser estudiada con mayor detalle. Esta falta de dependencia indicaría que las interfases entre subunidades de la proteína se mantienen unidas indistintamente del resto de la estructura y, por lo tanto, una posibilidad es que el mecanismo de despliegamiento siga el siguiente modelo:



donde  $N_4$  es el tetrámero en estado nativo y  $D_4$  es el tetrámero desplegado. No obstante, cabe la posibilidad de que en el intervalo de concentraciones evaluado no se vea disociación aparente porque nos encontramos en la porción asintótica de la curva de dependencia de la  $T_m$  con la concentración. Regularmente, experimentos a menor concentración deberían de ayudar a determinar si ese es el caso, pero una menor cantidad de proteína comprometería la señal de DSC. Una alternativa viable es agregar concentraciones bajas de desnaturalizante y observar si disminuye el valor de la  $T_m$  en espectros de DC a concentraciones de proteína aún menores, esperando que dicha disminución sea se alcance a observar por ese método.

### *Análisis de los datos bajo control termodinámico*

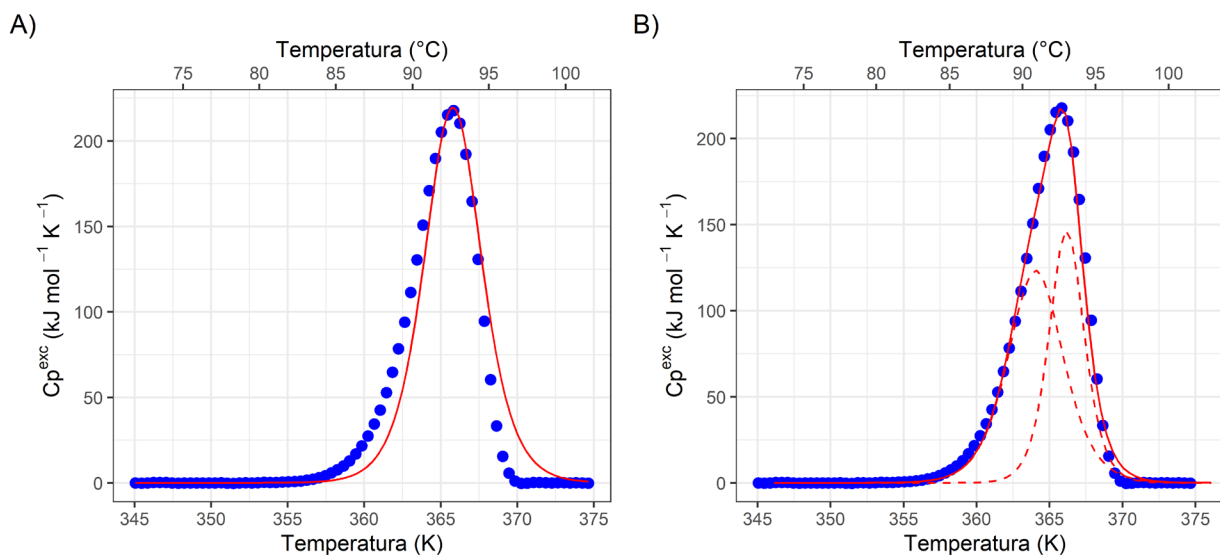
Regularmente, cuando una proteína sigue el modelo irreversible de dos estados se elimina la posibilidad de obtener un análisis de termodinámica al equilibrio. En algunos casos, sin embargo, es posible hacer dicha interpretación si los cambios de  $T_m$  con respecto a la velocidad de barrido son despreciables. Esto se puede observar más claramente si los datos de dependencia presentan un comportamiento asintótico hacia velocidades de barrido más elevadas. De ser así, a la velocidad de barrido más elevada se considera que las distorsiones cinéticas son demasiado pequeñas para ser tomadas en consideración, por lo que, con ciertas licencias, se puede hacer un análisis de termodinámica al equilibrio (Ibarra-Molero & Sanchez-Ruiz, 2006; Thórólfsson et al., 2002; Vogl, Jatzke, Hinz, Benz, & Huber, 1997).

La dependencia de la  $T_m$  con la velocidad de barrido en MfoTIM-75 parece seguir un comportamiento cuasi-asintótico (Figura 17B). Las diferencias entre las  $T_m$  son pequeñas y parecen disminuir mientras mayor es la velocidad de barrido, por lo que esta proteína es candidata de ser evaluada como si su desplegamiento estuviera bajo control termodinámico (ver Figura 20).

Una primera consideración que se tiene que hacer para este tipo de análisis es que el cálculo del  $\Delta C_p$  se ve alterado por las modificaciones irreversibles presentes durante la desnaturalización, así que no se puede obtener un valor completamente confiable de este parámetro. La segunda consideración que se debe de hacer es que este análisis sólo se puede realizar en los experimentos de DSC hechos a la mayor velocidad de barrido posible. Por estas razones, se decidió someter al ajuste a la traza calorimétrica obtenida a  $180 \text{ K h}^{-1}$  bajo una ecuación de desplegamiento con transiciones independientes de más de dos estados (Tabla 10). Esta ecuación considera que el  $\Delta C_p$  es constante, por lo cual se requiere restar la línea base. La razón por la que se escogió este modelo es que no hay simetría en la traza calorimétrica, lo que implica que hay más de un proceso implicado en el desplegamiento.

**Tabla 10. Parámetros del desplegamiento térmico por un modelo de control termodinámico.** Los errores corresponden al promedio de los diferentes ajustes al modelo.

Parámetro	Una sola transición	Dos transiciones	
		Primer pico	Segundo pico
$T_m$ (°C)	92.6	$90.9 \pm 0.3$	$93 \pm 0.1$
$\Delta H_{cal}$ (kJ mol <sup>-1</sup> )	$1099 \pm 34$	$626 \pm 61$	$497 \pm 61$
$\Delta H_{vH}$ (kJ mol <sup>-1</sup> )	$648 \pm 25$	$631 \pm 19$	$1000 \pm 59$
Criterio calorimétrico ( $\Delta H_{vH} / \Delta H_{cal}$ )	0.59	1.01	2.01



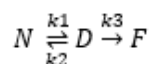
**Figura 20. Ajuste de MfoTIM-75 a modelo de desplegamiento de más de dos estados con transiciones independientes.** Se muestra la transición de la traza calorimétrica obtenida a  $180 \text{ K h}^{-1}$  de MfoTIM-75 a  $0.4 \text{ mg mL}^{-1}$  en amortiguador de fosfatos  $10 \text{ mM}$ ,  $\text{pH } 8.0$ . Las líneas rojas continuas representan el ajuste global, mientras que las líneas rojas discontinuas indican las transiciones independientes. A) Ajuste considerando una sola transición. B) Ajuste considerando dos transiciones independientes.

El análisis del área bajo la curva de una traza calorimétrica nos indica el cambio en entalpía calorimétrica ( $\Delta H_{\text{cal}}$ ). Debido a que este valor constituye al total de la transición, se considera que es independiente de la vía de plegamiento presente. Por otro lado, a partir del modelo se puede hacer una estimación del valor del cambio en entalpía a partir de la ecuación de van't Hoff ( $\Delta H_{\text{vH}}$ ). Este valor constituye una estimación del proceso si éste fuera de dos estados. A la comparación de estos dos valores ( $\Delta H_{\text{vH}}/\Delta H_{\text{cal}}$ ) se le conoce como el criterio calorimétrico, el cual nos dice si una transición puede ser evaluada como un modelo de dos estados. Si el cociente no da un valor de uno, esta medida nos indica que tenemos diferentes unidades cooperativas de plegamiento. El valor obtenido fue de 0.59 (Tabla 10), lo que nos indica que hay más de una unidad cooperativa de plegamiento, por lo que se analizaron los datos con dos transiciones independientes (Figura 20B). Estos nuevos valores indican que la primera transición cumple con lo esperado para un modelo de dos estados, mientras que la segunda ( $\Delta H_{\text{vH}}/\Delta H_{\text{cal}} > 1$ ) implica la presencia de un proceso irreversible (Sanchez-Ruiz, 1995). Aunque el criterio calorimétrico indica otra cosa con los resultados actuales, es posible especular que la falta de corrimiento apreciable en la  $T_m$  con la concentración se deba a que hay dos picos bajo la misma endoterma. Si el primer pico estuviera relacionado con la disociación y el segundo con el desplegamiento irreversible, el corrimiento del primero al incrementar la concentración de proteína se vería “enmascarado” por el segundo pico. Lo anterior se podría corroborar experimentalmente con ensayos a diferentes concentraciones, pero con una velocidad de barrido elevada. En cualquier caso,

nuestros datos sugieren que el desplegamiento de MfoTIM-75 sigue el siguiente esquema de desnaturalización:



Este mecanismo es igual al descrito por Lumry & Eyring (1954) en la que el estado nativo (N) es estable termodinámicamente respecto a las formas desplegadas (D), pero estas últimas pasan por alteraciones irreversibles hacia un estado final (F):



La constante de equilibrio está dada por  $Keq = k_1/k_2$  y posee una constante cinética de primer orden  $k$ . Curiosamente, el modelo irreversible de dos estados que se utilizó para evaluar los datos bajo control cinético es una variación del modelo de Lumry-Eyring donde  $k_3 \gg k_2$ , por lo que la gran mayoría de las moléculas desplegadas pasan hacia el estado final (Sanchez-Ruiz, 1992). Este comportamiento podría explicar el porqué se encontró reversibilidad parcial del proceso, además de que es un claro indicio de que el desplegamiento de esta proteína está controlado tanto por factores cinéticos como por factores termodinámicos.

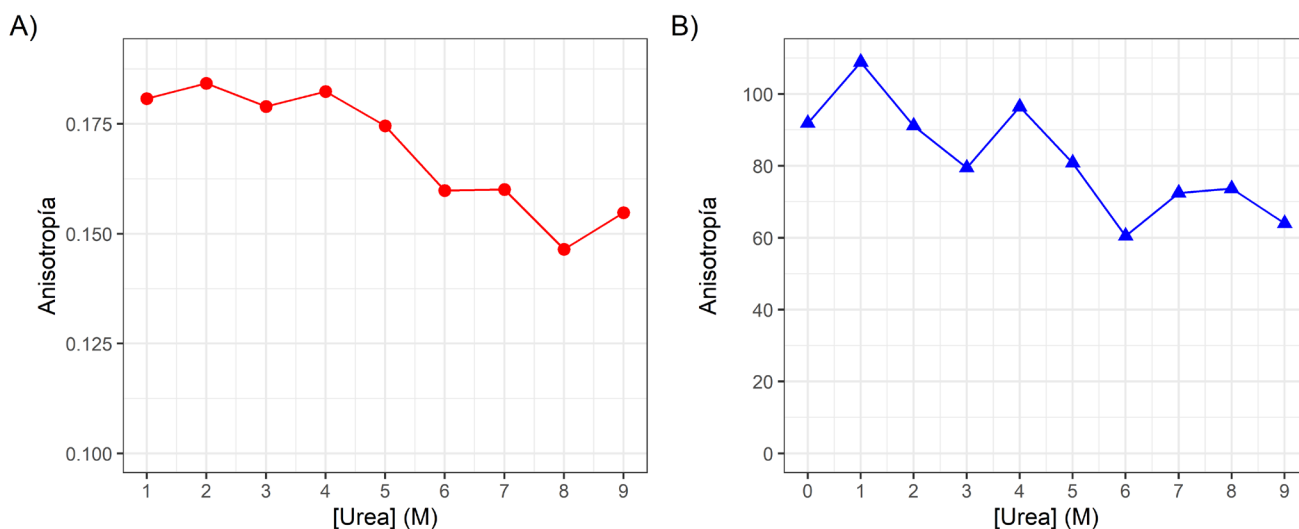
### **Denaturalización por agentes químicos**

Antes de proseguir con la presentación de los resultados y la discusión de esta subsección, cabe resaltar algunas características de los estudios realizados.

Para evaluar el desplegamiento provocado por agentes químicos usualmente se recomienda medir el efecto del desnaturalizante por el mayor número de métodos posibles, ya que ello nos puede dar una mayor cantidad de información sobre los cambios asociados a diferentes elementos estructurales de la proteína. Durante el trabajo de licenciatura (Muñiz Trejo, 2017), sólo se obtuvieron curvas de desplegamiento por dicroísmo circular a 222 nm y actividad catalítica de la MjaTIM-85 con histidinas. Es decir, se evaluó el estado de la estructura secundaria y del sitio activo. Sin embargo, no se implementó un método para medir los cambios terciarios, ya que proteínas sin triptófanos no presentan transiciones cuando se sigue la señal por el máximo de fluorescencia ni por el centro espectral de masa durante la desnaturalización química. Como alternativa, encontramos que Forsyth y Matthews (2002) pudieron caracterizar el desplegamiento terciario de la enzima indol-3-glicerol fosfato sintasa del arqueón termófilo *Sulfolobus solfataricus*, el cual carece de triptófanos en su secuencia y posee una topología de barril  $(\beta/\alpha)_8$ , por cambios en la anisotropía de las tirosinas. Debido a las semejanzas

con nuestros modelos de estudio, decidimos emplear este método para incluir la medición de la anisotropía en el desplegamiento químico de nuestras enzimas, incluida la versión con histidinas de MjaTIM-85.

En el trabajo anterior se determinó también que la urea en un intervalo de 0 a 8.5 M a 25 °C no tiene efectos importantes en la estructura/función de la MjaTIM-85 con histidinas después de una semana de incubación. Por esta razón se hizo un estudio preliminar del efecto de este desnaturizante en MfoTIM-75. Similar al caso anterior, la urea afecta sólo parcialmente al plegamiento de MfoTIM-75, ya que no se alcanza a observar una transición completa (ver Figura 21). En el caso de la actividad, se observa una pérdida del ~40% y una ligera disminución en la anisotropía hacia los 9 M del desnaturizante después de 72 horas de equilibrio a 25 °C. Debido a esto, se descartó utilizar a la urea para seguir la caracterización del desplegamiento químico de las ArqTIM. El clorhidrato de guanidinio (GdnHCl), en contraparte, sí tuvo en efecto sobre su plegamiento y todas las muestras alcanzaban el equilibrio aproximadamente a las 12 horas de incubación a 25 °C, por lo que toda la caracterización se hizo con este desnaturizante.



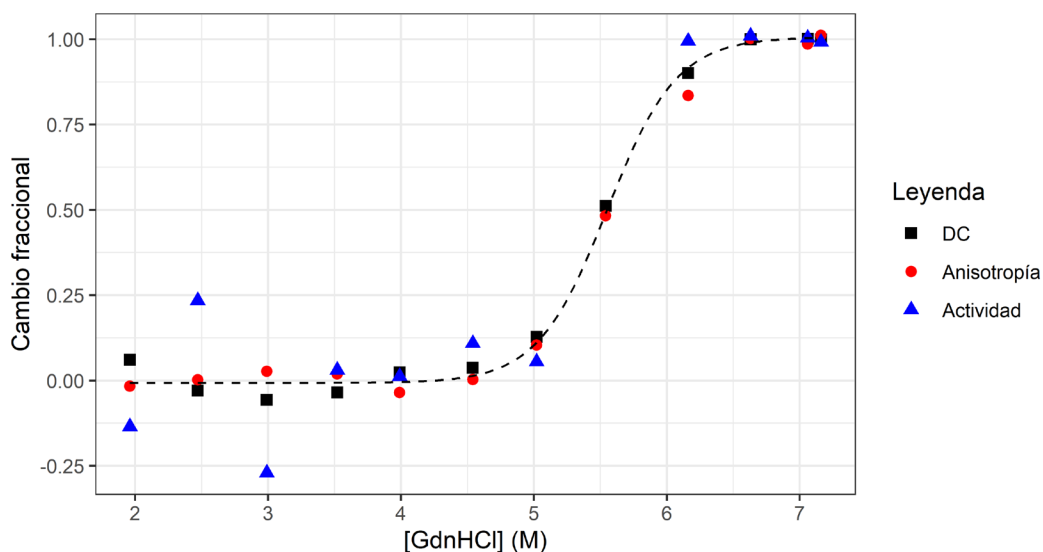
**Figura 21. Efecto de la urea sobre el plegamiento de MfoTIM-75.** Se siguió la señal de anisotropía (panel izquierdo) y actividad (panel derecho) de 0.1 mg mL<sup>-1</sup> de la proteína después de incubar por 72 horas a 25°C a diferentes concentraciones de urea en un amortiguador de fosfato de sodio 10 mM, pH 8.0.

Tanto la urea como el GdnHCl son agentes caotrópicos (*i. e.* rompen la estructura del agua) que incrementan la solubilidad de las moléculas no polares por medio del efecto de *salting-in* (Hribar, Southall, Vlachy, & Dill, 2002; Zhang, 2012); sin embargo, se sabe que el GdnHCl tiene una capacidad desnaturizante superior a la de la urea debido a que se disocia en las especies iónicas Gdn<sup>+</sup> y Cl<sup>-</sup> en solución, por lo que sus propiedades de hidratación son diferentes e interrumpe las interacciones

electrostáticas de las proteínas (Mason *et al.*, 2017). La fuerza iónica que provee el GdnHCl a bajas concentraciones puede estabilizar a las proteínas, por lo que conviene incubar a las proteínas con alguna otra alguna otra sal monovalente para medir el efecto salino (Arakawa & Timasheff, 1984). En nuestros casos, probamos incubar a la proteína con cloruro de sodio y encontramos un efecto estabilizante en el intervalo de 0 a 2 M en las curvas medidas por anisotropía (Figura S4), por lo que excluimos los primeros puntos de todos los análisis para esta técnica.

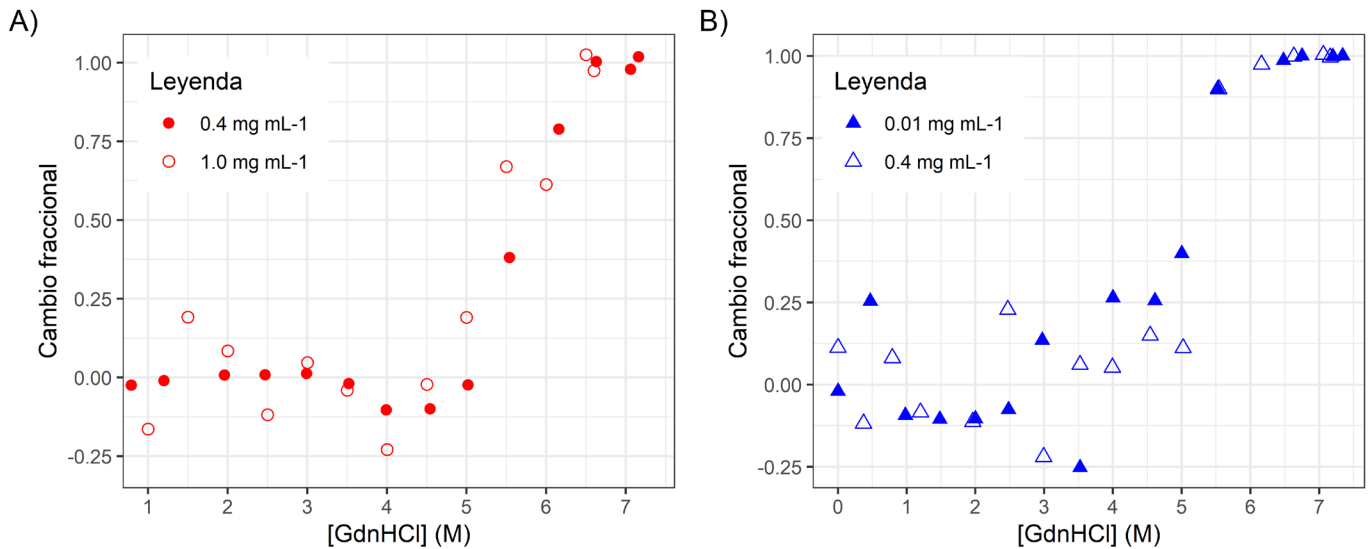
#### *MjaTIM-85 con etiqueta de histidinas*

Como se mencionó, los datos de desnaturalización de la versión con histidinas de MjaTIM-85 se complementaron con los datos de anisotropía y todos se normalizaron para la comparación entre métodos (Figura 22).



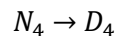
**Figura 22. Curva de desplegamiento con GdnHCl de MjaTIM-85 con histidinas a 25 °C.** Gráfico del cambio fraccional de la proteína a 0.4 mg mL<sup>-1</sup> con concentraciones ascendentes de GdnHCl según distintos métodos: anisotropía a 303 nm, dicroísmo circular a 222 nm y actividad catalítica. La línea discontinua indica el ajuste a una ecuación de Boltzmann con una Cm de 5.55 M. El tiempo de incubación fue de 24 horas.

Con estos nuevos datos confirmamos que el desplegamiento químico de esta proteína es monofásico con un valor de la concentración media de desnaturalización (Cm) aparente de 5.5 M, lo cual es congruente con lo reportado anteriormente (Muñiz Trejo, 2017). Lo que también fue consistente es la falta de reversibilidad en su desplegamiento. Con el propósito de abonar al mecanismo de desplegamiento, se obtuvieron curvas con concentraciones distintas de proteína por anisotropía y actividad (Figura 23).



**Figura 23. Curvas de desnaturalización de MjaTIM-85 con histidinas a distintas concentraciones de proteína.** Panel izquierdo: Anisotropía. Panel derecho: Actividad catalítica específica. Las concentraciones de proteína utilizadas se especifican en la leyenda.

Aún con el ruido de los datos, en MjaTIM-85 con histidinas no parece haber desplazamiento alguno de la curva de desplegamiento, por lo que se podría inferir que el desplegamiento se da sin disociación del tetrámero. Nuevamente, esto tiene el matiz de que el proceso no está al equilibrio, por lo que el desplazamiento de la curva puede no presentarse a diferentes concentraciones de la proteína. A diferencia de lo expuesto para el desplegamiento térmico de MfoTIM-75, aquí se exploró un intervalo de concentración de 0.01 a 1 mg mL<sup>-1</sup> de proteína, por lo que parece ser que efectivamente la proteína se despliega sin disociación:

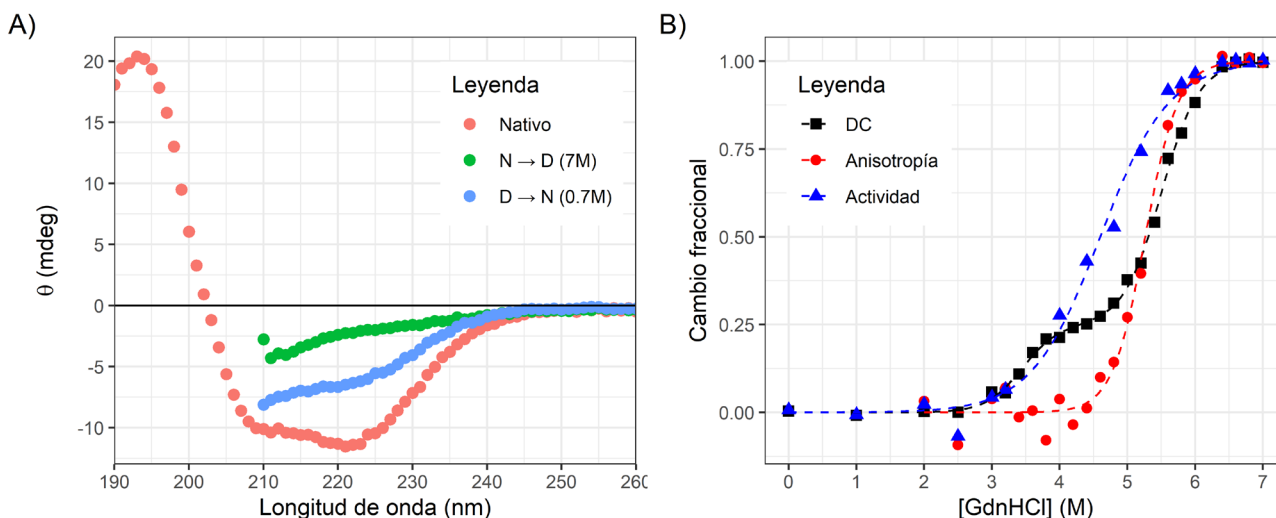


En la literatura hay ejemplos de proteínas multiméricas que no alcanzan a disociarse completamente hasta monómeros con agentes caotrópicos (v. gr. la  $\beta$ -lactoglobulina bovina A y la concanavalina A, véase Sawyer & Puckridge, 1973), pero no encontramos ejemplos donde no se observe ningún grado de disociación. Debido a lo inusual de este resultado, necesitamos repetir el experimento procurando minimizar el ruido a 0.01, 0.1 y 1 mg mL<sup>-1</sup> por la mayor cantidad de métodos posible y realizar repeticiones para considerar el error estadístico en los datos. De igual forma se propone evaluar el desplegamiento por cromatografía de exclusión molecular con el fin de determinar si hay cambios en el tamaño de partícula que nos permitan reconocer si hay disociación previa al desplegamiento completo de las muestras. Para estudiar si había cambios en el estado oligomérico se midió la curva por dispersión dinámica de luz; sin embargo, esta técnica no tiene la sensibilidad suficiente para

detectar la diferencia de tamaño entre dímeros y tetrameros, por lo que no se pudo reconocer si el tamaño de partícula cambiaba antes de la transición, aunque sí se observó la misma transición monitoreada por actividad y anisotropía (datos no mostrados).

### MjaTIM-85

A continuación, se exploró el desplegamiento con GdnHCl de la versión sin histidinas de MjaTIM-85.



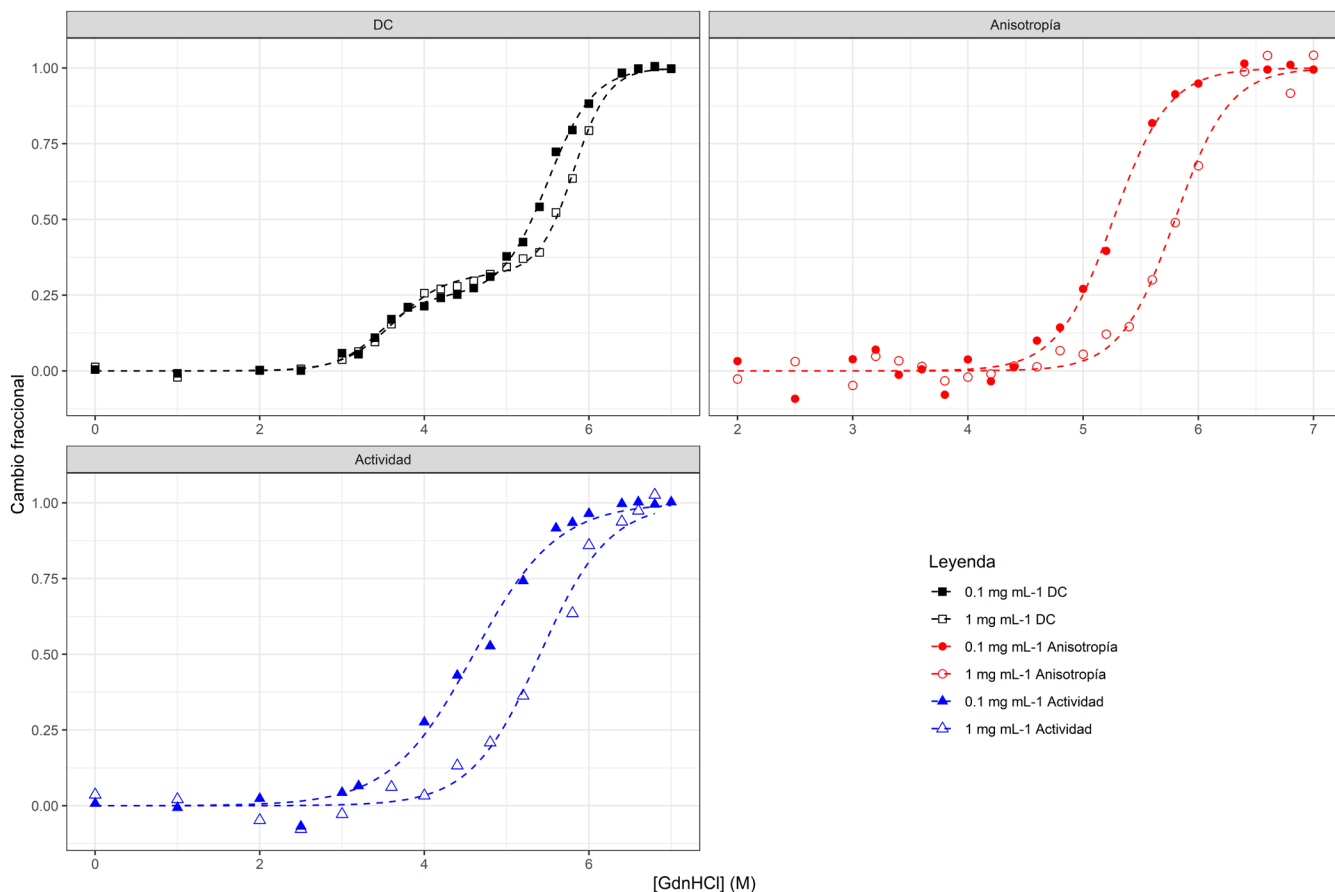
**Figura 24. Desplegamiento con GdnHCl de MjaTIM-85 a 25 °C.** A). Prueba de reversibilidad del desplegamiento. B) Curva de desnaturalización obtenida con  $0.1 \text{ mg mL}^{-1}$  según distintos métodos. Las líneas discontinuas reflejan el ajuste a la ecuación de Boltzmann o doble Boltzmann, según el caso. Tiempo de incubación: 24 horas.

Parecido a lo observado en la otra versión, la desnaturalización no fue reversible (Figura 24A), sin embargo (y para nuestra sorpresa), la curva de desnaturalización de MjaTIM-85 sin etiqueta de histidinas es completamente diferente a la de su versión con etiqueta (Figura 24B). En primera instancia, es muy ostensible como diferentes métodos presentan transiciones a diferentes concentraciones y en segunda, se observa una curva bifásica en la desnaturalización seguida por dicroísmo circular. Esto nos indica claramente la presencia de intermediarios al equilibrio. En este caso era posible suponer que algunas de las curvas presentarían dependencia a la concentración, lo cual nos ayudaría a delinear el mecanismo de plegamiento.

**Tabla 11. Valores de  $C_m$  de la desnaturalización de MjaTIM-85 sin histidinas por método empleado.**

Técnica		$0.1 \text{ mg mL}^{-1}$	$1 \text{ mg mL}^{-1}$
Dicroísmo	$C_{m1}$	3.44	3.63
	$C_{m2}$	5.5	5.81
Anisotropía	$C_m$	5.26	5.8
Actividad	$C_m$	4.6	5.4





**Figura 25. Curvas de desnaturalización de MjaTIM-85 sin histidinas a dos concentraciones de proteína.** Los resultados se muestran por método de análisis para facilitar la visualización. En símbolos llenos corresponden a  $0.1 \text{ mg mL}^{-1}$  y los vacíos a  $1 \text{ mg mL}^{-1}$  de proteína. Se muestran los datos por dicroísmo circular a  $222\text{nm}$ , ajuste a doble Boltzmann (negro); por anisotropía de tirosinas a  $303 \text{ nm}$ , ajuste a Boltzmann (rojo); por actividad catalítica específica, ajuste a Boltzmann (azul).

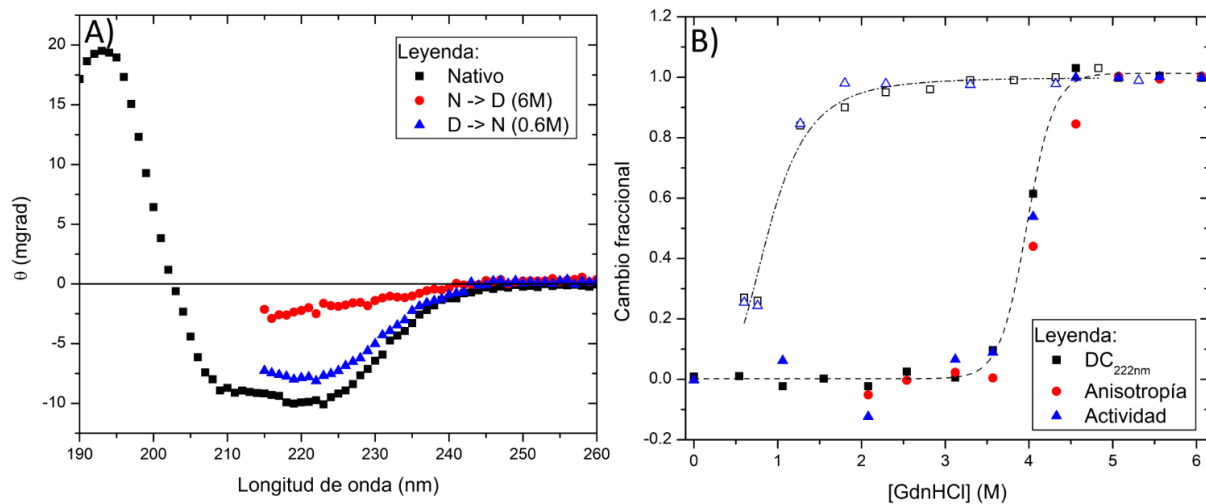
Como se observa en la Figura 25 y en la Tabla 11, las cuatro transiciones presentan dependencia con la concentración de proteína. Suponiendo que una proteína tetramérica se disociara de monómero en monómero, es decir, que siguiera la siguiente secuencia: tetramero  $\rightarrow$  trímero  $\rightarrow$  dímero  $\rightarrow$  monómero, el número máximo teórico de transiciones dependientes de concentración debería ser tres, pero los datos parecen indicar que son cuatro. Esta última observación puede deberse a errores en la medición. Como se aprecia en la Tabla 11, a  $1 \text{ mg mL}^{-1}$ , la  $C_{m2}$  por dicroísmo y la  $C_m$  por anisotropía presentan el mismo valor, mientras que a  $0.1 \text{ mg mL}^{-1}$  tienen un valor muy similar. Adicionalmente, en la Figura 24B también podemos observar cómo ambas curvas cruzan muy cerca la una de la otra. Hasta tener mediciones por triplicado que nos permitan calcular la desviación estándar podremos discernir cuáles transiciones sí son diferentes y así afinar el mecanismo de plegamiento propuesto.

Independientemente de precisar los datos de las  $C_m$ , resulta evidente que la etiqueta de histidinas está modificando de manera categórica el comportamiento de la MjaTIM-85 cuando se expone al

despliegamiento por GdnHCl, y aunque no hay reversibilidad que nos permita comparar con valores termodinámicos, sí podemos conjeturar que está modulando la estabilidad de la proteína bajo estas condiciones. El efecto que tiene la etiqueta de histidinas sobre la estabilidad ha sido explorado antes en la literatura. En 2018, Booth y colaboradores evaluaron el valor del despliegamiento térmico de diez proteínas por fluorimetría diferencial de barrido y encontraron que, en la mayoría de las condiciones de pH y fuerza iónica probadas, la presencia de la etiqueta de histidinas disminuía el valor de la  $T_m$  respecto al presentado sin etiqueta. Por otro lado, Aslantas & Surmeli (2019) encontraron que la termoestabilidad y la actividad del citocromo P450 119 incrementan cuando se coloca una etiqueta de histidinas en el N-terminal. En nuestro caso, el acompañamiento de esta etiqueta no parece modificar de manera importante la termoestabilidad, pero sí tiene un efecto sobre la actividad y, sobre todo, en el patrón de plegamiento, ya que no se observaron intermediarios al equilibrio cuando la etiqueta está presente. Resulta evidente que, para un estudio de carácter evolutivo, es necesario eliminar componentes extra que nos pueden estar sesgando la interpretación que se hace de los datos.

#### MfoTIM-75

Finalmente, se probó el efecto del cloruro de guanidinio sobre el plegamiento de MfoTIM-75 y su reversibilidad ante este agente.



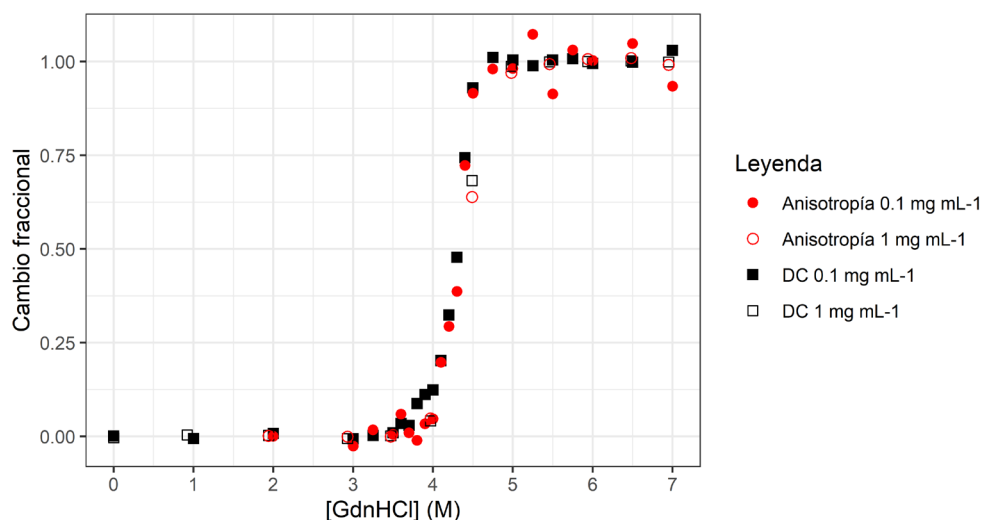
**Figura 26. Despliegamiento de MfoTIM-75 con GdnHCl a 25 °C.** A). Prueba de reversibilidad del despliegamiento. B) Curva de desnaturalización obtenida con 0.1 mg mL<sup>-1</sup> según distintos métodos. En símbolos llenos la curva de desnaturalización, en símbolos vacíos la curva de renaturalización. Las líneas discontinuas sólo son indicativas de la histéresis. Tiempo de incubación: 24 horas.

En cuanto a la reversibilidad (Figura 26A), pudimos observar que la proteína recuperaba señal de dicroísmo durante la renaturalización, lo cual fue un primer indicio de que sí podemos recuperar a la proteína. Sin embargo, cuando evaluamos el regreso desde 6 M en una curva a distintas concentraciones (Figura 26B), notamos a simple vista que hay una histéresis muy marcada entre la

vía de desplegamiento y la de replegamiento. Es particularmente importante destacar que no alcanzamos a recuperar completamente la señal ni de dicroísmo, ni actividad, ya que en la concentración de GdnHCl más baja de replegamiento (0.6 M) aún no termina la transición. La histéresis es un efecto cinético en el que un sistema no sólo reacciona según sus condiciones actuales, sino que depende de su historia particular, es decir, el comportamiento del sistema depende de la vía. Esto, a nivel de proteína, nos indica que su paisaje energético es complejo e intricado, con trampas cinéticas y termodinámicas en el camino (Andrews, Capraro, Sulkowska, Onuchic, & Jennings, 2013). La histéresis es común en proteínas multidominio o multiméricas como la titina (Minajeva, Kulke, Fernandez, & Linke, 2001), el colágeno (Davis & Bächinger, 1993) o algunas cápsides virales (Singh & Zlotnick, 2003), pero también se ha observado para algunas triosafosfato isomerasa como la de levadura, la de conejo (Nájera, Costas, & Fernández-Velasco, 2003) y la de la bacteria hipertermófila *Thermotoga marítima* (Beaucamp, Schurig, & Jaenicke, 1997).<sup>10</sup> Estas trampas cinéticas pueden implicar una gran cantidad de tiempo para alcanzar un verdadero equilibrio, por lo que las proteínas con estas características se suelen estudiar en condiciones de pseudo-equilibrio (Andrews et al., 2013). En nuestros experimentos encontramos que no había cambios detectables en las señales de desplegamiento/replegamiento después de 72 horas de equilibrio, así que falta claridad sobre cuál de las transiciones observadas es la que corresponde a los datos al equilibrio o si convergen en algún punto intermedio. Sin embargo, para el análisis al pseudo-equilibrio decidimos tomar como referencia la curva de desplegamiento, pues es la que más asemeja a lo observado para MjaTIM-85 en cuanto a los valores de la  $C_m$  (Tabla 12).

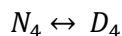
---

<sup>10</sup> Un aspecto que vale la pena resaltar sobre la TIM de *T. marítima* es que es la única otra TIM tetramérica identificada fuera del dominio Archaea hasta la fecha; sin embargo, esta TIM está fusionada covalentemente con otra enzima glicolítica, la fosfoglicerato cinasa, y su estado de asociación se da por la interacción de esta última, por lo que no se considera el mismo tipo de tetrámero (Beaucamp, Hofmann, Kellerer, & Jaenicke, 1997).



**Figura 27. Curvas de desplegamiento con GdnHCl a distintas concentraciones de MfoTIM-75.** Se muestran los resultados con distintos métodos del desplegamiento. En símbolos llenos a 0.1 mg mL<sup>-1</sup> y en símbolos vacíos a 1 mg mL<sup>-1</sup>. Tiempo de incubación: 24 horas.

Primero se evaluó el efecto de la concentración en la curva para determinar si existen cambios en el estado oligomérico de MfoTIM-75 durante el desplegamiento (Figura 27). Parecido a lo encontrado con los datos térmicos, tampoco pudimos identificar un corrimiento de la  $C_m$  entre las concentraciones evaluadas (0.1 y 1 mg mL<sup>-1</sup>), lo que nos lleva a concluir que el desplegamiento es de dos estados, sin disociación y reversible, pero con histéresis:



Dado que los datos no están por completo al equilibrio, cabe la posibilidad de que la ausencia de disociación sea parte de un efecto cinético. No obstante, es notable que el mecanismo de desplegamiento descrito sea tan similar al observado para MjaTIM-85 con histidinas, donde tampoco hay dependencia aparente de la  $C_m$  con la concentración de proteína. Esto recalca la importancia de caracterizar a las proteínas en condiciones similares, pues mantener la etiqueta de histidinas pudo conducir a una mala interpretación de la conservación de esta vía durante el curso evolutivo de estas proteínas homólogas.

Ya con lo anterior delineado, se procedió a determinar la estabilidad conformacional aparente en condiciones de pseudo-equilibrio para esta proteína a 25°C. Para esto se tuvo que determinar la dependencia del  $\Delta G$  con la concentración de desnaturalizante ( $m$ ) y a partir de eso estimar el valor de estabilidad en ausencia de guanidinio ( $\Delta G_{H_2O}$ ) y la  $C_m$  para una proteína de dos estados (Tabla 12 y Figura 28).

Tabla 12. Parámetros termodinámicos del desplegamiento químico de MfoTIM-75 a 25°C.

Método	$\Delta G_{H_2O}$ (kJ mol <sup>-1</sup> )	$m$ (kJ mol <sup>-1</sup> M <sup>-1</sup> )	$C_m$ (M)
Ajuste de los datos de dicroísmo	96.15 ± 1.47	22.35 ± 0.34	4.29 ± 0.001
Ajuste de los datos de anisotropía	102.87 ± 2.32	23.79 ± 0.53	4.27 ± 0.003
Ajuste de los datos normalizados	91.74 ± 1.29	21.35 ± 0.3	4.29
Ajuste lineal	101.31 ± 8.7	23.68 ± 2.04	4.28
Promedio	<b>98 ± 5.1</b>	<b>22.8 ± 1.2</b>	<b>4.28 ± 0.01</b>

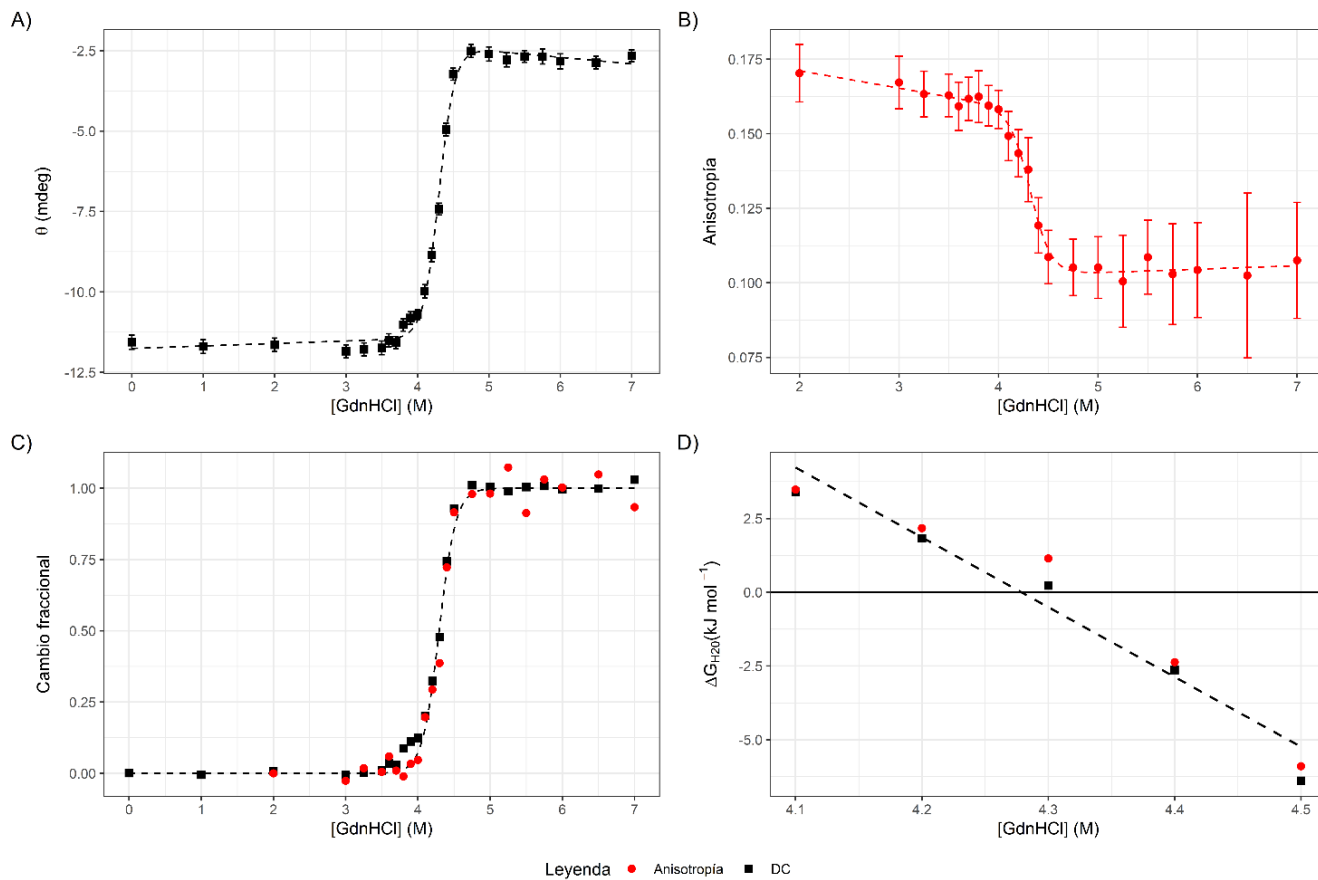
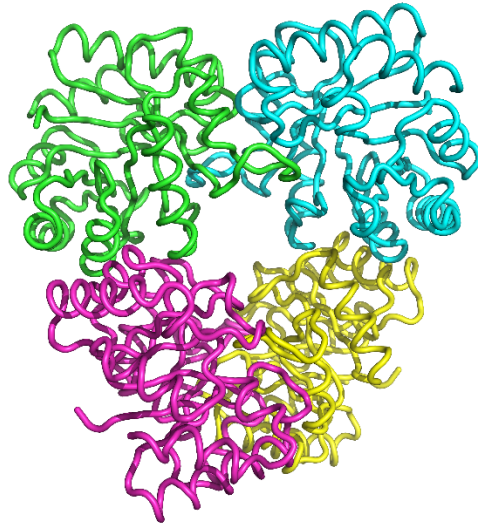


Figura 28. Ajustes de la curva de desnaturalización de MfoTIM-75 a un modelo de dos estados. Ajuste de los datos crudos de A) dicroísmo circular en elipticidad a 222 nm y B) anisotropía a 303 nm de emisión con la ecuación de Santoro & Bolen (1988, SB). Las barras representan la desviación estándar de las mediciones de los equipos. C) Ajuste de los datos normalizados de dicroísmo y anisotropía por SB. D) Extrapolación lineal de los datos de la transición ajustados al modelo de Greene & Pace (1974).



**Figura 29. Modelo estructural de MfoTIM-75.** Generado por Robetta por modelaje comparativo con la MjaTIM-85 (PDB: 2H6R) con valor de confianza de 0.87.

Entre las distintas curvas encontramos congruencia para los tres parámetros evaluados. Adicionalmente, se utilizó el valor de la  $m$  para estimar el cambio en el área accesible al solvente ( $\Delta ASA$ ) durante la desnaturalización de la MfoTIM-75 por medio de la correlación descrita por Myers, Pace, & Scholtz (1995) en la que la  $m = 859 + 0.22(\Delta ASA)$ . Para contextualizar este dato, se determinó también el  $\Delta ASA$  teórico de la desnaturalización completa de esta proteína por medio de la relación, también descrita por Myers y colaboradores (*ibid.*), entre el  $\Delta ASA$  y el número de residuos de una proteína ( $\Delta ASA = -907 + 93(\#residuos)$ ). De igual manera, se generó un modelo de la proteína por homología con el servidor Robetta (Kim, Chivian, & Baker, 2004, <http://rosetta.bakerlab.org/>) a partir de la estructura cristalográfica de la MjaTIM-85 (Figura 29), del cual se calculó su  $\Delta ASA$  de desnaturalización por medio del servidor VADAR (Willard *et al.*, 2003, <http://vadar.wishartlab.com/>). Los resultados se resumen en la Tabla 13.

**Tabla 13. Cálculos del  $\Delta ASA$  (Å) de desnaturalización de MfoTIM-75.** Los porcentajes indican a cuánto equivale el valor obtenido con respecto a los valores calculados.

Método		Calculado a partir del número de aa	Calculado a partir del modelo estructural
		81305	121678
Obtenido a partir del valor de la $m$	20864	26 %	17 %

Lo que estos datos nos indican de manera cualitativa es que la cantidad de área expuesta por la desnaturalización con GdnHCl es tan sólo una pequeña fracción del área total que debería exhibirse si la proteína se disociara al tiempo que se despliega. Esto abona a la evidencia de que esta proteína

se despliega sin perder su estado oligomérico, muy probablemente con presencia de estructura residual.

En cuanto al valor de  $\Delta G$  obtenido, como ya hemos discutido, no es comparable con otras TIM pues se obtuvo de datos al pseudo-equilibrio. Aunque implica algunas limitantes técnicas y operativas, quizá se podría aprovechar la  $T_m$  tan elevada de esta proteína para soslayar el problema de la histéresis al incubar las muestras a una temperatura mayor. Alternativamente, se podrían generar datos de estabilidad a la dilución que nos permitan observar el comportamiento de las subunidades a diferentes concentraciones y su contribución a la estabilidad de la MfoTIM-75.<sup>11</sup>

### ***Función, estabilidad y evolución de proteínas termófilas, parte II***

Una de las propiedades que más comúnmente se atribuyen a las proteínas globulares es que son marginalmente estables. Eso quiere decir que su  $\Delta G$  de plegamiento ronda los 20 a 41 kJ mol<sup>-1</sup> (ó 5 a 10 kcal mol<sup>-1</sup>), la cual se considera una energía apenas suficiente para mantenerlas estables en condiciones fisiológicas (Taverna & Goldstein, 2002; Williams, Pollock, & Goldstein, 2006). Esta noción de marginalidad, desde una perspectiva fisicoquímica, surge de la oposición entre las fuerzas que desestabilizan a una proteína, como la entropía conformacional, y las fuerzas que la estabilizan, como el efecto hidrofóbico y las interacciones moleculares no covalentes (Privalov & Khechinashvili, 1974). Desde el punto evolutivo, la estabilidad marginal se ha visto como una adaptación de las proteínas para mantener su funcionalidad, esto se explica porque se observa un balance, un *trade-off*, entre características estructurales que confieren actividad y las que confieren estabilidad, ya que ambas parecen ser opuestas (Beadle & Shoichet, 2002; DePristo, Weinreich, & Hartl, 2005; Závodszky, Kardos, Svingor, & Petsko, 1998).<sup>12</sup> Este razonamiento recuerda lo que ya se había descrito para la actividad de proteínas termófilas a temperatura ambiente en el capítulo anterior: como la rigidez incrementa mientras disminuye la temperatura, la actividad debe disminuir concomitantemente, por lo que la estabilidad de una proteína siempre es marginal en el intervalo de temperatura óptima de un organismo.

Estas ideas sobre la estabilidad y la función se contraponen a lo encontrado en las TermoTIM y en las TIM de *Pyrococcus furiosus* y *Thermococcus onnurineus*, mencionadas anteriormente (Sharma & Guptasarma, 2015). Este no es un fenómeno único para las TIM de arqueas. En *Thermotoga maritima*,

---

<sup>11</sup> Generamos algunos datos preliminares al respecto de ambas TermoTIM (Figura S5 y Tabla S5), pero tienen la limitante aquí es que no hay certeza si la disociación se presenta de tetrámeros a monómeros, tetrámeros a dímeros, o tetrámeros a dímeros a monómeros, pues desconocemos si los dímeros mantienen actividad catalítica.

<sup>12</sup> Nótese nuevamente la distinción entre las causas próximas que nos ofrece la perspectiva fisicoquímica y las causas últimas que se proponen desde una perspectiva evolutiva para el fenómeno de estabilidad marginal (ver la Introducción de este texto).

por ejemplo, se han identificado varias proteínas como la fosforibosil antranilato isomerasa (Sternier et al., 1996), la indolglicerol fosfato sintasa (Merz, Knöchel, Jansonius, & Kirschner, 1999) y la L-isoaspartil metiltransferasa (Ichikawa & Clarke, 1998), cuya actividad es elevada e, incluso, puede ser mayor que las de sus homólogas mesófilas en *E. coli* a temperatura ambiente.

Este tipo de hallazgos y una visión crítica al paradigma adaptacionista han generado evidencia que sugiere que la estabilidad marginal y el balance (*trade-off*) estructura/función no son características adaptativas.<sup>13</sup> En un experimento seminal del grupo de la premio Nobel de Química 2018, Frances Arnold, con el citocromo P450 se mostró que una estabilidad elevada confiere un potencial adaptativo mayor, ya que permite explorar nuevas funciones cuya adquisición implica una serie de mutaciones desestabilizantes. Los autores encontraron que la función y la estabilidad no son inherentemente incompatibles, sino que una estabilidad elevada puede incrementar la *evolucionabilidad* de una proteína (Bloom, Labthavikul, Otey, & Arnold, 2006).

La evolucionabilidad (*sensu* Wagner & Altenberg, 1996) es la habilidad de un sistema biológico de producir variación *fenotípica* que es *heredable* y *adaptativa*.<sup>14</sup> Para entender este concepto es necesario distinguir entre lo que es la variación y la variabilidad. La *variación* se refiere a las diferencias presentes entre individuos de una población, especie o clado, mientras que la *variabilidad* remite al potencial de una característica fenotípica de variar, en otras palabras, qué tan fácil es que un fenotipo determinado cambie. De manera general, la facilidad con la que un fenotipo puede modificarse depende de qué tan robusto es ante el cambio genético, en ese sentido, la *robustez* constituye una medida de la neutralidad del cambio. Así, un fenotipo robusto será capaz de explorar un mayor número de trasfondos genéticos sin experimentar heterogeneidad adaptativa (Payne & Wagner, 2019; Wagner, 2008).

De este marco interpretativo se derivan algunas observaciones importantes. La primera es que proteínas cuya relación estabilidad/actividad no está tan estrechamente vinculada tienen una mayor variabilidad heredable en su secuencia, porque los cambios no afectan de manera tan importante la función. En ese sentido, el origen de la variación observada en estas secuencias, de manera general, no es adaptativo pues no se compromete la actividad (y por lo tanto la adecuación) tan fácilmente. En

---

<sup>13</sup> En punto hay que recordar que la noción de *trade-off* es precisamente uno de los puntos de crítica de Gould & Lewontin (1979). Cuando se encuentra que un fenotipo determinado es óptimo porque de serlo se hace a expensas de la optimización de otro. Se invoca la idea de *trade-off* para explicar que estas características son subóptimas porque deben contribuir lo mejor posible al todo sin dejar de lado una noción adaptativa.

<sup>14</sup> *Evolucionabilidad* es un término polisémico y su interpretación varía de autor a autor. En este texto nos referiremos a la acepción acuñada por la escuela de Wagner & Altenberg (1996) y actualizada por Payne & Wagner (2019) (No son el mismo Wagner). Para más información sobre otros usos del término, el filósofo Massimo Pigliucci ofrece una revisión histórica de sus distintos usos en su artículo *Is evolvability evolvable?* (2008).



su lugar, como la actividad es robusta en un intervalo amplio de temperatura, los cambios observados en la secuencia no están constreñidos a mantener la funcionalidad ante un ambiente que se modifica. Para el caso de las TermoTIM, si bien sólo son reflejo de la variación existe, el que la actividad elevada a baja temperatura sea recurrente para otras TIM de arqueas indica que esta característica se debe considerar en cuenta cuando se estudie la evolución de estas enzimas.

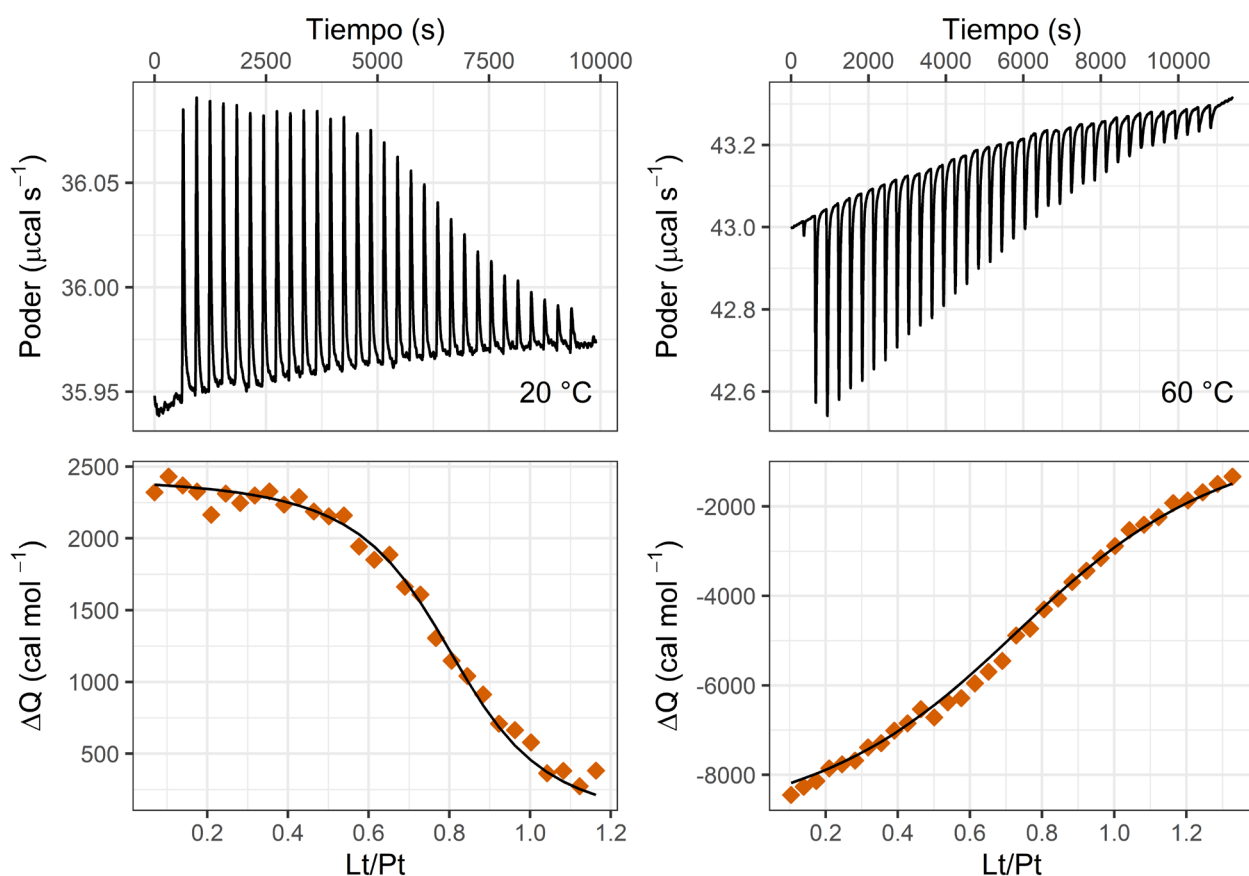
La segunda tiene que ver con la relación entre la secuencia de aminoácidos y las otras características biofísicas de estas proteínas. Si la secuencia puede variar sin que se comprometa la adecuación del organismo, los cambios observados pueden interpretarse de la neutralidad del cambio, ya no están dados por procesos adaptativos vinculados a la pérdida o adquisición de una función. Además, como se mencionó en los antecedentes, la triosafosfato isomerasa es también muy robusta a la mutación a nivel de secuencia, por lo que el origen de la variación en otras características biofísicas como la energía de activación, la entalpía de desnaturalización, la estabilidad conformacional, la reversibilidad ante la desnaturalización, etc., probablemente tampoco sea adaptativo. Como regla general, parece ser que mientras los cambios no bajen de cierto umbral, tanto en estabilidad (Bershtein *et al.*, 2006) como en actividad (Newton *et al.*, 2015), no hay una penalización sobre el valor de adecuación. En ese sentido, buena parte de la variación observada entre las TermoTIM puede ser simplemente un mero subproducto de la neutralidad del cambio de la secuencia ante la mutación y de la actividad ante la temperatura.

La termoestabilidad, como se mencionó, parece ser una excepción a esto porque por lo general sí está correlacionada con la temperatura de crecimiento del organismo en muchas familias de proteínas. MfoTIM-75 tiene una  $T_m$  veinte grados menor que MjaTIM-85, por lo que esa relación parece cumplirse también (aunque se admite que dos puntos no son suficientes para establecer una correlación). Reconocer cuáles de los cambios de la secuencia están vinculados con la modulación de la termoestabilidad, ya ambas son tetrámeras, será fundamental para establecer el origen de los cambios en las secuencias de las TermoTIM.

Pensando en la robustez de la actividad con la temperatura, quisimos profundizar más en el análisis del reconocimiento molecular de las TermoTIM. Además, hasta ahora, no se han hecho estudios de reconocimiento en TIM tetraméricas, por lo que iniciar estudios de esta naturaleza sería una buena pieza de información para seguir discutiendo acerca de cuáles características cambian a partir de explorar la neutralidad del espacio de secuencias. Los resultados de estos ensayos se discuten en el siguiente capítulo.

### Caracterización de la unión a ligando

Debido a orden emitida por nuestra máxima casa de estudios al respecto de suspender las labores de investigación en el contexto de la contingencia sanitaria internacional por COVID-19 (Dirección General de Comunicación Social, 2020), en esta sección del trabajo sólo se obtuvieron resultados preliminares de un solo experimento de la unión de las TermoTIM con un análogo fosforilado del estado de transición, el ácido fosfoglicolohidroxámico (PGH) por calorimetría de titulación isotérmica (ITC). Pese a que presentan algunos valores numéricos, el análisis se realizó bajo una perspectiva principalmente cualitativa.



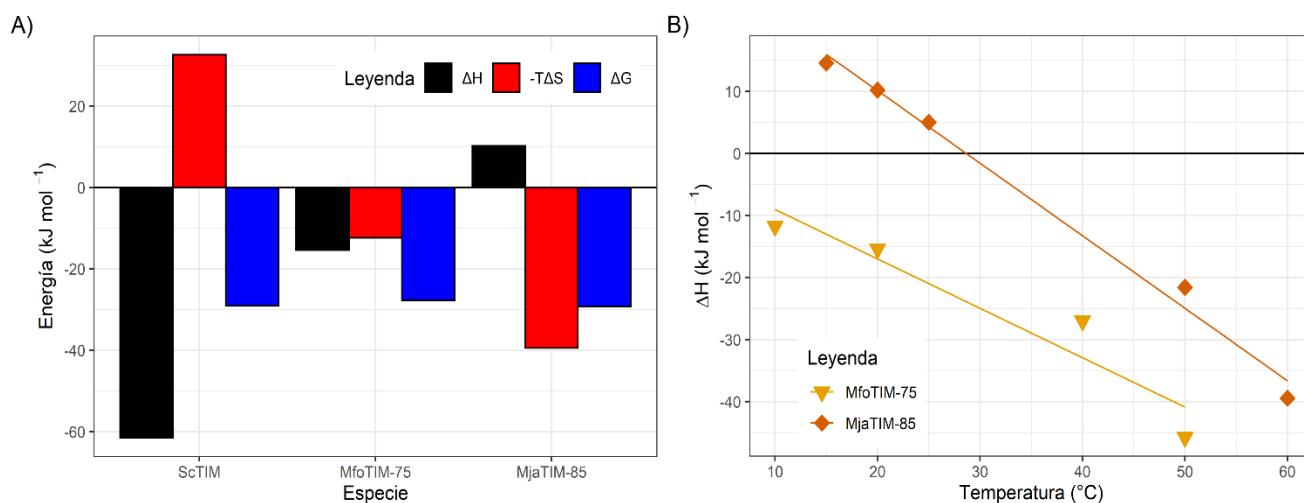
**Figura 30. ITC de unión a PGH de las MjaTIM-85 a dos temperaturas.** En los paneles superiores se observa el termograma de unión en la temperatura especificada. En los paneles inferiores, las isothermas a partir de la integración de los picos, se muestran los datos observados con puntos y la línea de ajuste. Amortiguador de trietanolamina (TEA) 100 mM, pH 7.4

El resultado más llamativo de estos experimentos fueron la isothermas de MjaTIM-85 (Figura 30), ya que, a bajas temperaturas, la señal de entalpía es endotérmica ( $\Delta H > 0$ ), mientras que a temperaturas elevadas la señal es exotérmica ( $\Delta H < 0$ ). Las isothermas para cada temperatura de las TermoTIM y los valores respectivos de cada ajuste se encuentran en las figuras Figura S6 y Figura S7 y en las Tabla S6 y Tabla S7.

**Tabla 14. Firmas termodinámicas de unión a PGH de distintas TIM a 25°C.** En azul se indican las contribuciones favorables a la unión y en naranja las contribuciones desfavorables.

TIM <sup>a</sup>	$\Delta G_U$ (kJ mol <sup>-1</sup> )	$\Delta H_U$ (kJ mol <sup>-1</sup> )	$-T\Delta S_U$ (kJ mol <sup>-1</sup> )	Referencia
MjaTIM-85	-29.2	5.0	-34.2	Este trabajo
MfoTIM-75	-27.8	-15.4	-12.4	
NpTIM	-26.3	-17.1	-9.2	Romero Romero, 2017
GoTIM	-24.3	-45.6	21.3	
CpTIM	-28	-39.3	11.7	
DrTIM	-27.2	-60.2	33	
ScoTIM	-27.2	-54.8	27.6	
ScTIM	-29.1	-61.5	32.6	Serratos <i>et al.</i> , 2011

El resultado a 25 °C fue novedoso, porque las firmas termodinámicas de unión identificadas hasta la fecha para la TIM han sido siempre guiadas por entalpía, es decir, el proceso de unión siempre se había identificado como exotérmico. Por otro lado, salvo el caso de NpTIM, la entropía de unión siempre había sido desfavorable (Tabla 14). No obstante, a esta temperatura MjaTIM-85 presenta una firma completamente invertida, mientras que MfoTIM-75 tiene tanto una entalpía como una entropía favorables (Figura 31A).



**Figura 31. Firmas termodinámicas de unión a PGH de las TermoTIM.** En el panel A se muestra la firma típica de unión de PGH a una TIM ejemplificada por ScTIM a 25 °C (Serratos *et al.*, 2011). Junto se muestran las firmas de MjaTIM-85 a 25°C y MfoTIM-75 a 20 °C. En el panel B se muestra la dependencia del  $\Delta H_U$  con la temperatura. El  $\Delta C_{pU}$  se determinó como la pendiente del ajuste lineal de los valores presentados. MjaTIM-85  $\Delta C_{pU} = -1.17$  kJ mol<sup>-1</sup> K<sup>-1</sup>; MfoTIM-75  $\Delta C_{pU} = -0.796$  kJ mol<sup>-1</sup> K<sup>-1</sup>.

El análisis a diferentes temperaturas indica un  $\Delta C_{pU}$  negativo para ambas proteínas (Figura 31B), por lo que el valor del  $\Delta H$  va disminuyendo conforme se incrementa la temperatura de los experimentos. Esto nos indica que, al intervalo de temperatura del organismo, la firma termodinámica es guiada por entalpía, tal como sucede con las otras TIM mesófilas. No obstante, el resultado obtenido a 25 °C (y en general menor a 30 °C) nos parece llamativo, ya que como describimos antes, la actividad catalítica a esa temperatura sigue siendo elevada.

Los resultados de un experimento de ITC surgen de la contribución de distintas fuerzas moleculares jugando un papel simultáneo en la formación espontánea del complejo proteína-ligando. Por ejemplo, los grados de restricción de las cadenas laterales de residuos, del ligando, así como de las moléculas de agua e iones presentes dan origen a los efectos entrópicos, mientras que la ruptura de enlaces con el agua y la formación de enlaces entre los complejos moleculares (proteína/proteína o proteína/ligando, según sea el caso) ofrecen una serie de contribuciones entálpicas (Doyle, 1997). Debido a la suma de contribuciones de diferentes términos, la interpretación fenomenológica de los datos termodinámicos que la ITC nos proporciona puede ser complicada (Frasca, 2016), aunque ya se tienen algunas pautas delineadas para comenzar a dar razón de los datos observados.

Un  $\Delta C_{pU}$  negativo se utiliza como un indicativo general de la preponderancia de las interacciones hidrofóbicas en el proceso. Esto se debe a que la capa de agua que rodea a las especies libres (capa de solvatación) se vuelve más débil conforme incrementa la temperatura, lo que tiene como consecuencia una disminución del  $\Delta H$  mientras más elevada es la temperatura (Holdgate & Ward, 2005; Ward & Holdgate, 2001). En cuanto al valor de entropía, se sabe que la interacción entre una proteína y su ligando conduce a su desolvatación, lo cual resulta energéticamente favorable. En este sentido, la desolvatación implica que se rompen las interacciones del agua con las especies involucradas. La ruptura de interacciones resulta en una entalpía desfavorable, pero su liberación contribuye positivamente a la entropía (Claveria-Gimeno, Vega, Abian, & Velazquez-Campoy, 2017; Freire, 2008). En la unión de lectinas con carbohidratos, por ejemplo, se da una importante remoción de moléculas de agua en la superficie de interacción, lo que contribuye favorablemente al  $\Delta G$  por la naturaleza entrópica del proceso (Srinivas, Bhanuprakash Reddy, & Surolia, 1999). Otro buen ejemplo es la unión de la GroEL con proteínas desplegadas, la cual se da completamente por la desolvatación, por lo que la firma termodinámica es desfavorable en cuanto a la entalpía (Lin, Schwarz, & Eisenstein, 1995). Esta relación puede ser la clave de la firma invertida de MjaTIM-85, a diferencia de MfoTIM-75, es posible que la superficie de unión esté altamente solvatada por debajo de los 30 °C y la formación de enlaces con el ligando no es capaz de compensar la pérdida de los mismos. La obtención de

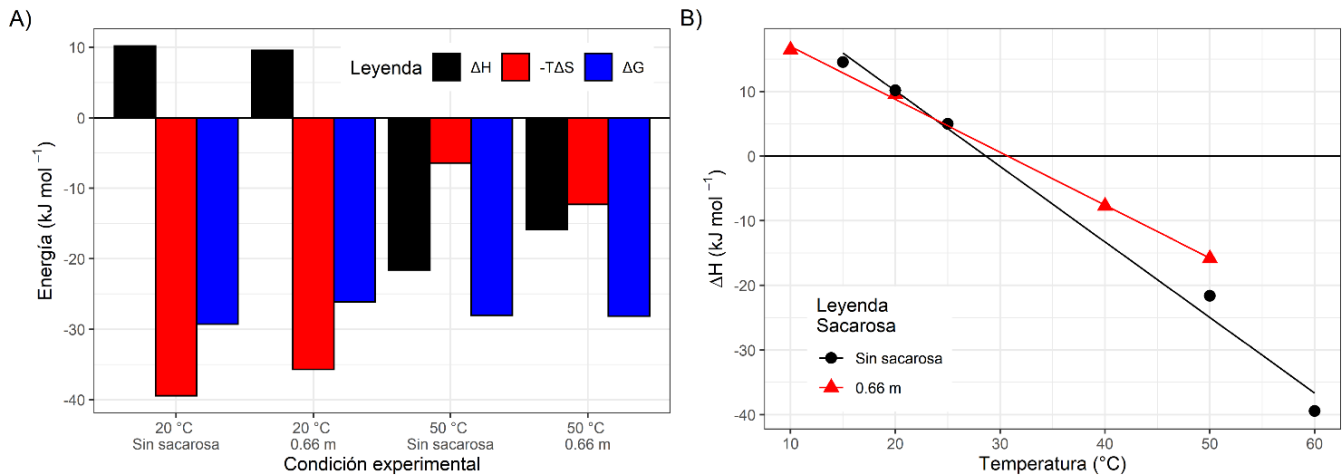
información estructural de ambas proteínas con y sin ligando puede ofrecernos más información al respecto de las moléculas de agua involucradas durante la interacción.

Uno de los factores que contribuye positivamente a la entalpía son las moléculas de agua que quedan atrapadas en la superficie de unión y que no pueden ser liberadas al resto del agua bulto (Claveria-Gimeno et al., 2017). Experimentalmente es posible determinar el número de moléculas de agua que quedan atrapadas durante la unión bajo condiciones de estrés osmótico. Esto se hace graficando el efecto que distintas concentraciones de osmolitos tienen sobre la  $K_A$  de la proteína, de la pendiente de esta relación se puede obtener el número de moléculas de agua atrapadas ( $\Delta N_w$ ) bajo la ecuación 22 (Parsegian, Rand, & Rau, 1995):

$$\frac{\delta \ln K_A}{\delta [\text{osmolito}]} = -55.6 \times \Delta N_w \quad (22)$$

Lo anterior se basa en que los osmolitos disminuyen la actividad de la mayor parte de las moléculas de agua presentes, pero no son capaces de acceder a las moléculas de agua que estén involucradas en los complejos proteína-ligando. En 2018, Serratos y colaboradores calcularon la cantidad de moléculas de agua atrapadas en la superficie de unión de la ScTIM con tres osmolitos diferentes: la sacarosa, la glicina betaína y el etilenglicol, con lo que confirmaron que el agua actúa como un puente entre la enzima y los ligandos por medio de una superficie hidratada. El número de moléculas de agua atrapadas depende el tipo y tamaño del osmolito, siendo la sacarosa la molécula que atrapa el mayor número de moléculas y el etilenglicol el menor.

Con este antecedente, decidimos explorar el efecto de la sacarosa sobre la unión del PGH en las TermoTIM, en aras de obtener datos sobre el número de aguas atrapadas en las superficies de interacción a diferentes temperaturas. No obstante, por las condiciones contingentes mencionadas al principio de esta sección, sólo pudimos obtener datos de la unión con 0.66 molal de sacarosa de MjaTIM-85 (Figura 32), por lo que no se determinaron valores numéricos concretos. Las isotermas completas se muestran en la Figura S8 y los parámetros correspondientes en la Tabla S8.



**Figura 32. Firma termodinámica de la MjaTIM-85 con sacarosa.** Se obtuvieron firmas de unión de PGH en presencia de 0.66 molal de sacarosa a distintas temperaturas. En el panel A se hace la comparación de las firmas obtenidas con y sin el osmolito a 20 y 50 °C. En el panel B se observa la dependencia del  $\Delta H$  con la temperatura.  $\Delta C_{pU}$  sin sacarosa =  $-1.17 \text{ kJ mol}^{-1} \text{ K}^{-1}$ ;  $\Delta C_{pU}$  con 0.66 molal de sacarosa =  $-0.82 \text{ kJ mol}^{-1} \text{ K}^{-1}$ .

En estos últimos datos podemos observar que la presencia de sacarosa en el medio disminuye el  $\Delta C_{pU}$ , por lo que los efectos de la temperatura sobre el  $\Delta H_U$  son menos ostensibles. A 20 °C podemos ver que la contribución entrópica es menor, pero a 50 °C esta señal es mayor que cuando no hay osmolitos, lo que representa un primer indicio de que sí hay un efecto de la capa de solvatación que modula la interacción con el ligando.

### ***Función, estabilidad y evolución de proteínas, parte III***

El que la firma termodinámica de unión a PGH de las TermoTIM en el intervalo de temperatura óptima de crecimiento del individuo sea igual que las observadas en organismos mesófilos nos podría llevar a considerar que, efectivamente, hay una ventaja adaptativa en este patrón para la función de la triosafosfato isomerasa (Tabla 14). Sin embargo, no hay que perder de vista que la actividad catalítica de estas TIM sigue siendo “perfecta” o casi perfecta a temperatura ambiente. En ese sentido, cabe preguntarse si la firma termodinámica, o más específicamente, las determinantes estructurales de la unión están jugando un papel en la adecuación del organismo. Lo anterior nos conduce una vez más a replantearnos cuáles son los aspectos del fenotipo que son sujetos de selección y cuáles pueden presentar variación sólo como consecuencia de la exploración neutra del espacio de secuencia.

Un antecedente de esta noción podemos encontrarlo en el trabajo de Hart y colaboradores (2014), donde los autores se propusieron entender las raíces históricas de la adquisición de la termoestabilidad de la ribonucleasa H1 entre bacterias mesófilas y termófilas. Para ello se valieron de la reconstrucción de secuencias ancestrales a modo de trazar los caminos que distintos parámetros siguieron en el curso

de la evolución de estos organismos. Ellos pudieron identificar que había una relación estadística entre el incremento de la  $T_m$  con el tipo de ambiente en el que habitan los organismos. Sin embargo, los parámetros fisicoquímicos que describen la curva de estabilidad de una proteína, y por tanto definen el valor de la  $T_m$ , como son el  $\Delta C_p$ , el  $\Delta H$  de desplegamiento y la temperatura de máxima estabilidad ( $T_s$ ), fluctúan indistintamente del ambiente en el que viven los organismos. Esto es un claro indicio de que la selección natural no es capaz de “sensar” los mecanismos subyacentes a la estabilización. Los autores comparan este fenómeno con lo que se observa en algunos procesos ontogenéticos o fenotipos determinados que se conservan por selección, pero cuyas bases genéticas o mecanismos moleculares pueden modificarse durante la evolución, denominado “*deriva de sistemas del desarrollo* [developmental system drift]” (True & Haag, 2001).<sup>15</sup>

Nuestros datos no son suficientes para determinar si estamos ante un escenario como el anterior, o sobre cuáles son los efectos de la robustez mutacional en la presencia de ciertos rasgos, por lo cual es importante encontrar condiciones de expresión y purificación de las MesoTIM, de modo que se pueda finalizar el estudio por el método comparativo filogenético y determinar cuáles características representan plesiomorfías y cuáles apomorfías del grupo. Además, para complementar la información filogenética es fundamental encontrar maneras de evaluar el efecto de la adecuación sobre la población sobre estas arqueas o sobre un modelo de estudio pertinente.

Además de estos experimentos, sería relevante tratar de obtener versiones diméricas de estas proteínas para tratar de entender la neutralidad del cambio y sus consecuencias fisicoquímicas en términos de termoestabilidad y actividad. Asimismo, la reconstrucción de secuencias ancestrales de las TIM del dominio Archaea nos ayudaría a precisar los caminos evolutivos de las modificaciones en el estado oligomérico que se han descrito en la literatura.

---

<sup>15</sup> En nuestra opinión, la razón por la que estos autores no encuentran correlación entre esas propiedades fisicoquímicas y la temperatura de crecimiento radica en el problema de atomización del fenotipo planteado por Gould & Lewontin (1979) (ver Introducción). Dado que hay una relación física indisoluble entre  $T_m$ ,  $\Delta C_p$ ,  $\Delta H$  y  $T_s$  descrita matemáticamente por la ecuación de Gibbs-Helmholtz, no podemos tratar a estos parámetros como “partes” independientes que pueden estar sujetas a selección natural.

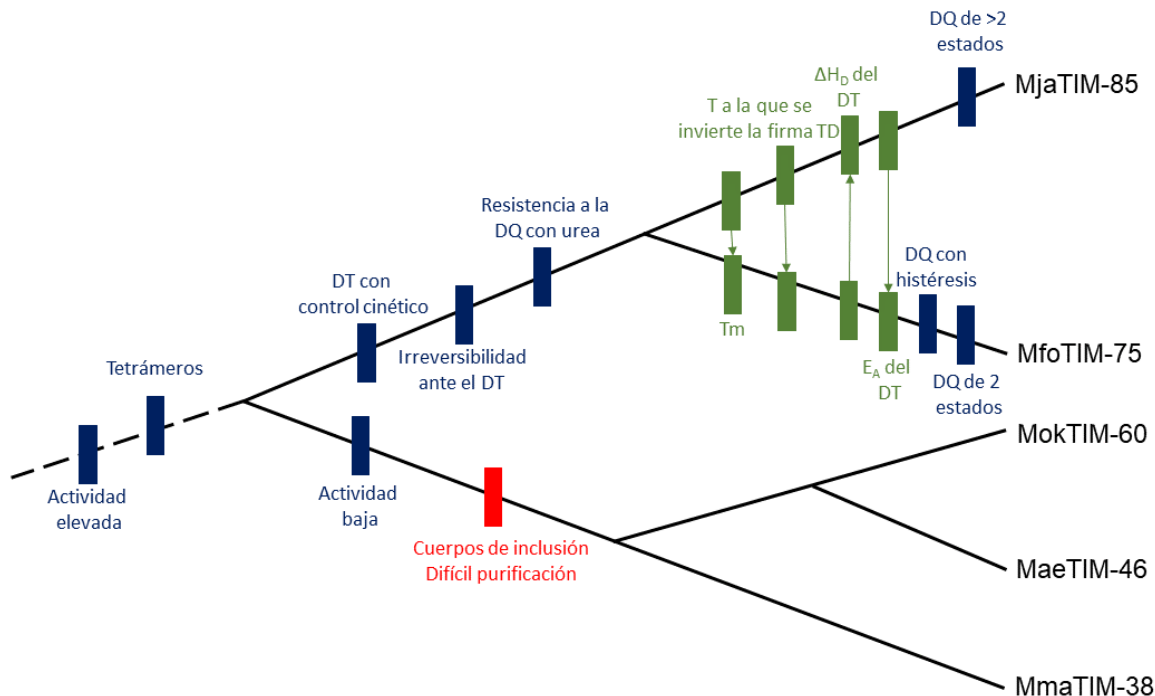
## RESUMEN DE RESULTADOS

Parámetro	MjaTIM-85	MfoTIM-75	MokTIM-60	MaeTIM-46	MmaTIM-38
<i>Espectros de dicroísmo circular</i>					
<b>Hélices <math>\alpha</math></b>	38%	33%	22%	22%	20%
<b>Hojas <math>\beta</math></b>	18%	22%	30%	29%	30%
<b>Estructura no repetitiva</b>	45%	44%	48%	48%	49%
<i>Espectros de fluorescencia</i>					
<b>Intensidad máx (UA)</b>	264.1	309.8	127.4	127.6	128.8
<b><math>\lambda</math> máx (nm)</b>	302	303	328	326	327
<i>Cinética enzimática a 25 °C</i>					
<b><math>V_{máx}</math> (<math>\mu\text{mol min}^{-1} \text{mg}^{-1}</math>)</b>	1723.4 $\pm$ 46.8	2989.8 $\pm$ 158.7	24.15	31.47	20.41
<b><math>K_M</math> (mM)</b>	0.26 $\pm$ 0.03	0.21 $\pm$ 0.06	1.14	0.99	1.23
<b><math>k_{cat}</math> (<math>\text{min}^{-1}</math>)</b>	(3.67 $\pm$ 0.1) $\times 10^4$	(7.0 $\pm$ 0.4) $\times 10^4$	5.67 $\times 10^2$	7.55 $\times 10^2$	4.67 $\times 10^2$
<b>Eficiencia catalítica (<math>M^{-1} \text{s}^{-1}</math>)</b>	(7.5 $\pm$ 0.2) $\times 10^7$	(1.62 $\pm$ 0.3) $\times 10^8$	2.4 $\times 10^5$	3.69 $\times 10^5$	1.84 $\times 10^5$
<i>Estado oligomérico</i>					
<b>Radio de Stokes (Å)</b>	40.72	40.95	ND	ND	ND
<b>Peso molecular experimental (kDa)</b>	87.82	89.18	ND	ND	ND
<b>Estado oligomérico</b>	Tetramero	Tetramero	ND	ND	ND
<i>Desnaturalización térmica</i>					
<b><math>T_m</math> (°C)</b>	111.6 $\pm$ 0.6	91.7 $\pm$ 0.8	ND	ND	ND
<b><math>\Delta H_D</math> (<math>\text{kJ mol}^{-1}</math>)</b>	717.5 $\pm$ 66.5	1142.5 $\pm$ 38.7	ND	ND	ND
<b><math>E_A</math> (<math>\text{kJ mol}^{-1}</math>)</b>	829.6 $\pm$ 131.7	575.1 $\pm$ 5.3	ND	ND	ND



Parámetro	MjaTIM-85	MfoTIM-75	MokTIM-60	MaeTIM-46	MmaTIM-38
<i>Desnaturalización química con GdnHCl</i>					
<b>C<sub>m</sub></b> (M)	3.44, 4.6, 5.26, 5.5	4.28 ± 0.01	ND	ND	ND
<b>m</b> (kJ mol <sup>-1</sup> M <sup>-1</sup> )	ND	22.8 ± 1.2	ND	ND	ND
<b>ΔG<sub>H2O</sub></b> (kJ mol <sup>-1</sup> )	ND	98 ± 5.1	ND	ND	ND
<i>Unión a PGH</i>					
<b>ΔH<sub>u</sub> a 25 °C</b> (kJ mol <sup>-1</sup> )	5.0	-15.4	ND	ND	ND
<b>-TΔS<sub>u</sub> a 25 °C</b> (kJ mol <sup>-1</sup> )	-34.2	-12.4	ND	ND	ND
<b>ΔG<sub>u</sub> a 25 °C</b> (kJ mol <sup>-1</sup> )	-29.2	-27.8	ND	ND	ND
<b>ΔC<sub>pu</sub></b> (kJ mol <sup>-1</sup> K <sup>-1</sup> )	-1.17	-0.796	ND	ND	ND

## CONCLUSIONES



**Figura 33. Relación de las características biofísicas caracterizadas con la filogenia de la TIM de Methanococcales.** La topología del árbol corresponde a la obtenida en la filogenia de la TIM de arqueas (ver material suplementario A). En color azul se indican variables categóricas, en verde variables numéricas y en rojo se indica un problema compartido en la expresión heteróloga. Las flechas verdes indican si el valor numérico es mayor o menor respecto a la rama contigua. DQ, desnaturalización química. DT, desnaturalización térmica. TD, termodinámica.

Para un estudio comparativo es importante determinar cuál fue la sucesión temporal con la que se adquirieron las características descritas experimentalmente durante la filogenia de los organismos estudiados. En la Figura 33 se observa la relación propuesta entre las características biofísicas determinadas experimentalmente exploradas en este trabajo con respecto a la filogenia de la TIM de Methanococcales.

La actividad baja de las MesoTIM, de demostrarse que sea el caso y no un problema de plegamiento, resulta de gran importancia, pues tiene implicaciones evolutivas importantes. Desde el argumento metabólico, no se puede explicar esta diferencia en la actividad porque las TermoTIM son muy activas incluso a temperatura ambiente, por lo que no parece tener un origen adaptativo. Tampoco se podría argumentar a favor de la adaptación al ambiente mesófilo, pues MokTIM-60 es de origen termófilo y la actividad baja es una característica compartida con MaeTIM-46 y MmaTIM-38. Sin embargo, las dificultades en la expresión soluble y la purificación limitan la información que se puede obtener de estas proteínas.

En cuanto al estado oligomérico, dado que otras TIM de arqueas presentan la forma tetramérica, es razonable suponer que ésta es una característica simplesiomórfica. Lo anterior, por supuesto, hasta no tener evidencia de lo contrario. Además, MfoTIM-75 y MjaTIM-85, pese a que ambas son tetrámeras, poseen valores de  $T_m$  muy diferentes entre ellas, por lo que el cómo se modula la termoestabilidad en estas enzimas parece tener un origen estructural distinto a la tetramerización. Obtener información cristalográfica de MfoTIM-75 será determinante para conocer con precisión cuáles factores se pueden asociar a este cambio. Aunque sólo se presentan dos datos, parece haber una relación entre la temperatura óptima de crecimiento del organismo y el valor de la  $T_m$  de su TIM. De esta manera, la termoestabilidad podría presentarse como una adaptación desde el punto de vista del valor de la  $T_m$ , mas no así de los cambios estructurales específicos que tienen que suceder durante la filogenia para alcanzarlo.

Sobre el tema de la termoestabilidad, en la literatura al respecto de proteínas termófilas se encuentran descritos dos tipos de moléculas, aquellas que no poseen actividad a baja temperatura porque adquieren una rigidez estructural que no les permite llevar a cabo adecuadamente su función, y aquellas personas que siguen realizando la catálisis de manera efectiva en estas condiciones. Claramente, MfoTIM-75 y MjaTIM-85 pertenecen a este último grupo. Esto tiene implicaciones evolutivas, ya que, como se mencionó anteriormente, se ha relacionado a la termoestabilidad con la evolucionabilidad. Desde el punto de vista funcional, el que tengan actividad elevada a baja temperatura puede implicar que, en el proceso evolutivo de los Methanococcales, estas secuencias de TIM no requieren pasar por muchas modificaciones que alteren su actividad o su flexibilidad, por lo que las diferencias observadas pueden tener un origen neutro, al menos en cuanto a temperatura se refiere. De ser así, se debe considerar cómo es que la neutralidad del cambio afecta la variabilidad de otras características biofísicas que no están relacionadas directamente con la actividad como lo es, en este caso, la energía de activación del desplegamiento térmico, la estabilidad ante la desnaturalización química, la irreversibilidad, así como las vías de plegamiento. De manera indirecta, esto también obliga a replantearse si estas características pueden analizarse como “partes” evolutivas independientes del un fenotipo de una proteína, o si cambian en consecuencia de otros factores. Desafortunadamente, la falta de datos de las especies mesófilas seleccionadas no permite hacer conclusiones más categóricas al respecto, pero se plantea que su origen no es adaptativo hasta este punto del estudio.

Los datos de unión a ligando, si bien son preliminares, resultaron en una observación valiosa para el estudio histórico de estas proteínas, ya que el tipo de unión no afecta la actividad desde el punto de vista energético —sucede igual si es exotérmica o endotérmica—, pero resulta curioso que en todas las TIM en las cuales se tiene datos la unión es exotérmica en el intervalo de temperatura de

crecimiento del organismo, por lo que no es del todo claro si esto tiene una razón histórica particular o si es una simple coincidencia. Estudios comparativos al respecto pueden dar mucha información sobre el tema.

La hipótesis inicial, enmarcada en una aproximación heurística, resultó adecuada para comenzar a aproximarse a qué otros agentes causales pueden estar implicados en la variación de algunas de las características exploradas en este trabajo, aunque aún pueden hacerse preguntas acerca del rol concreto que las fuerzas evolutivas pueden tener en los cambios históricos de las proteínas termófilas. Al respecto, en *The Dialectical Biologist* (1987), Richard Levins y Richard Lewontin argumentan:

[...] cada característica tiene propiedades adicionales además de aquellas que fueron seleccionadas inicialmente. Estas propiedades —las consecuencias no seleccionadas de la selección— crean nuevas posibilidades y nuevas vulnerabilidades, y bajo condiciones distintas, *estas mismas propiedades pueden convertirse en el objeto principal de la selección*. Además, el significado evolutivo de una característica puede cambiar drásticamente de grupo en grupo o a través del tiempo. Los huesos en el oído medio de los mamíferos formaron parte de la mandíbula de nuestros ancestros y, más atrás, tuvieron su origen en las agallas de los peces (p. 41).<sup>16</sup>

De manera análoga, no se puede descartar que la variación presente en algunas características biofísicas sea sólo una consecuencia imprevista de otros cambios en la proteína que se dan como una respuesta adaptativa. Del mismo modo, tampoco se puede negar que estas características puedan adquirir un rol adaptativo bajo otras circunstancias ambientales. En cualquier caso, este planteamiento recalca la importancia de realizar estudios específicos que nos permitan determinar la historia evolutiva —las causas últimas— del origen de la variación en proteínas.

---

<sup>16</sup> La traducción del inglés original y el énfasis son míos.

## PERSPECTIVAS

- Implementar otras estrategias para mejorar la purificación de las MesoTIM (v. gr. columnas de interacción hidrofóbica).
- Realizar duplicados y triplicados de las curvas de desnaturalización para definir con mayor precisión el mecanismo de desplegamiento de MjaTIM-85 sin histidinas
- Refinar la obtención de datos y hacer triplicados de las firmas termodinámicas de unión a PGH de MjaTIM-85 y MfoTIM-75.
- Concluir los experimentos de unión a PGH con sacarosa a 20 y 50 °C para analizar el efecto del agua en las cavidades durante la unión
- Obtener estructuras cristalográficas con y sin ligando de las TermoTIM.
- Plantear experimentos que consideren el contexto poblacional para obtener información directa sobre el efecto que algunas características biofísicas tienen sobre la adecuación de un organismo.

## REFERENCIAS

- Agilent Technologies. (2015). ArtixExpress Competent Cells and ArtixExpress (DE3) Competent Cells Instruction Manual. Recuperado el 20 de mayo de 2020, a partir de <https://www.agilent.com/cs/library/usermanuals/Public/230191.pdf>
- Albery, W. J., & Knowles, J. R. (1976). Evolution of enzyme function and the development of catalytic efficiency. *Biochemistry*, *15*(25), 5631–5640. <http://doi.org/10.1021/bi00670a032>
- Alvarez, M., Zeelen, J. P., Mainfroid, V., Rentier-Delrue, F., Martial, J. A., Wyns, L., ... Maes, D. (1998). Triosephosphate Isomerase (TIM) of the Psychrophilic Bacterium *Vibrio marinus*: KINETIC AND STRUCTURAL PROPERTIES. *Journal of Biological Chemistry*, *273*(4), 2199–2206. <http://doi.org/10.1074/jbc.273.4.2199>
- Andrews, B. T., Capraro, D. T., Sulkowska, J. I., Onuchic, J. N., & Jennings, P. A. (2013). Hysteresis as a Marker for Complex, Overlapping Landscapes in Proteins. *The Journal of Physical Chemistry Letters*, *4*(1), 180–188. <http://doi.org/10.1021/jz301893w>
- Anfinsen, C. B., Haber, E., Sela, M., & White, F. H. J. (1961). The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain. *Proceedings of the National Academy of Sciences of the United States of America*, *47*(9), 1309–1314. <http://doi.org/10.1073/pnas.47.9.1309>
- Arakawa, T., & Timasheff, S. N. (1984). Protein stabilization and destabilization by guanidinium salts. *Biochemistry*, *23*(25), 5924–5929. <http://doi.org/10.1021/bi00320a005>
- Aslantas, Y., & Surmeli, N. B. (2019). Effects of N-Terminal and C-Terminal Polyhistidine Tag on the Stability and Function of the Thermophilic P450 CYP119. *Bioinorganic Chemistry and Applications*, *2019*, 8080697. <http://doi.org/10.1155/2019/8080697>
- Baranauskiene, L., Kuo, T.-C., Chen, W.-Y., & Matulis, D. (2019). Isothermal titration calorimetry for characterization of recombinant proteins. *Current Opinion in Biotechnology*, *55*, 9–15. <http://doi.org/10.1016/J.COPBIO.2018.06.003>
- Battistuzzi, F. U., & Hedges, S. B. (2009). Archaeobacteria. En S. B. Hedges & S. Kumar (Eds.), *The Timetree of life* (pp. 101–105). New York, NY: Oxford University Press.
- Beadle, B. M., & Shoichet, B. K. (2002). Structural Bases of Stability–function Tradeoffs in Enzymes. *Journal of Molecular Biology*, *321*(2), 285–296. [http://doi.org/https://doi.org/10.1016/S0022-2836\(02\)00599-5](http://doi.org/https://doi.org/10.1016/S0022-2836(02)00599-5)
- Beaucamp, N., Hofmann, A., Kellerer, B., & Jaenicke, R. (1997). 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*, *6*(10), 2159–2165. <http://doi.org/10.1002/pro.5560061010>
- Beaucamp, N., Schurig, H., & Jaenicke, R. (1997). The PGK-TIM fusion protein from *Thermotoga maritima* and its constituent parts are intrinsically stable and fold independently. *Biological Chemistry*, *378*(7), 679–685. <http://doi.org/10.1515/bchm.1997.378.7.679>
- Bershtein, S., Segal, M., Bekerman, R., Tokuriki, N., & Tawfik, D. S. (2006). Robustness–epistasis link shapes the fitness landscape of a randomly drifting protein. *Nature*, *444*(7121), 929–932. <http://doi.org/10.1038/nature05385>
- Bio-Rad Laboratories. (s/f). Gel Filtration Standard Instruction Manual. Recuperado el 20 de mayo de 2020, a partir de <http://www.bio-rad.com/webroot/web/pdf/lsr/literature/MSLIT-102E.pdf>
- Bloom, J. D., Labthavikul, S. T., Otey, C. R., & Arnold, F. H. (2006). Protein stability promotes evolvability. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(15), 5869–5874. <http://doi.org/10.1073/pnas.0510098103>

- Booth, W. T., Schlachter, C. R., Pote, S., Ussin, N., Mank, N. J., Klapper, V., ... Chruszcz, M. (2018). Impact of an N-terminal Polyhistidine Tag on Protein Thermal Stability. *ACS Omega*, 3(1), 760–768. <http://doi.org/10.1021/acsomega.7b01598>
- Bräsen, C., Esser, D., Rauch, B., & Siebers, B. (2014). Carbohydrate metabolism in Archaea: current insights into unusual enzymes and pathways and their regulation. *Microbiology and molecular biology reviews : MMBR*, 78(1), 89–175. <http://doi.org/10.1128/MMBR.00041-13>
- Brautigam, C. A., Zhao, H., Vargas, C., Keller, S., & Schuck, P. (2016). Integration and global analysis of isothermal titration calorimetry data for studying macromolecular interactions. *Nature Protocols*, 11(5), 882–894. <http://doi.org/10.1038/nprot.2016.044>
- Bringer, C., Spradlin, S., Cobani, L., & Evilia, C. (2018). The more adaptive to change, the more likely you are to survive: Protein adaptation in extremophiles. *Seminars in Cell & Developmental Biology*, 84, 158–169. <http://doi.org/10.1016/j.semdb.2017.12.016>
- Chang, R. (2010). *Química* (10ma ed.). McGraw-Hill.
- Chu, C.-H., Lai, Y.-J., Huang, H., & Sun, Y.-J. (2008). Kinetic and structural properties of triosephosphate isomerase from *Helicobacter pylori*. *Proteins*, 71(1), 396–406. <http://doi.org/10.1002/prot.21709>
- Claveria-Gimeno, R., Vega, S., Abian, O., & Velazquez-Campoy, A. (2017). A look at ligand binding thermodynamics in drug discovery. *Expert Opinion on Drug Discovery*, 12(4), 363–377. <http://doi.org/10.1080/17460441.2017.1297418>
- Costas, M., Rodríguez-Larrea, D., De Maria, L., Borchert, T. V., Gómez-Puyou, A., & Sanchez-Ruiz, J. M. (2009). Between-Species Variation in the Kinetic Stability of TIM Proteins Linked to Solvation-Barrier Free Energies. *Journal of Molecular Biology*, 385(3), 924–937. <http://doi.org/10.1016/j.jmb.2008.10.056>
- Davis, J. M., & Bächinger, H. P. (1993). Hysteresis in the triple helix-coil transition of type III collagen. *The Journal of Biological Chemistry*, 268(34), 25965–25972.
- DePristo, M. A., Weinreich, D. M., & Hartl, D. L. (2005). Missense meanderings in sequence space: a biophysical view of protein evolution. *Nature Reviews Genetics*, 6(9), 678–687. <http://doi.org/10.1038/nrg1672>
- Despa, F., Orgill, D. P., & Lee, R. C. (2005). Effects of Crowding on the Thermal Stability of Heterogeneous Protein Solutions. *Annals of Biomedical Engineering*, 33(8), 1125–1131. <http://doi.org/10.1007/s10439-005-5780-8>
- Dhaunta, N., Arora, K., Chandrayan, S. K., & Guptasarma, P. (2013). Introduction of a thermophile-sourced ion pair network in the fourth beta/alpha unit of a psychrophile-derived triosephosphate isomerase from *Methanococcus burtonii* significantly increases its kinetic thermal stability. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 1834(6), 1023–1033. <http://doi.org/http://dx.doi.org/10.1016/j.bbapap.2013.01.001>
- Diamant, S., Eliahu, N., Rosenthal, D., & Goloubinoff, P. (2001). Chemical chaperones regulate molecular chaperones in vitro and in cells under combined salt and heat stresses. *The Journal of biological chemistry*, 276(43), 39586–91. <http://doi.org/10.1074/jbc.M103081200>
- Dill, K. A., & MacCallum, J. L. (2012). The Protein-Folding Problem, 50 Years On. *Science*, 338(6110), 1042 LP-1046. <http://doi.org/10.1126/science.1219021>
- Dirección General de Comunicación Social. (2020). La UNAM informa. Recuperado el 20 de mayo de 2020, a partir de [https://www.dgcs.unam.mx/boletin/bdboletin/2020\\_236bis.html](https://www.dgcs.unam.mx/boletin/bdboletin/2020_236bis.html)
- Dong, Y., Liao, M., Meng, X., & Somero, G. N. (2018). Structural flexibility and protein adaptation to temperature: Molecular dynamics analysis of malate dehydrogenases of marine molluscs. *Proceedings of*

*the National Academy of Sciences*, 115(6), 1274 LP-1279. <http://doi.org/10.1073/pnas.1718910115>

- Doyle, M. L. (1997). Characterization of binding interactions by isothermal titration calorimetry. *Current Opinion in Biotechnology*, 8(1), 31–35. [http://doi.org/10.1016/S0958-1669\(97\)80154-1](http://doi.org/10.1016/S0958-1669(97)80154-1)
- Dumas, P. (2019). Rigorous equations for isothermal titration calorimetry: theoretical and practical consequences. *bioRxiv*, 512780. <http://doi.org/10.1101/512780>
- Eldredge, N. (2008). Hierarchies and the Sloshing Bucket: Toward the Unification of Evolutionary Biology. *Evolution: Education and Outreach*, 1(1), 10–15. <http://doi.org/10.1007/s12052-007-0007-6>
- Fields, P. A. (2001). Review: Protein function at thermal extremes: balancing stability and flexibility. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, 129(2–3), 417–431. [http://doi.org/10.1016/S1095-6433\(00\)00359-7](http://doi.org/10.1016/S1095-6433(00)00359-7)
- Fields, P. A., Dong, Y., Meng, X., & Somero, G. N. (2015). Adaptations of protein structure and function to temperature: there is more than one way to 'skin a cat'. *The Journal of Experimental Biology*, 218(12), 1801 LP-1811. <http://doi.org/10.1242/jeb.114298>
- Finkelstein, A. V. (2018). 50+ Years of Protein Folding. *Biochemistry (Moscow)*, 83(1), S3–S18. <http://doi.org/10.1134/S000629791814002X>
- Fleming, L. (2017). Two Dogmas of Biology. *Philosophy, Theory, and Practice in Biology*. Recuperado a partir de <http://philsci-archive.pitt.edu/14523/>
- Forsyth, W. R., & Matthews, C. R. (2002). Folding Mechanism of Indole-3-glycerol Phosphate Synthase from *Sulfolobus solfataricus*: A Test of the Conservation of Folding Mechanisms Hypothesis in ( $\beta\alpha$ )<sub>8</sub> Barrels. *Journal of Molecular Biology*, 320(5), 1119–1133. [http://doi.org/10.1016/S0022-2836\(02\)00557-0](http://doi.org/10.1016/S0022-2836(02)00557-0)
- Frasca, V. (2016). Biophysical characterization of antibodies with isothermal titration calorimetry. *Journal of Applied Bioanalysis*, 2(3), 90–102. <http://doi.org/10.17145/jab.16.013>
- Freire, E. (2008). Do enthalpy and entropy distinguish first in class from best in class? *Drug Discovery Today*, 13(19–20), 869–874. <http://doi.org/10.1016/j.drudis.2008.07.005>
- Gardner, N. W., Monroe, L. K., Kihara, D., & Park, C. (2016). Energetic Coupling between Ligand Binding and Dimerization in *Escherichia coli* Phosphoglycerate Mutase. *Biochemistry*, 55(12), 1711–1723. <http://doi.org/10.1021/acs.biochem.5b00980>
- Gayathri, P., Banerjee, M., Vijayalakshmi, A., Azeez, S., Balaram, H., Balaram, P., & Murthy, M. R. N. (2007). Structure of triosephosphate isomerase (TIM) from *Methanocaldococcus jannaschii*. *Acta Crystallographica Section D*, 63(2), 206–220. Recuperado a partir de <https://doi.org/10.1107/S0907444906046488>
- GE Healthcare. (s/f). Size exclusion chromatography. Principles and Methods. Recuperado el 1 de agosto de 2020, a partir de <https://cdn.gelifesciences.com/dmm3bwsv3/AssetStream.aspx?mediaformatid=10061&destinationid=10016&assetid=11639>
- Gell, D. A., Grant, R. P., & Mackay, J. P. (2012). The Detection and Quantitation of Protein Oligomerization. En J. M. Matthews (Ed.), *Protein Dimerization and Oligomerization in Biology* (pp. 19–41). New York, NY: Springer New York. [http://doi.org/10.1007/978-1-4614-3229-6\\_2](http://doi.org/10.1007/978-1-4614-3229-6_2)
- Godfrey-Smith, P. (2001). Three Kinds of Adaptationism. En E. Sober & S. H. Orzack (Eds.), *Adaptationism and Optimality* (pp. 335–357). Cambridge: Cambridge University Press. <http://doi.org/DOI:10.1017/CBO9780511609084.012>
- Gómez-Puyou, A., Saavedra-Lira, E., Becker, I., Zubillaga, R. A., Rojo-Domínguez, A., & Pérez-Montfort, R.



- (1995). Using evolutionary changes to achieve species-specific inhibition of enzyme action — studies with triosephosphate isomerase. *Chemistry & Biology*, 2(12), 847–855. [http://doi.org/10.1016/1074-5521\(95\)90091-8](http://doi.org/10.1016/1074-5521(95)90091-8)
- Gould, S. J., & Lewontin, R. C. (1979). The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 205(1161), 581 LP-598. Recuperado a partir de <http://rspb.royalsocietypublishing.org/content/205/1161/581.abstract>
- Goyal, N., Zhou, Z., & Karimi, I. A. (2016). Metabolic processes of *Methanococcus maripaludis* and potential applications. *Microbial Cell Factories*, 15(1), 107. <http://doi.org/10.1186/s12934-016-0500-0>
- Greene, R. F. J., & Pace, C. N. (1974). Urea and guanidine hydrochloride denaturation of ribonuclease, lysozyme, alpha-chymotrypsin, and beta-lactoglobulin. *The Journal of Biological Chemistry*, 249(17), 5388–5393.
- Greenfield, N. J. (2007). Using circular dichroism spectra to estimate protein secondary structure. *Nat. Protocols*, 1(6), 2876–2890. Recuperado a partir de <http://dx.doi.org/10.1038/nprot.2006.202>
- Han, J.-H., Batey, S., Nickson, A. A., Teichmann, S. A., & Clarke, J. (2007). The folding and evolution of multidomain proteins. *Nature Reviews Molecular Cell Biology*, 8(4), 319–330. <http://doi.org/10.1038/nrm2144>
- Harms, M. J., & Thornton, J. W. (2013). Evolutionary biochemistry: revealing the historical and physical causes of protein properties. *Nature reviews. Genetics*, 14(8), 559–71. <http://doi.org/10.1038/nrg3540>
- Hart, K. M., Harms, M. J., Schmidt, B. H., Elya, C., Thornton, J. W., & Marqusee, S. (2014). Thermodynamic system drift in protein evolution. *PLoS biology*, 12(11), e1001994. <http://doi.org/10.1371/journal.pbio.1001994>
- Holdgate, G. A., & Ward, W. H. J. (2005). Measurements of binding thermodynamics in drug discovery. *Drug Discovery Today*, 10(22), 1543–1550. [http://doi.org/10.1016/S1359-6446\(05\)03610-X](http://doi.org/10.1016/S1359-6446(05)03610-X)
- Hong, P., Koza, S., & Bouvier, E. S. P. (2012). Size-Exclusion Chromatography for the Analysis of Protein Biotherapeutics and their Aggregates. *Journal of Liquid Chromatography & Related Technologies*, 35(20), 2923–2950. <http://doi.org/10.1080/10826076.2012.743724>
- Horikoshi, K., Takai, K., & Inoue, A. (2002). *Methanothermococcus okinawensis* sp. nov., a thermophilic, methane-producing archaeon isolated from a Western Pacific deep-sea hydrothermal vent system. *International Journal of Systematic and Evolutionary Microbiology*, 52(4), 1089–1095. <http://doi.org/10.1099/00207713-52-4-1089>
- Hribar, B., Southall, N. T., Vlachy, V., & Dill, K. A. (2002). How Ions Affect the Structure of Water. *Journal of the American Chemical Society*, 124(41), 12302–12311. <http://doi.org/10.1021/ja026014h>
- Ibarra-Molero, B., & Sanchez-Ruiz, J. M. (2006). Differential Scanning Calorimetry of Proteins: an Overview and Some Recent Developments BT - Advanced Techniques in Biophysics. En J. L. R. Arrondo & A. Alonso (Eds.), (pp. 27–48). Berlin, Heidelberg: Springer Berlin Heidelberg. [http://doi.org/10.1007/3-540-30786-9\\_2](http://doi.org/10.1007/3-540-30786-9_2)
- Ichikawa, J. K., & Clarke, S. (1998). A highly active protein repair enzyme from an extreme thermophile: the L-isoaspartyl methyltransferase from *Thermotoga maritima*. *Archives of Biochemistry and Biophysics*, 358(2), 222–231. <http://doi.org/10.1006/abbi.1998.0830>
- Jaenicke, R., & Sterner, R. (2013). Life at High Temperatures. En E. Rosenberg, E. F. DeLong, S. Lory, E. Stackebrandt, & F. Thompson (Eds.), *The Prokaryotes: Prokaryotic Communities and Ecophysiology* (pp. 337–374). Berlin, Heidelberg: Springer Berlin Heidelberg. [http://doi.org/10.1007/978-3-642-30123-0\\_55](http://doi.org/10.1007/978-3-642-30123-0_55)
- Jelesarov, I., & Bosshard, H. R. (1999). Isothermal titration calorimetry and differential scanning calorimetry as

complementary tools to investigate the energetics of biomolecular recognition. *Journal of Molecular Recognition*, 12(1), 3–18. [http://doi.org/10.1002/\(SICI\)1099-1352\(199901/02\)12:1<3::AID-JMR441>3.0.CO;2-6](http://doi.org/10.1002/(SICI)1099-1352(199901/02)12:1<3::AID-JMR441>3.0.CO;2-6)

- Johnson, C. M. (2013). Differential scanning calorimetry as a tool for protein folding and stability. *Archives of Biochemistry and Biophysics*, 531(1–2), 100–109. <http://doi.org/http://dx.doi.org/10.1016/j.abb.2012.09.008>
- Jones, W. J., Leigh, J. A., Mayer, F., Woese, C. R., & Wolfe, R. S. (1983). *Methanococcus jannaschii* sp. nov., an extremely thermophilic methanogen from a submarine hydrothermal vent. *Archives of Microbiology*, 136(4), 254–261. <http://doi.org/10.1007/BF00425213>
- Jones, W. J., Paynter, M. J. B., & Gupta, R. (1983). Characterization of *Methanococcus maripaludis* sp. nov., a new methanogen isolated from salt marsh sediment. *Archives of Microbiology*, 135(2), 91–97. <http://doi.org/10.1007/BF00408015>
- Joubert, F., Neitz, A. W. H., & Louw, A. I. (2001). Structure-based inhibitor screening: A family of sulfonated dye inhibitors for malaria parasite triosephosphate isomerase. *Proteins: Structure, Function, and Bioinformatics*, 45(2), 136–143. <http://doi.org/10.1002/prot.1133>
- Kandror, O., Sherman, M., Rhode, M., & Goldberg, A. L. (1995). Trigger factor is involved in GroEL-dependent protein degradation in *Escherichia coli* and promotes binding of GroEL to unfolded proteins. *The EMBO Journal*, 14(23), 6021–6027.
- Kassambara, A. (2020). ggpubr: “ggplot2” Based Publication Ready Plots. Recuperado a partir de <https://cran.r-project.org/package=ggpubr>
- Kelly, S. M., Jess, T. J., & Price, N. C. (2005). How to study proteins by circular dichroism. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 1751(2), 119–139. <http://doi.org/https://doi.org/10.1016/j.bbapap.2005.06.005>
- Kendall, M. M., Liu, Y., Sieprawska-Lupa, M., Stetter, K. O., Whitman, W. B., & Boone, D. R. (2006). *Methanococcus aeolicus* sp. nov., a mesophilic, methanogenic archaeon from shallow and deep marine sediments. *International Journal of Systematic and Evolutionary Microbiology*, 56(7), 1525–1529. Recuperado a partir de <http://ijs.microbiologyresearch.org/content/journal/ijsem/10.1099/ijs.0.64216-0>
- Kim, D. E., Chivian, D., & Baker, D. (2004). Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Research*, 32(Web Server issue), W526–W531. <http://doi.org/10.1093/nar/gkh468>
- Kohlhoff, M., Dahm, A., & Hensel, R. (1996). Tetrameric triosephosphate isomerase from hyperthermophilic Archaea. *FEBS Letters*, 383(3), 245–250. [http://doi.org/10.1016/0014-5793\(96\)00249-9](http://doi.org/10.1016/0014-5793(96)00249-9)
- Krietsch, W. K. G., Pentchev, P. G., Klingenburg, H., Hofstätter, T., & Bücher, T. (1970). 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*, 14(2), 289–300. <http://doi.org/10.1111/j.1432-1033.1970.tb00289.x>
- Krissinel, E., & Henrick, K. (2007). Inference of Macromolecular Assemblies from Crystalline State. *Journal of Molecular Biology*, 372(3), 774–797. <http://doi.org/https://doi.org/10.1016/j.jmb.2007.05.022>
- Kumar, S., & Nussinov, R. (2001). How do thermophilic proteins deal with heat? *Cellular and Molecular Life Sciences CMLS*, 58(9), 1216–1233. <http://doi.org/10.1007/PL00000935>
- Lakowicz, J. R. (2006). *Principles of Fluorescence Spectroscopy* (3ra ed.). Boston, MA: Springer-Verlag. <http://doi.org/https://doi.org/10.1007/978-0-387-46312-4>
- Lanfear, R., Calcott, B., Ho, S. Y. W., & Guindon, S. (2012). PartitionFinder: Combined Selection of Partitioning Schemes and Substitution Models for Phylogenetic Analyses. *Molecular Biology and Evolution*, 29(6), 1695–1701. <http://doi.org/10.1093/molbev/mss020>

- Lanfear, R., Frandsen, P. B., Wright, A. M., Senfeld, T., & Calcott, B. (2016). PartitionFinder 2: New Methods for Selecting Partitioned Models of Evolution for Molecular and Morphological Phylogenetic Analyses. *Molecular Biology and Evolution*, 34(3), msw260. <http://doi.org/10.1093/molbev/msw260>
- Le, S. Q., Dang, C. C., & Gascuel, O. (2012). Modeling Protein Evolution with Several Amino Acid Replacement Matrices Depending on Site Rates. *Molecular Biology and Evolution*, 29(10), 2921–2936. <http://doi.org/10.1093/molbev/mss112>
- Levins, R., & Lewontin, R. C. (1987). *The Dialectical Biologist*. Harvard University Press.
- Lewens, T. (2009). Seven types of adaptationism. *Biology & Philosophy*, 24, 161–182. <http://doi.org/10.1007/s10539-008-9145-7>
- Lewontin, R. C. (1978). Adaptation. *Scientific American*, 239(3), 212–218.
- Lin, Z., Schwarz, F. P., & Eisenstein, E. (1995). The Hydrophobic Nature of GroEL-Substrate Binding. *Journal of Biological Chemistry*, 270(3), 1011–1014. <http://doi.org/10.1074/JBC.270.3.1011>
- Lindberg, M., Tångrot, J., & Oliveberg, M. (2002). Complete change of the protein folding transition state upon circular permutation. *Nature Structural Biology*, 9(11), 818–822. <http://doi.org/10.1038/nsb847>
- Lumry, R., & Eyring, H. (1954). Conformation Changes of Proteins. *The Journal of Physical Chemistry*, 58(2), 110–120. <http://doi.org/10.1021/j150512a005>
- Lynch, M. (2007). The frailty of adaptive hypotheses for the origins of organismal complexity. *Proceedings of the National Academy of Sciences*, 104(suppl 1), 8597 LP-8604. <http://doi.org/10.1073/pnas.0702207104>
- Mason, P. E., Brady, J. W., Neilson, G. W., & Dempsey, C. E. (2017). The Interaction of Guanidinium Ions with a Model Peptide. *Biophysical Journal*, 93(1), L04–L06. <http://doi.org/10.1529/biophysj.107.108290>
- Mateus, A., Bobonis, J., Kurzawa, N., Stein, F., Helm, D., Hevler, J., ... Savitski, M. M. (2018). Thermal proteome profiling in bacteria: probing protein state in vivo. *Molecular systems biology*, 14(7), e8242. <http://doi.org/10.15252/msb.20188242>
- Maynard Smith, J. (1970). Natural selection and the concept of a protein space. *Nature*, 225(5232), 563–564. <http://doi.org/10.1038/225563a0>
- Mayr, E. (1961). Cause and Effect in Biology. *Science*, 134(3489), 1501 LP-1506. <http://doi.org/10.1126/science.134.3489.1501>
- Merck. (2020). Glycerol-3-Phosphate Dehydrogenase (GDH) from rabbit muscle. Recuperado el 29 de mayo de 2020, a partir de [https://www.sigmaaldrich.com/catalog/product/roche/gdhro?lang=es&region=MX&gclid=CjwKCAjwq832BRA5EiwACvCWsfECie8iL57ExfDDFsprQrZhiGxmfXm2H\\_zlhcGq88UpcYwt1eRwbx0CaOcQAvD\\_BwE](https://www.sigmaaldrich.com/catalog/product/roche/gdhro?lang=es&region=MX&gclid=CjwKCAjwq832BRA5EiwACvCWsfECie8iL57ExfDDFsprQrZhiGxmfXm2H_zlhcGq88UpcYwt1eRwbx0CaOcQAvD_BwE)
- Merz, A., Knöchel, T., Jansonius, J. N., & Kirschner, K. (1999). The hyperthermostable indoleglycerol phosphate synthase from *Thermotoga maritima* is destabilized by mutational disruption of two solvent-exposed salt bridges. *Journal of Molecular Biology*, 288(4), 753–763. <http://doi.org/10.1006/jmbi.1999.2709>
- MicroCal. (1998). DSC Data Analysis in Origin Tutorial Guide Version 5.0. Recuperado el 20 de mayo de 2020, a partir de <https://ctrstbio.org.uic.edu/manuals/dscOrigin.pdf>
- Miller, M. A., Pfeiffer, W., & Schwartz, T. (2010). Creating the CIPRES Science Gateway for inference of large phylogenetic trees. *2010 Gateway Computing Environments Workshop, GCE 2010*. <http://doi.org/10.1109/GCE.2010.5676129>
- Minajeva, A., Kulke, M., Fernandez, J. M., & Linke, W. A. (2001). Unfolding of titin domains explains the viscoelastic behavior of skeletal myofibrils. *Biophysical Journal*, 80(3), 1442–1451.

[http://doi.org/10.1016/S0006-3495\(01\)76116-4](http://doi.org/10.1016/S0006-3495(01)76116-4)

- Muñiz Trejo, R. (2017). *Producción y caracterización fisicoquímica comparativa de la triosafosfato isomerasa de tres arqueas del orden Methanococcales*. México, 89 p. (Asesor: Daniel Alejandro Fernández Velasco, Tesis de licenciatura en Biología, Facultad de Ciencias, Universidad Nacional Autónoma de México).
- Myers, J. K., Pace, C. N., & Scholtz, J. M. (1995). Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Science : A Publication of the Protein Society*, 4(10), 2138–2148. <http://doi.org/10.1002/pro.5560041020>
- Nájera, H., Costas, M., & Fernández-Velasco, D. (2003). Thermodynamic characterization of yeast triosephosphate isomerase refolding: Insights into the interplay between function and stability as reasons for the oligomeric nature of the enzyme. *The Biochemical journal*, 370, 785–792. <http://doi.org/10.1042/BJ20021439>
- Newton, M. S., Arcus, V. L., & Patrick, W. M. (2015). Rapid bursts and slow declines: on the possible evolutionary trajectories of enzymes. *Journal of The Royal Society Interface*, 12(107), 20150036–20150036. <http://doi.org/10.1098/rsif.2015.0036>
- Ogbunugafor, C. B. (2020). A Reflection on 50 Years of John Maynard Smith's "Protein Space". *Genetics*, 214(4), 749 LP-754. <http://doi.org/10.1534/genetics.119.302764>
- Oliveira, C., & Domingues, L. (2018). Guidelines to reach high-quality purified recombinant proteins. *Applied Microbiology and Biotechnology*, 102(1), 81–92. <http://doi.org/10.1007/s00253-017-8623-8>
- Olson, M. E., & Arroyo-Santos, A. (2015). How to study adaptation (and why to do it that way). *The Quarterly Review of Biology*, 90(2), 167–191.
- Olson, M. E., & Rosell, J. A. (2015). Adaptación. En A. Becerra Bracho, A. Castañeda Sortibrán, & D. Piñero Dalmau (Eds.), *Evolución orgánica* (pp. 222–279). Ciudad de México: Universidad Nacional Autónoma de México.
- OriginLab. (s/f-a). Boltzmann. Recuperado el 1 de julio de 2020, a partir de <https://www.originlab.com/doc/origin-help/boltzmann-fitfunc>
- OriginLab. (s/f-b). DoubleBoltzmann. Recuperado el 1 de julio de 2020, a partir de <https://www.originlab.com/doc/Origin-Help/DoubleBoltzmann-FitFunc>
- OriginLab. (s/f-c). Origin. Northampton, MA: OriginLab Corporation. Recuperado a partir de <https://www.originlab.com>
- Park, S. H., Kim, H. S., Park, M. S., Moon, S., Song, M. K., Park, H. S., ... Han, B. W. (2016). Structure and Stability of the Dimeric Triosephosphate Isomerase from the Thermophilic Archaeon *Thermoplasma acidophilum*. *PLoS ONE*, 10(12), e0145331. Recuperado a partir de <http://dx.doi.org/10.1371/journal.pone.0145331>
- Parsegian, V. A., Rand, R. P., & Rau, D. C. B. T.-M. in E. (1995). [3] Macromolecules and water: Probing with osmotic stress. En *Energetics of Biological Macromolecules* (Vol. 259, pp. 43–94). Academic Press. [http://doi.org/https://doi.org/10.1016/0076-6879\(95\)59039-0](http://doi.org/https://doi.org/10.1016/0076-6879(95)59039-0)
- Payne, J. L., & Wagner, A. (2019). The causes of evolvability and their evolution. *Nature Reviews Genetics*, 20(1), 24–38. <http://doi.org/10.1038/s41576-018-0069-z>
- Pei, J., Kim, B. H., & Grishin, N. V. (2008). PROMALS3D: A tool for multiple protein sequence and structure alignments. *Nucleic Acids Research*, 36(7), 2295–2300. <http://doi.org/10.1093/nar/gkn072>
- Peimbert, M., Domínguez-Ramírez, L., Tellez, L. A., Blancas, L. M., & Fernández-Velasco, D. A. (2008). Triosephosphate isomerase. En E. García-Hernández & D. A. Fernández-Velasco (Eds.), *Advances in*

*Protein Physical Chemistry* (pp. 193–217). Transworld Research Network.

- Pigliucci, M. (2008). Is evolvability evolvable? *Nature Reviews Genetics*, 9(1), 75–82. <http://doi.org/10.1038/nrg2278>
- Piñero, Á., Muñoz, E., Sabín, J., Costas, M., Bastos, M., Velázquez-Campoy, A., ... Cotelo, C. (2019). AFFINImeter: A software to analyze molecular recognition processes from experimental data. *Analytical Biochemistry*, 577, 117–134. <http://doi.org/10.1016/j.ab.2019.02.031>
- Plaxco, K. W., Riddle, D. S., Grantcharova, V., & Baker, D. (1998). Simplified proteins: minimalist solutions to the 'protein folding problem'. *Current Opinion in Structural Biology*, 8(1), 80–85. [http://doi.org/https://doi.org/10.1016/S0959-440X\(98\)80013-4](http://doi.org/https://doi.org/10.1016/S0959-440X(98)80013-4)
- Prendergast, F. G., Hampton, P. D., & Jones, B. (1984). Characteristics of tyrosinate fluorescence emission in .alpha.- and .beta.-purothionins. *Biochemistry*, 23(26), 6690–6697. <http://doi.org/10.1021/bi00321a063>
- Privalov, P. L., & Khechinashvili, N. N. (1974). A thermodynamic approach to the problem of stabilization of globular protein structure: a calorimetric study. *Journal of Molecular Biology*, 86(3), 665–684. [http://doi.org/10.1016/0022-2836\(74\)90188-0](http://doi.org/10.1016/0022-2836(74)90188-0)
- Purich, D. L. (2010a). Initial-Rate Kinetics of One-Substrate Enzyme-Catalyzed Reactions. En D. L. B. T.-E. K. C. & C. Purich (Ed.), (pp. 287–334). Boston: Elsevier. <http://doi.org/https://doi.org/10.1016/B978-0-12-380924-7.10005-5>
- Purich, D. L. (2010b). Practical Aspects of Measuring Initial Rates and Reaction Parameters. En D. L. Purich (Ed.), *Enzyme Kinetics: Catalysis & Control. A Reference of Theory and Best-Practice Methods* (pp. 215–285). Boston: Elsevier. <http://doi.org/https://doi.org/10.1016/B978-0-12-380924-7.10004-3>
- Quezada, A. G., Díaz-Salazar, A. J., Cabrera, N., Pérez-Montfort, R., Piñero, Á., & Costas, M. (2017). Interplay between Protein Thermal Flexibility and Kinetic Stability. *Structure*, 25(1), 167–179. <http://doi.org/http://dx.doi.org/10.1016/j.str.2016.11.018>
- R Core Team. (2020). R: A Language and Environment for Statistical Computing. Viena: R Foundation for Statistical Computing. Recuperado a partir de <https://www.r-project.org/>
- Razvi, A., & Scholtz, J. M. (2006). Lessons in stability from thermophilic proteins. *Protein science : a publication of the Protein Society*, 15(7), 1569–78. <http://doi.org/10.1110/ps.062130306>
- Robert, X., & Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Research*, 42(W1), W320–W324. <http://doi.org/10.1093/nar/gku316>
- Rodriguez-Larrea, D., Minning, S., Borchert, T. V., & Sanchez-Ruiz, J. M. (2006). Role of Solvation Barriers in Protein Kinetic Stability. *Journal of Molecular Biology*, 360(3), 715–724. <http://doi.org/https://doi.org/10.1016/j.jmb.2006.05.009>
- Romero-Romero, S., Fernández-Velasco, D. A., & Costas, M. (2018). Estabilidad termodinámica de proteínas. *Educación Química*, 29(3), 3–17. <http://doi.org/10.22201/fq.18708404e.2018.3.64699>
- Romero Romero, S. (2011). *Control cinético y termodinámico en el desplegamiento térmico de la triosafosfato isomerasa proveniente de especies bacterianas*. México, 190 p. (Asesor: Daniel Alejandro Fernández Velasco, Tesis de maestría en Ciencias, Universidad Nacional Autónoma de México).
- Romero Romero, S. (2017). *El paisaje conformacional de las proteínas: estudio de barriles TIM naturales y diseñadas de novo*. México, 368 p. (Asesor: Daniel Alejandro Fernández Velasco Tesis de doctorado en Ciencias, Universidad Nacional Autónoma de México).
- Ronimus, R. S., & Morgan, H. W. (2003). Distribution and phylogenies of enzymes of the Embden-Meyerhof-Parnas pathway from archaea and hyperthermophilic bacteria support a gluconeogenic origin of

metabolism. *Archaea*, 1(3), 199–221. Recuperado a partir de <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2685568/>

Rosenberg, A., & Bouchard, F. (2020). Fitness. En E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 202). Metaphysics Research Lab, Stanford University. Recuperado a partir de <https://plato.stanford.edu/archives/spr2020/entries/fitness/>

Rothschild, L. J., & Mancinelli, R. L. (2001). Life in extreme environments. *Nature*, 409(6823), 1092–1101. Recuperado a partir de <http://dx.doi.org/10.1038/35059215>

Royer, C. A. (2006). Probing Protein Folding and Conformational Transitions with Fluorescence. *Chemical Reviews*, 106(5), 1769–1784. <http://doi.org/10.1021/cr0404390>

Sadeghi, M., Naderi-Manesh, H., Zarrabi, M., & Ranjbar, B. (2006). Effective factors in thermostability of thermophilic proteins. *Biophysical Chemistry*, 119(3), 256–270. <http://doi.org/http://dx.doi.org/10.1016/j.bpc.2005.09.018>

Sanchez-Ruiz, J. (2010). Protein kinetic stability. *Biophysical Chemistry*, 148(1–3), 1–15. <http://doi.org/10.1016/j.bpc.2010.02.004>

Sanchez-Ruiz, J. M. (1992). Theoretical analysis of Lumry-Eyring models in differential scanning calorimetry. *Biophysical Journal*, 61(4), 921–935. [http://doi.org/10.1016/S0006-3495\(92\)81899-4](http://doi.org/10.1016/S0006-3495(92)81899-4)

Sanchez-Ruiz, J. M. (1995). Differential Scanning Calorimetry of Proteins. En B. B. Biswas & S. Roy (Eds.), *Proteins: Structure, Function, and Engineering* (pp. 133–176). Boston, MA: Springer US. [http://doi.org/10.1007/978-1-4899-1727-0\\_6](http://doi.org/10.1007/978-1-4899-1727-0_6)

Santoro, M. M., & Bolen, D. W. (1988). Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl alpha-chymotrypsin using different denaturants. *Biochemistry*, 27(21), 8063–8068. <http://doi.org/10.1021/bi00421a014>

Savitzky, A., & Golay, M. J. E. (1964). Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry*, 36(8), 1627–1639. <http://doi.org/10.1021/ac60214a047>

Sawyer, W. H., & Puckridge, J. (1973). The Dissociation of Proteins by Chaotropic Salts. *Journal of Biological Chemistry*, 248, 8429–8433.

Scalley-Kim, M., & Baker, D. (2004). Characterization of the Folding Energy Landscapes of Computer Generated Proteins Suggests High Folding Free Energy Barriers and Cooperativity may be Consequences of Natural Selection. *Journal of Molecular Biology*, 338(3), 573–583. <http://doi.org/https://doi.org/10.1016/j.jmb.2004.02.055>

Schein, C. H. (1989). Production of Soluble Recombinant Proteins in Bacteria. *Bio/Technology*, 7(11), 1141–1149. <http://doi.org/10.1038/nbt1189-1141>

Serratos, I. N., Millán-Pacheco, C., Garza-Ramos, G., Pérez-Hernández, G., & Zubillaga, R. A. (2018). Exploring interfacial water trapping in protein-ligand complexes with multithermal titration calorimetry. *Biochimica et Biophysica Acta. Proteins and Proteomics*, 1866(3), 488–495. <http://doi.org/10.1016/j.bbapap.2018.01.001>

Serratos, I. N., Pérez-Hernández, G., Garza-Ramos, G., Hernández-Arana, A., González-Mondragón, E., & Zubillaga, R. A. (2011). Binding Thermodynamics of Phosphorylated Inhibitors to Triosephosphate Isomerase and the Contribution of Electrostatic Interactions. *Journal of Molecular Biology*, 405(1), 158–172. <http://doi.org/10.1016/J.JMB.2010.10.018>

Sharma, P., & Guptasarma, P. (2015). ‘Super-perfect’ enzymes: Structural stabilities and activities of recombinant triose phosphate isomerases from *Pyrococcus furiosus* and *Thermococcus onnurineus* produced in *Escherichia coli*. *Biochemical and Biophysical Research Communications* (Vol. 460).

- Shieh, J. S., & Whitman, W. B. (1987). Pathway of acetate assimilation in autotrophic and heterotrophic methanococci. *Journal of Bacteriology*, *169*(11), 5327 LP-5329. <http://doi.org/10.1128/jb.169.11.5327-5329.1987>
- Shieh, J., & Whitman, W. B. (1988). Autotrophic acetyl coenzyme A biosynthesis in *Methanococcus maripaludis*. *Journal of Bacteriology*, *170*(7), 3072 LP-3079. <http://doi.org/10.1128/jb.170.7.3072-3079.1988>
- Sikosek, T., & Chan, H. S. (2014). Biophysics of protein evolution and evolutionary protein biophysics. *Journal of the Royal Society, Interface*, *11*(100), 20140419. <http://doi.org/10.1098/rsif.2014.0419>
- Silverman, J. A., Balakrishnan, R., & Harbury, P. B. (2001). Reverse engineering the ( $\beta/\alpha$ )<sub>8</sub> barrel fold. *Proceedings of the National Academy of Sciences*, *98*(6), 3092–3097. <http://doi.org/10.1073/pnas.041613598>
- Singh, S., & Zlotnick, A. (2003). Observed hysteresis of virus capsid disassembly is implicit in kinetic models of assembly. *The Journal of Biological Chemistry*, *278*(20), 18249–18255. <http://doi.org/10.1074/jbc.M211408200>
- Smith, R. J. (2016). Explanations for adaptations, just-so stories, and limitations on evidence in evolutionary biology. *Evolutionary Anthropology*, *25*(6), 276–287. <http://doi.org/10.1002/evan.21495>
- Somero, G. N. (1995). Proteins and Temperature. *Annual Review of Physiology*, *57*(1), 43–68. <http://doi.org/10.1146/annurev.ph.57.030195.000355>
- Sreerama, N., Venyaminov, S. Y., & Woody, R. W. (2000). Estimation of Protein Secondary Structure from Circular Dichroism Spectra: Inclusion of Denatured Proteins with Native Proteins in the Analysis. *Analytical Biochemistry*, *287*(2), 243–251. <http://doi.org/https://doi.org/10.1006/abio.2000.4879>
- Sreerama, N., & Woody, R. W. (2000). Estimation of Protein Secondary Structure from Circular Dichroism Spectra: Comparison of CONTIN, SELCON, and CDSSTR Methods with an Expanded Reference Set. *Analytical Biochemistry*, *287*(2), 252–260. <http://doi.org/https://doi.org/10.1006/abio.2000.4880>
- Srinivas, V. R., Bhanuprakash Reddy, G., & Surolia, A. (1999). A predominantly hydrophobic recognition of H-antigenic sugars by winged bean acidic lectin: a thermodynamic study. *FEBS Letters*, *450*(3), 181–185. [http://doi.org/10.1016/s0014-5793\(99\)00502-5](http://doi.org/10.1016/s0014-5793(99)00502-5)
- Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*(9), 1312–1313. <http://doi.org/10.1093/bioinformatics/btu033>
- Sterner, R., Kleemann, G. R., Szadkowski, H., Lustig, A., Hennig, M., & Kirschner, K. (1996). Phosphoribosyl anthranilate isomerase from *Thermotoga maritima* is an extremely stable and active homodimer. *Protein Science: A Publication of the Protein Society*, *5*(10), 2000–2008. <http://doi.org/10.1002/pro.5560051006>
- Szilágyi, A., & Závodszy, P. (2000). Structural differences between mesophilic, moderately thermophilic and extremely thermophilic protein subunits: results of a comprehensive survey. *Structure*, *8*(5), 493–504. [http://doi.org/http://dx.doi.org/10.1016/S0969-2126\(00\)00133-7](http://doi.org/http://dx.doi.org/10.1016/S0969-2126(00)00133-7)
- Takai, K., Neelson, K. H., & Horikoshi, K. (2004). *Methanotorris formicicus* sp. nov., a novel extremely thermophilic, methane-producing archaeon isolated from a black smoker chimney in the Central Indian Ridge. *International Journal of Systematic and Evolutionary Microbiology*, *54*(4), 1095–1100. <http://doi.org/10.1099/ijs.0.02887-0>
- Takara Bio, I. (2017). Chaperone Plasmid Set. Recuperado el 19 de noviembre de 2018, a partir de [https://www.takarabio.com/assets/documents/User Manual/3340\\_e.v1701Da.pdf](https://www.takarabio.com/assets/documents/User Manual/3340_e.v1701Da.pdf)
- Taverna, D. M., & Goldstein, R. A. (2002). Why are proteins marginally stable? *Proteins*, *46*(1), 105–109. <http://doi.org/10.1002/prot.10016>

- Thauer, R. K. (2012). The Wolfe cycle comes full circle. *Proceedings of the National Academy of Sciences of the United States of America*, 109(38), 15084–5. <http://doi.org/10.1073/pnas.1213193109>
- Thórólfsson, M., Ibarra-Molero, B., Fojan, P., Petersen, S. B., Sanchez-Ruiz, J. M., & Martínez, A. (2002). L-Phenylalanine Binding and Domain Organization in Human Phenylalanine Hydroxylase: A Differential Scanning Calorimetry Study. *Biochemistry*, 41(24), 7573–7585. <http://doi.org/10.1021/bi0160720>
- Trentham, D. R., McMurray, C. H., & 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. *Biochemical Journal*, 114(1), 19 LP-24. Recuperado a partir de <http://www.biochemj.org/content/114/1/19.abstract>
- True, J. R., & Haag, E. S. (2001). Developmental system drift and flexibility in evolutionary trajectories. *Evolution & Development*, 3(2), 109–119. <http://doi.org/10.1046/j.1525-142x.2001.003002109.x>
- Vera, A., González-Montalbán, N., Arís, A., & Villaverde, A. (2007). The conformational quality of insoluble recombinant proteins is enhanced at low growth temperatures. *Biotechnology and bioengineering*, 96(6), 1101–6. <http://doi.org/10.1002/bit.21218>
- Vieille, C., & Zeikus, G. J. (2001). Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. *Microbiology and Molecular Biology Reviews : MMBR*, 65(1), 1–43. <http://doi.org/10.1128/MMBR.65.1.1-43.2001>
- Vogl, T., Jatzke, C., Hinz, H.-J., Benz, J., & Huber, R. (1997). Thermodynamic Stability of Annexin V E17G: Equilibrium Parameters from an Irreversible Unfolding Reaction. *Biochemistry*, 36(7), 1657–1668. <http://doi.org/10.1021/bi962163z>
- Wagner, A. (2008). Robustness and evolvability: a paradox resolved. *Proceedings of the Royal Society B: Biological Sciences*, 275(1630), 91–100. <http://doi.org/10.1098/rspb.2007.1137>
- Wagner, G. P. (2010). The measurement theory of fitness. *Evolution; International Journal of Organic Evolution*, 64(5), 1358–1376. <http://doi.org/10.1111/j.1558-5646.2009.00909.x>
- Wagner, G. P., & Altenberg, L. (1996). Perspective: Complex Adaptations and the Evolution of Evolvability. *Evolution*, 50(3), 967–976. <http://doi.org/10.2307/2410639>
- Walden, H., Bell, G. S., Russell, R. J. ., Siebers, B., Hensel, R., & Taylor, G. L. (2001). Tiny TIM: a small, tetrameric, hyperthermostable triosephosphate isomerase. *Journal of Molecular Biology*, 306(4), 745–757. <http://doi.org/10.1006/jmbi.2000.4433>
- Walden, H., Taylor, G. L., Lorentzen, E., Pohl, E., Lilie, H., Schramm, A., ... Hensel, R. (2004). Structure and Function of a Regulated Archaeal Triosephosphate Isomerase Adapted to High Temperature. *Journal of Molecular Biology*, 342(3), 861–875. <http://doi.org/10.1016/j.jmb.2004.07.067>
- Walker, J. M. (1996). The Bicinchoninic Acid (BCA) Assay for Protein Quantitation. En J. M. Walker (Ed.), *The Protein Protocols Handbook* (pp. 11–14). Totowa, NJ: Humana Press. [http://doi.org/10.1007/978-1-60327-259-9\\_3](http://doi.org/10.1007/978-1-60327-259-9_3)
- Ward, W. H., & Holdgate, G. A. (2001). Isothermal titration calorimetry in drug discovery. *Progress in Medicinal Chemistry*, 38, 309–376. [http://doi.org/10.1016/s0079-6468\(08\)70097-3](http://doi.org/10.1016/s0079-6468(08)70097-3)
- Whitman, W. B., & Jeanthon, C. (2006). Methanococcales. En M. Dworkin, S. Falkow, E. Rosenberg, K.-H. Schleifer, & E. Stackebrandt (Eds.), *The Prokaryotes: Volume 3: Archaea. Bacteria: Firmicutes, Actinomycetes* (pp. 257–273). New York, NY: Springer New York. [http://doi.org/10.1007/0-387-30743-5\\_13](http://doi.org/10.1007/0-387-30743-5_13)
- Whitmore, L., & Wallace, B. A. (2004). DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Research*, 32(suppl\_2), W668–W673. <http://doi.org/10.1093/nar/gkh371>



- Whitmore, L., & Wallace, B. A. (2008). Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases. *Biopolymers*, 89(5), 392–400. <http://doi.org/10.1002/bip.20853>
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. Recuperado a partir de <https://ggplot2.tidyverse.org>
- Willard, L., Ranjan, A., Zhang, H., Monzavi, H., Boyko, R. F., Sykes, B. D., & Wishart, D. S. (2003). VADAR: a web server for quantitative evaluation of protein structure quality. *Nucleic Acids Research*, 31(13), 3316–3319. <http://doi.org/10.1093/nar/gkg565>
- Williams, P. D., Pollock, D. D., & Goldstein, R. A. (2006). Functionality and the Evolution of Marginal Stability in Proteins: Inferences from Lattice Simulations. *Evolutionary Bioinformatics*, 2, 117693430600200020. <http://doi.org/10.1177/117693430600200013>
- Yu, J. P., Ladapo, J., & Whitman, W. B. (1994). Pathway of glycogen metabolism in *Methanococcus maripaludis*. *Journal of Bacteriology*, 176(2), 325–332. <http://doi.org/10.1128/jb.176.2.325-332.1994>
- Závodszky, P., Kardos, J., Svingor, & Petsko, G. A. (1998). Adjustment of conformational flexibility is a key event in the thermal adaptation of proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 95(13), 7406–7411. <http://doi.org/10.1073/pnas.95.13.7406>
- Zhang, J. (2012). Protein-Protein Interactions in Salt Solutions. En W. Cai (Ed.), *Protein-Protein Interactions - Computational and Experimental Tools* (pp. 359–376). InTech. <http://doi.org/10.5772/38056>

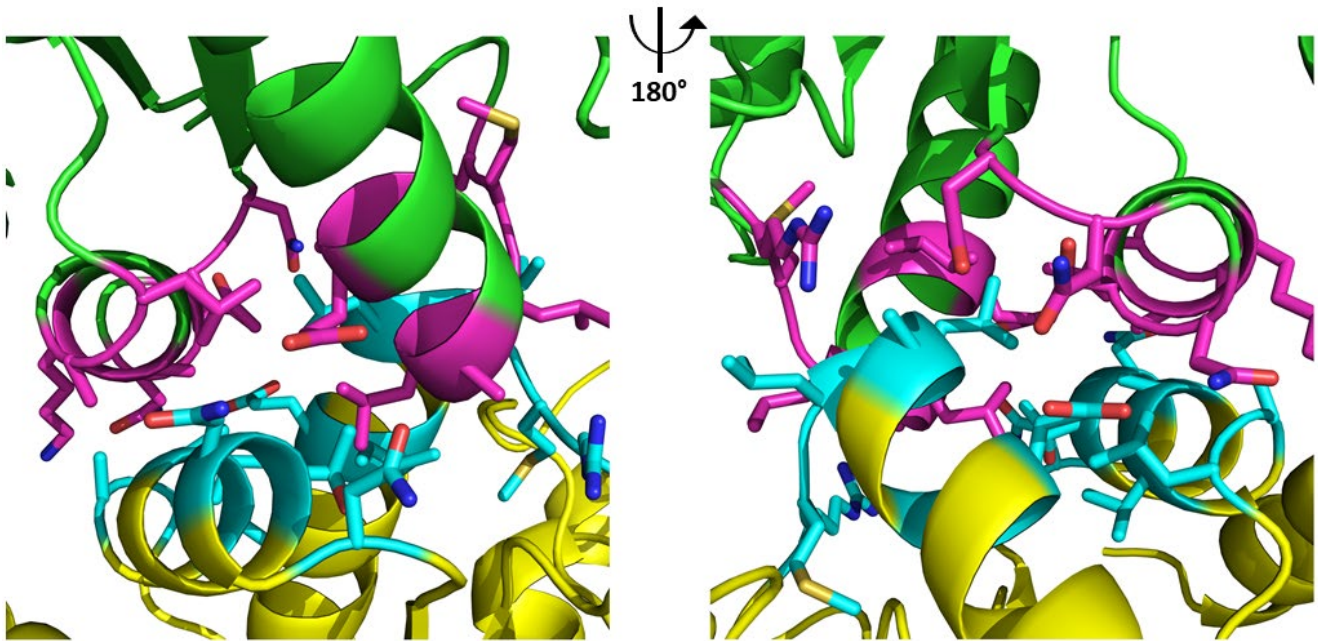
## MATERIAL SUPLEMENTARIO

### A. Detalles de la interfase de tetramerización de MjaTIM-85

**Tabla S1. Lista de pares de residuos que interactúan en la superficie del tetrámero.** Adaptado de Gayathri *et al.* 2007.

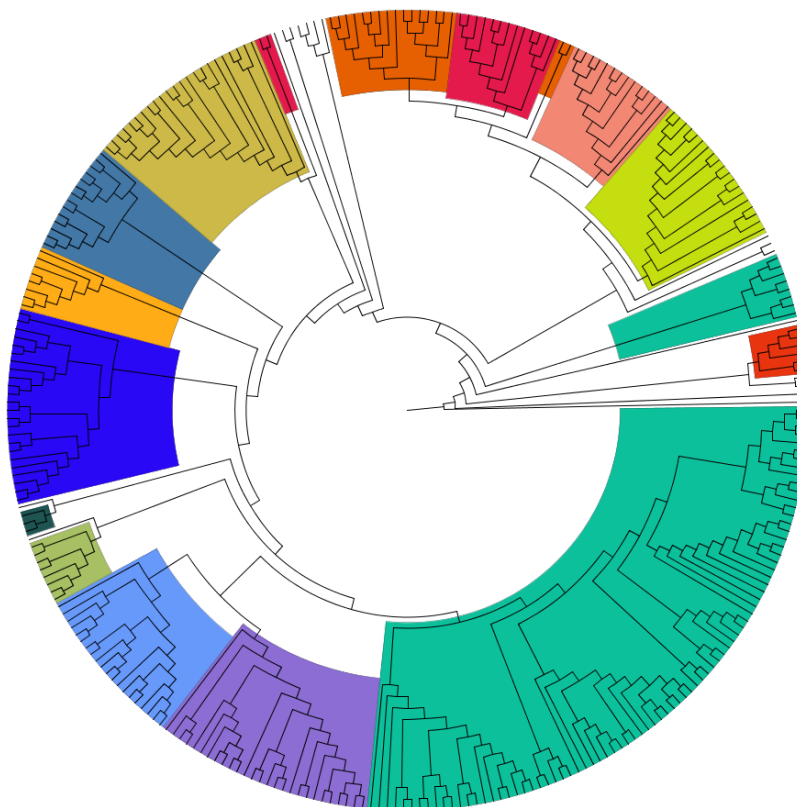
Subunidad A	Subunidad C	Tipo de interacción
Arg94	Ala98	Van der Waals
Met95	Leu96 Leu97	Van der Waals Puente de hidrógeno
Leu96	Arg94 Met95 Leu96	Van der Waals Van der Waals Van der Waals
Leu97	Met95 Ile100 Lys125	Puente de hidrógeno Van der Waals Van der Waals
Ala98	Arg94 Asn119	Van der Waals Van der Waals
Ile100	Leu97	Van der Waals
Glu101	Asn120 Asn122 Thr123	Van der Waals Puente de hidrógeno Puente de hidrógeno
Asn120	Glu101	Puente de hidrógeno
Asn122	Glu101	Puente de hidrógeno
Thr123	Glu101	Van der Waals
Ala126	Ala129 Leu130	Van der Waals Van der Waals
Val127	Leu97	Van der Waals
Ala129	Ala126	Van der Waals
Leu130	Ala126	Van der Waals

En la Tabla S1 y Figura S1 se observa el detalle de los residuos involucrados en la tetramerización de MjaTIM-85. Los datos provienen de Gayathri *et al.* (2007) donde se hizo el análisis de las interfases con el servidor web PISA (Krissinel & Henrick, 2007, <https://www.ebi.ac.uk/pdbe/pisa/>). Se repitió el análisis con el servidor para corroborar los datos y el programa no identificó otras interacciones además de los puentes de hidrógeno e interacciones de van der Waals ya descritos.



**Figura S1. Acercamiento a la superficie de tetramerización de MjaTIM-85.** Se indican los aminoácidos involucrados en la superficie del tetrámero de MjaTIM-85. En cian los residuos de la subunidad A y en magenta los residuos de la subunidad C (ver Tabla S1 para la lista completa). La asociación entre subunidades está dada por puentes de hidrógeno y enlaces de van der Waals (Gayathri et al. 2007, código PDB: 2H6R).

## B. Análisis filogenético de las TIM del dominio Archaea y selección de especies



**Figura S2. Árbol filogenético de máxima verosimilitud de la TIM de arqueas.** Para facilitar la visualización se eliminaron los nombres de las secuencias y se igualaron todas las longitudes de rama. Los colores indican los órdenes de arqueas según se especifica en la Tabla S2.

A partir de la base de datos compilada en el trabajo de licenciatura (Muñiz Trejo, 2017), se amplió el número de secuencias curadas manualmente del NCBI de 242 a 305. El alineamiento de las secuencias se realizó en el servidor PROMALS3D, el cual lo construye utilizando la información de la estructura tridimensional de archivos de coordenadas (Pei, Kim, & Grishin, 2008). Las estructuras de las TIM de *Methanocaldococcus jannaschii* (código PDB: 2H6R), *Pyrococcus woesei* (código PDB: 1HG3), *Thermoproteus tenax* (código PDB: 1W0M) y *Thermoplasma acidophilum* (código PDB: 5CSR) depositadas en el Protein Data Bank fueron provistas al servidor. El alineamiento resultante, por su extensión, se anexó al final de este documento. Dicho alineamiento se construyó utilizando el servidor ESPript 3.0 (Robert & Gouet, 2014), indicando la estructura secundaria a partir de la estructura tridimensional de la TIM de *Methanocaldococcus jannaschii*, además de una secuencia con un valor consenso de 0.7. Los residuos en casillas rojas indican una identidad completa para dicho residuo, los residuos con letra roja y en cajas azules indican aquellos que tienen una identidad mayor o igual a 0.7 para esa posición en el alineamiento, y los residuos en casillas verdes corresponden en posición a los residuos identificados en la superficie de tetramerización de MjaTIM.

Para construir el árbol filogenético se determinó primero el modelo de sustitución que mejor ajustara a los datos utilizando el programa PartitionFinderProtein 2 (Lanfear, Frandsen, Wright, Senfeld, & Calcott, 2016) con un

algoritmo de búsqueda codiciosa (Lanfear, Calcott, Ho, & Guindon, 2012) y búsqueda de esquemas de partición con la versión 8 de RAxML (Stamatakis, 2014). El mejor modelo de sustitución fue la matriz de sustitución mezclada LG4M+G (Le, Dang, & Gascuel, 2012) y no se encontró un esquema de partición para los datos. Con esta información se hizo la inferencia filogenética con la versión 8 de RAxML (Stamatakis, 2014) a través del servidor CIPRES (Miller, Pfeiffer, & Schwartz, 2010) y se obtuvo el árbol de máxima verosimilitud encontrado heurísticamente (**Figura S2**). El árbol se editó con el programa FigTree v 1.4.3 (<http://tree.bio.ed.ac.uk/software/figtree/>).

Durante el trabajo de licenciatura (Muñiz Trejo, 2017) se analizaron distintas secuencias de bacterias y eucariontes para enraizar el árbol, pero no se observaron cambios importantes en la topología del árbol de la TIM de arqueas, por lo que el árbol presentado aquí se enraizó utilizando únicamente como grupo externo a la secuencia de la TIM de *Saccharomyces cerevisiae* para facilitar la visualización. Como se encontró anteriormente, la filogenia de las TIM de arqueas parece agrupar de manera muy consistente a las arqueas a nivel de orden filogenético, aunque los Halobacteriales y los Thermoplasmatales no se agrupan todos de manera filogenética (el código de colores utilizado se presenta en la Tabla S2).

**Tabla S2. Código de colores del árbol filogenético de la TIM de arqueas**

Orden de Archaea	Color hexadecimal
<b>Phylum Crenarchaeota</b>	
Desulfurococcales	#E65F00
Sulfolobales	#F28773
Thermoproteales	#E51A4C
<b>Phylum Euryarchaeota</b>	
Archaeoglobales	#FFAC16
Halobacteriales	#0DC09C
Methanobacteriales	#2908F5
Methanocellales	#1E5553
<b>Methanococcales</b>	#4277A6
Methanomassilococcales	#A7C063
Methanomicrobiales	#6799FB
Methanosarcinales	#8C6DD3
Thermococcales	#CCB948
Thermoplasmatales	#E9350F
<b>Phylum Thaumarchaeota</b>	
Nitrosopumilales	#C4DE0F

**Tabla S3. Resumen de características de las especies seleccionadas y secuencias solicitadas**

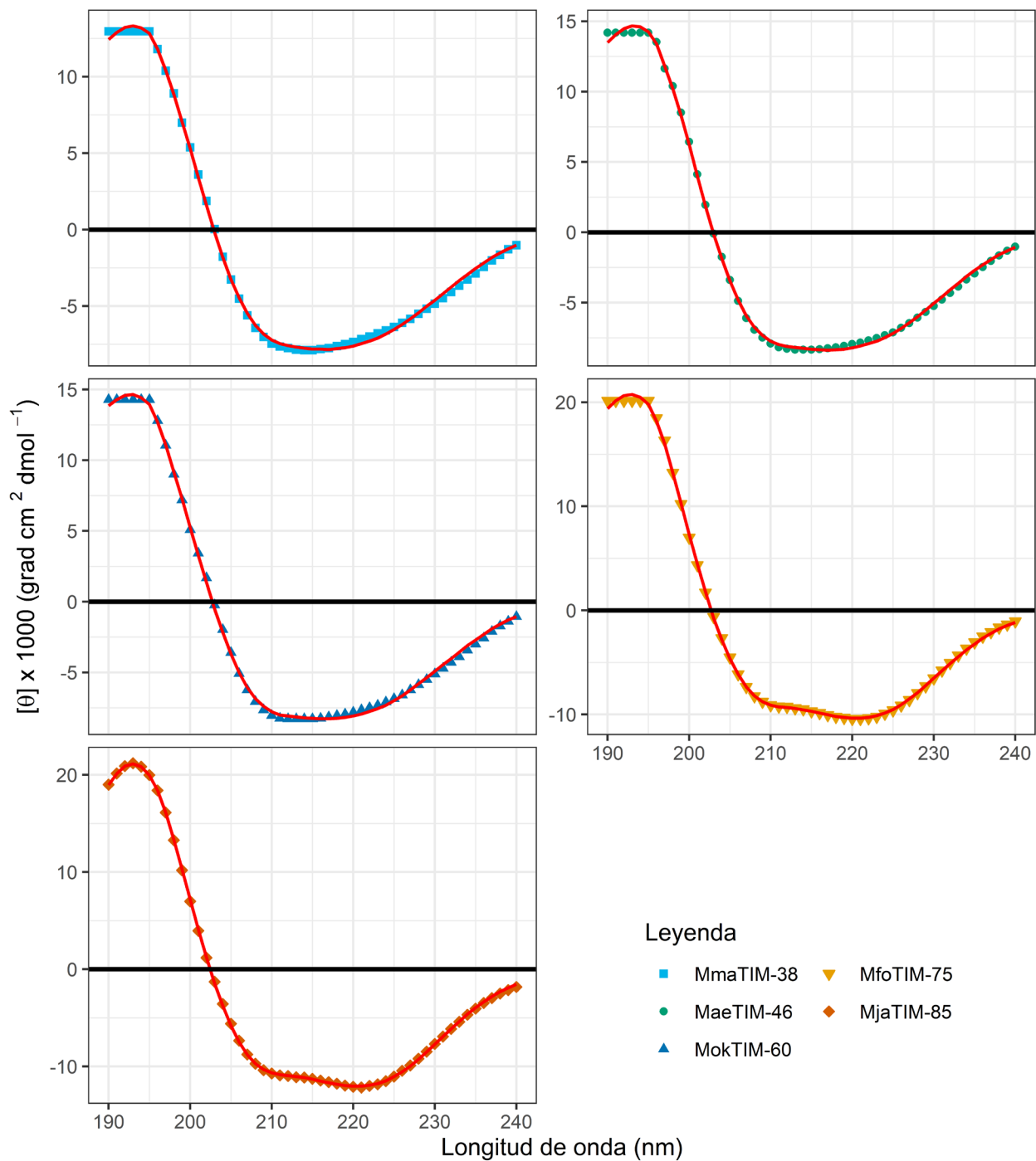
Especie	Abreviatura y código de color	Lugar de recolección	Características celulares	Intervalo de temperatura de crecimiento (°C) (óptimo)	Intervalo de pH de crecimiento (óptimo)	Referencia
<i>Methanocaldococcus jannaschii</i> DSM 2661	MjaTIM-85 #D55E00	Fumarola hidrotermal blanca en la dorsal del Pacífico Oriental.	Cocos irregulares móviles flagelados. Diámetro de hasta 1.5 µm.	80 – 85 (85)	5.2 – 7.0 (6.0)	Jones, Leigh, Mayer, Woese, & Wolfe, 1983
<i>Methanotorrus formicicus</i> Mc-S-70	MfoTIM-75 #E69F00	Fumarola hidrotermal negra en el campo Karei en la dorsal del Océano Índico.	Cocos irregulares. Diámetro de 0.8 a 1.5 µm.	55 – 83 (75)	6.0 – 8.5 (6 – 7)	Takai, Nealson, & Horikoshi, 2004
<i>Methanothermococcus okinawensis</i> IH1	MokTIM-60 #0072B2	Ventila hidrotermal profunda del Iheya Ridge, fosa de Okinawa, Japón.	Cocos irregulares móviles con un diámetro de 1.0 a 1.5 µm.	40 – 75 (60 – 65)	4.5 – 8.5 (6 – 7)	Horikoshi, Takai, & Inoue, 2002
<i>Methanococcus aeolicus</i> Nankai-3	MaeTIM-46 #009E73	Sedimentos marinos de la fosa Nankai, Japón.	Cocos de 1.5 a 2 µm de diámetro.	17 – 55 (46)	4.3 – 7.5	Kendall <i>et al.</i> , 2006
<i>Methanococcus maripaludis</i> C7	MmaTIM-38 #00B4E9	Sedimentos de marismas cerca de la Pawley, Carolina del Sur, Estados Unidos.	Cocos pleomórficos de diámetro 1.2 a 1.66 µm.	18 – 47 (38)	6.8 – 7.2	Jones, Paynter, & Gupta, 1983

```

>Methanocaldococcus_jannaschii
MLIVINYKTYNESIGNRGLIEAKIAEKVSEESGITIGVAPQFVDLRMIVENVNIPVYAQHDNINPGSHTGHILAEAIKDCGCKGTLINHSEKRMLLADIEAVINKCKNLGLETIVCTNNINTSK
AVAALSPDYIAVEPPELIGTGIPVSKANPEVVEGTVRVAVKEINKDKVLCGAGISKGEDVKAALDLGAEGVLLASGVVKAKNVEEAIRELIKFI
>Methanotorrus_formicicus
MKPVIIINYKTYMESIGKRGLEIAKIAEKVSEESGITIGVAPQFVDLRMITESVNIIPVYAQHIDAI SPGSHTGHILAESIKDCGVKGTLLNHSERNILLSIDDKVIEIAKNLGLLETIVCTNNINIV
SKAVALNPNYIAIEPPELIGTGIPVSKANPDIVEGTVKEVRKINKDKVILCGAGISTGSDVKAALDLGAGVLLASGVVKAKDIEKAIKIRDLISEI
>Methanothermococcus_okinawensis
MIKPIIIINYKYAESVGEKGLKIAKAAEKVSEESGISIGVCPQFLDLRMISENVNIPVYAQHFDVAVSPGSHTGHVLAETLKDCLNGSLLNHSEKRMLADLEKSIEIAKNYNLETIVCTNNIG
VSKAVALNPNMIAIEPPELIGTGIPVSKANPEVVEGTVKEVRGINKDKVILCGAGISKGEDVSSALELGACVLLASGVVKSVDVEGSIQELIKHI
>Methanococcus_aeolicus
MVI IINSPIIIINYKYSESVGKGLNIAKACEKISEESGISIGVAPQYLDLKTVDNVIIPVYAQHFDVAVGAGSNTGYVFAETIADCGLNGSLLNHSEKRMVLADLEEA INLANKYNFESVVC T
NNIVVSKAVVMNPTMLAIEPPELIGSGIPVSKANPEIVENTVNEVRKLNKTVTILCGAGISKGEDVSSAMELGADGVLLASGVVKAKNVEESIRELLRNI
>Methanococcus_maripaludis
MIINYKTYTESIGKSGLVIKAAEKVSDSESGVTIGVAPQFLDLRMILEEVNIIPVYAQHMDAVNPGSSTGHILPEVLKDGAVTGTLLNHSEKRMLLADLEKIEISIKLGLKTVCTNNITVSKAV
SALNPTSIAVEPPELIGSGMPVSKANPEVVSGTVREVREINKNVEILCGAGISKGEDVKSALDELGTNGVLLASGVVKAKCVEDAIRELIAEI

```

### C. Caracterización de la estructura secundaria

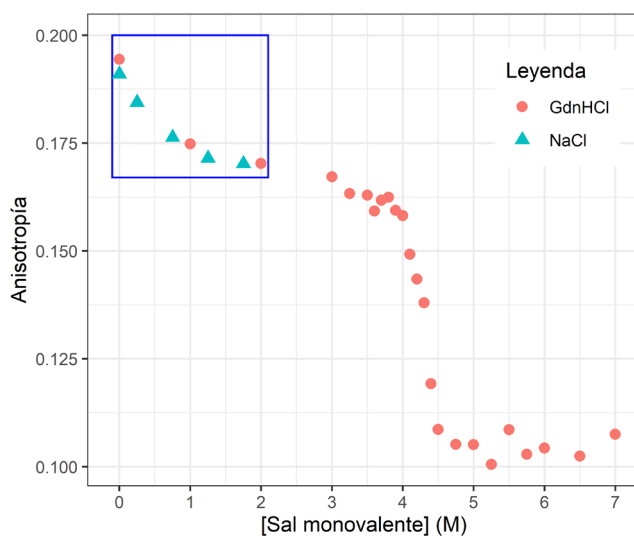


**Figura S3. Espectros de deconvolución de las ArqTIM.** Los datos de elipticidad fueron normalizados a elipticidad promedio por residuo ( $[\theta]$ ) y fueron suavizados con el método de Savitzky-Golay (1964) del programa Origin (OriginLab, s/f-c). El análisis de deconvolución se llevó a cabo con el programa CDSSTR (Sreerama et al., 2000; Sreerama & Woody, 2000) a través del servidor en línea DichroWeb (Whitmore & Wallace, 2004, 2008; <http://dichroweb.cryst.bbk.ac.uk/html/home.shtml>). Las líneas continuas representan el ajuste del análisis de deconvolución.

D. Desnaturalización térmica y química

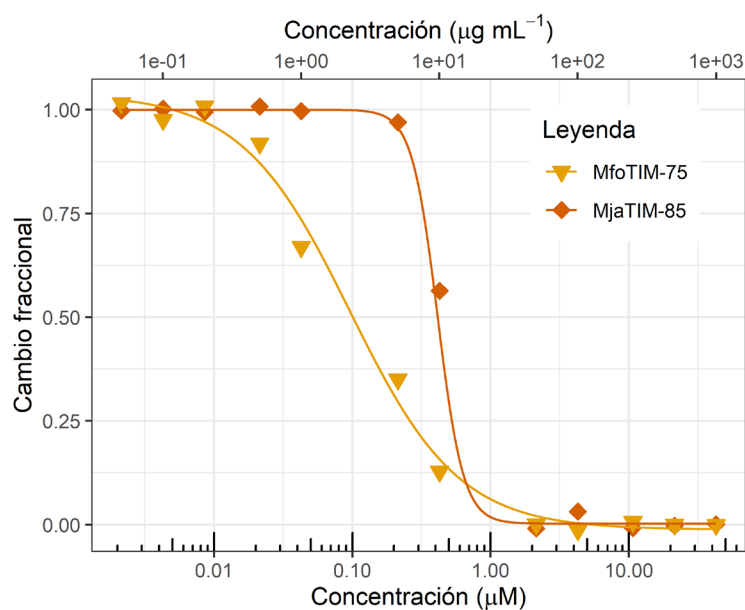
**Tabla S4. Resultados de las diferentes aproximaciones para el cálculo de la energía de activación en  $\text{kJ mol}^{-1}$ .**

Método	MfoTIM-75	MjaTIM-85
Promedio de los ajustes individuales a un modelo de dos estados irreversible	$574 \pm 9$	$895 \pm 35$
Gráfico de Arrhenius	$574 \pm 9$	$895 \pm 35$
Variación de la $T_m$ con respecto a la velocidad de barrido	$570 \pm 1$	$632 \pm 1$
Ancho del pico a la mitad de la altura de la traza calorimétrica (Ancho promedio $^{\circ}\text{C}$ )	$582 \pm 28$ (4.67)	$895 \pm 92$ (3.71)
Promedio de los métodos	$575 \pm 5$	$830 \pm 132$



**Figura S4. Efecto de los cationes monovalentes sobre la anisotropía de las tirosinas.** En el recuadro azul se observa que el cloruro de sodio tiene un efecto sobre la señal de anisotropía en el intervalo de 0 a 2 M al igual que el clorhidrato de guanidinio.





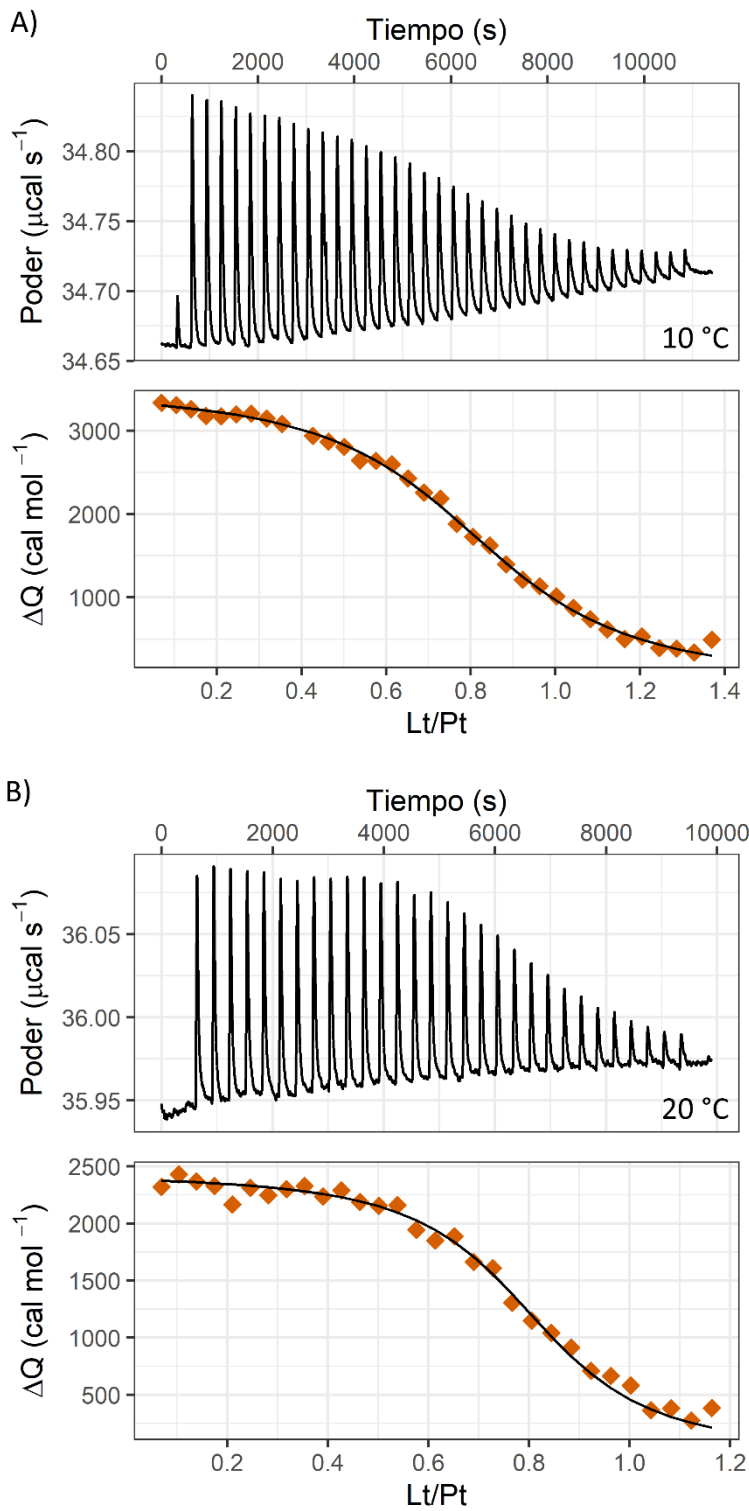
**Figura S5. Estabilidad a la dilución de las TermoTIM.** Se grafica la actividad catalítica normalizada con respecto a la concentración de las proteínas después de incubar por 48 horas a 25°C. Las medidas de actividad se realizaron con 5 ng de proteína finales. Los ajustes corresponden a una curva logística.

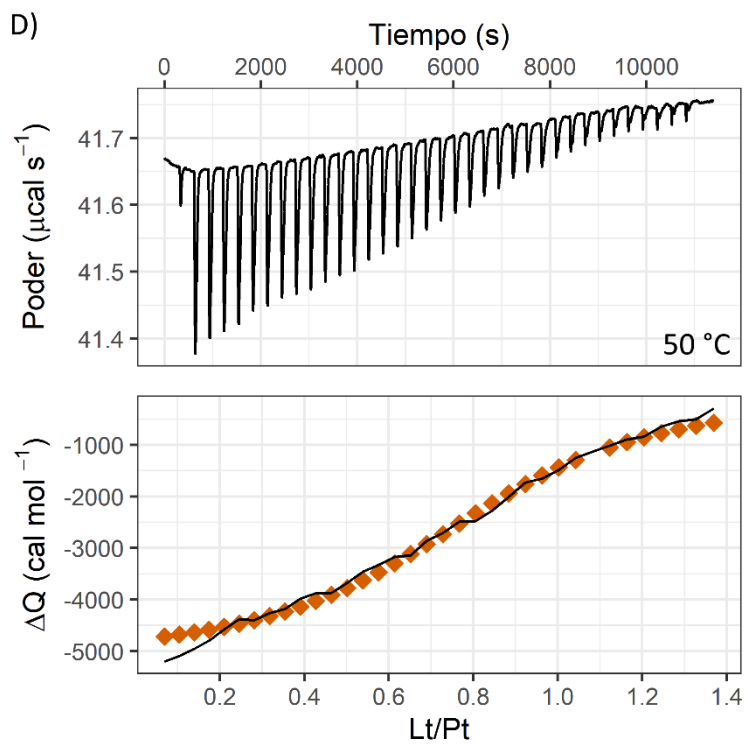
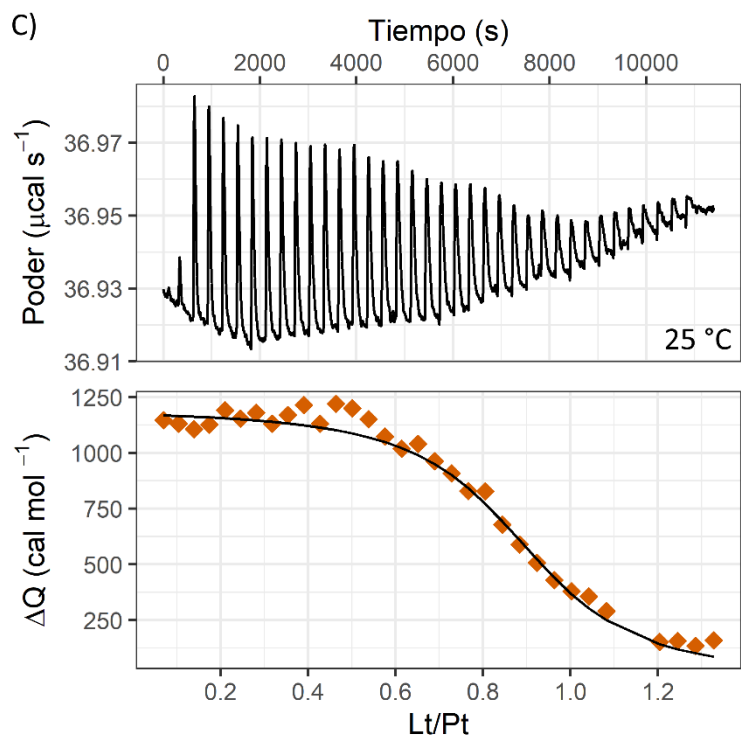
**Tabla S5. Valores de la constante de disociación ( $K_D$ ) observados para la dilución de las TermoTIM.**

TIM	$K_D$ ( $\mu\text{M}$ )	$K_D$ ( $\mu\text{g mL}^{-1}$ )
MfoTIM-75	0.0965	2.27
MjaTIM-85	0.4525	10.55

Para estos análisis es importante reiterar que se desconoce, hasta este punto, si la disociación de los tetrámeros por dilución concluye con monómeros, dímeros o pasa por ambos (esto que podría ocurrir si tanto los dímeros como los monómeros son inactivos, pero no se observaría en la gráfica), por lo que no se ofrece un análisis más detallado sobre la estabilidad a la dilución.

E. Caracterización de la unión a ligando





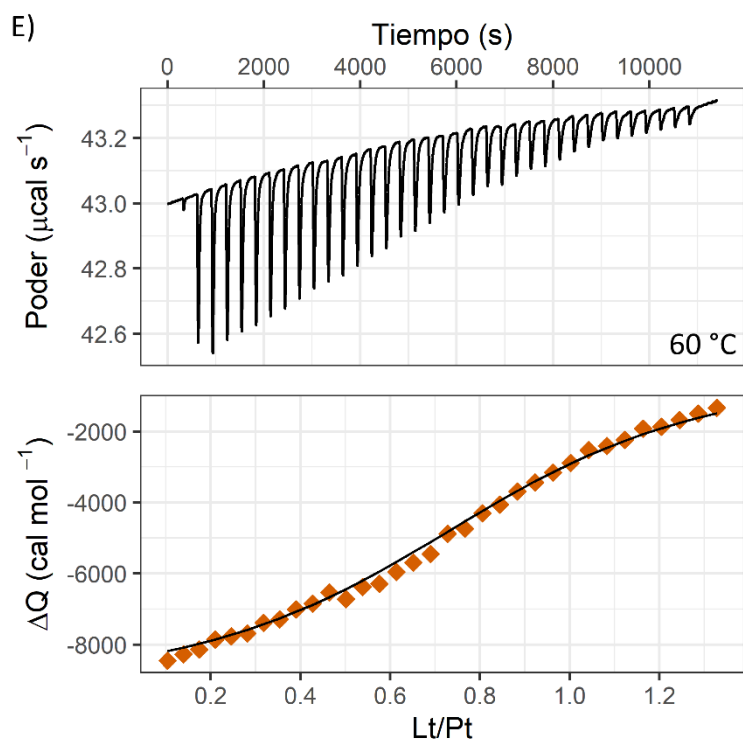
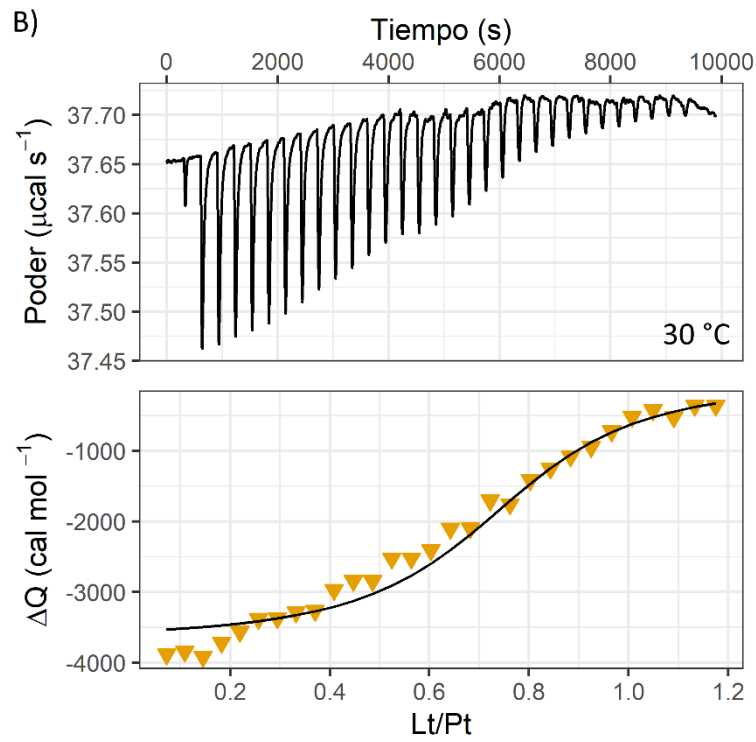
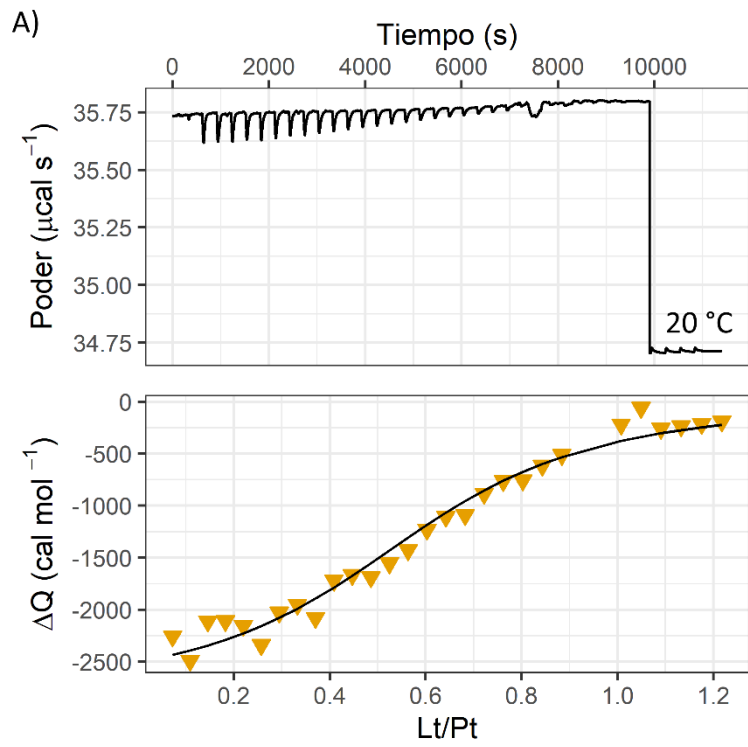


Figura S6. Isotermas de unión de MjaTIM-85 con PGH a varias temperaturas.

Tabla S6. Parámetros de las isotermas de unión de MjaTIM-85 bajo un modelo de unión 1:1

Temperatura (°C)	15	20	25	50	60
n	0.847	0.807	0.897	0.832	0.868
$K_A$ ( $\text{M}^{-1}$ )	$5.83 \times 10^4$	$1.65 \times 10^5$	$1.29 \times 10^5$	$3.38 \times 10^4$	$2.17 \times 10^4$
$K_D$ (M)	$1.72 \times 10^{-5}$	$6.08 \times 10^{-6}$	$7.73 \times 10^{-6}$	$2.96 \times 10^{-5}$	$4.61 \times 10^{-5}$
$\Delta G$ ( $\text{cal mol}^{-1}$ )	-6283.17	-6996.95	-6973.90	-6695.73	-6609.61
$\Delta H$ ( $\text{cal mol}^{-1}$ )	3478.10	2432.8	1195.6	-5162.9	-9424.9
$-T\Delta S$ ( $\text{cal mol}^{-1}$ )	-9761.27	-9429.75	-8169.50	-1532.83	2815.29
$\Delta S$ ( $\text{cal mol}^{-1} \text{K}^{-1}$ )	33.88	32.17	27.40	4.74	-8.45
Valor c	25	71	56	15	9



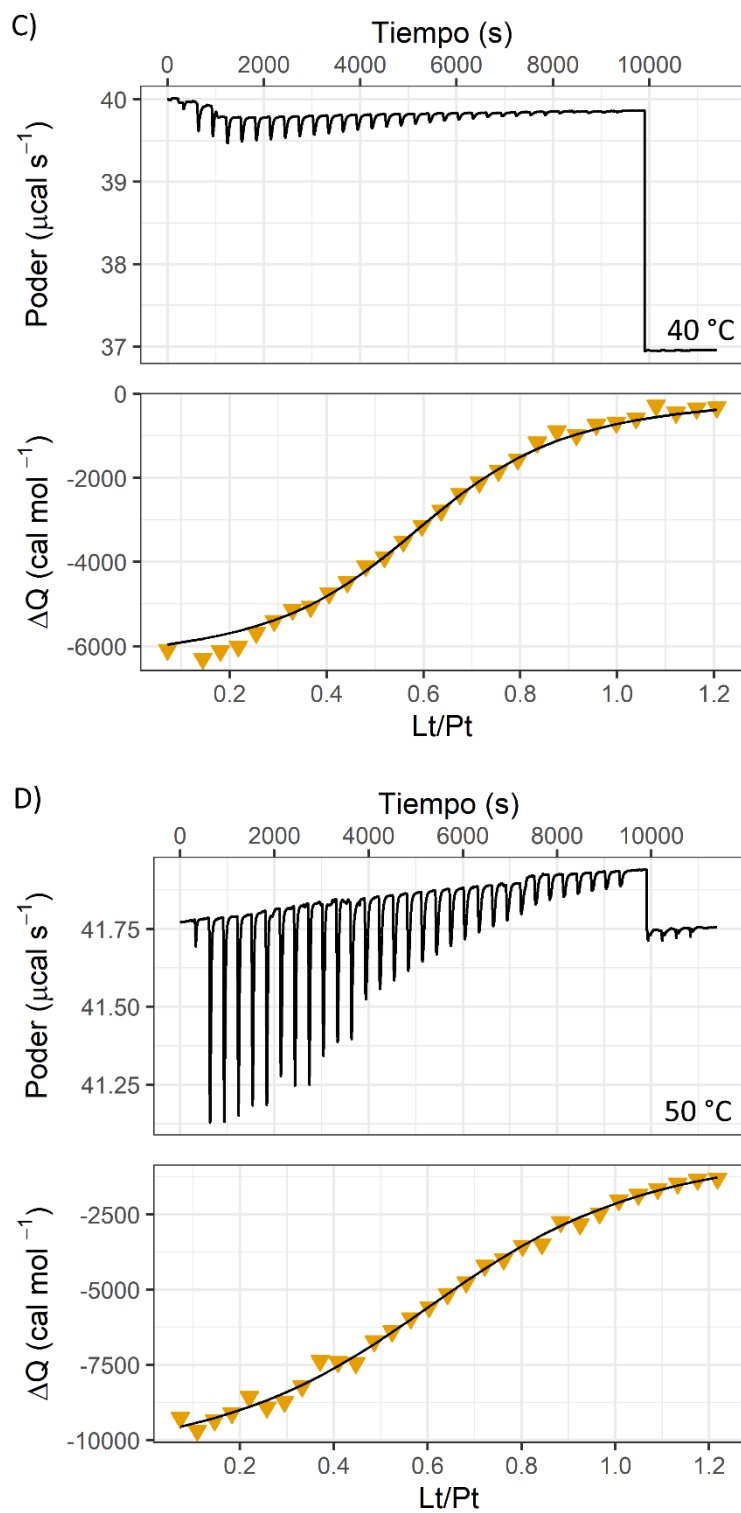
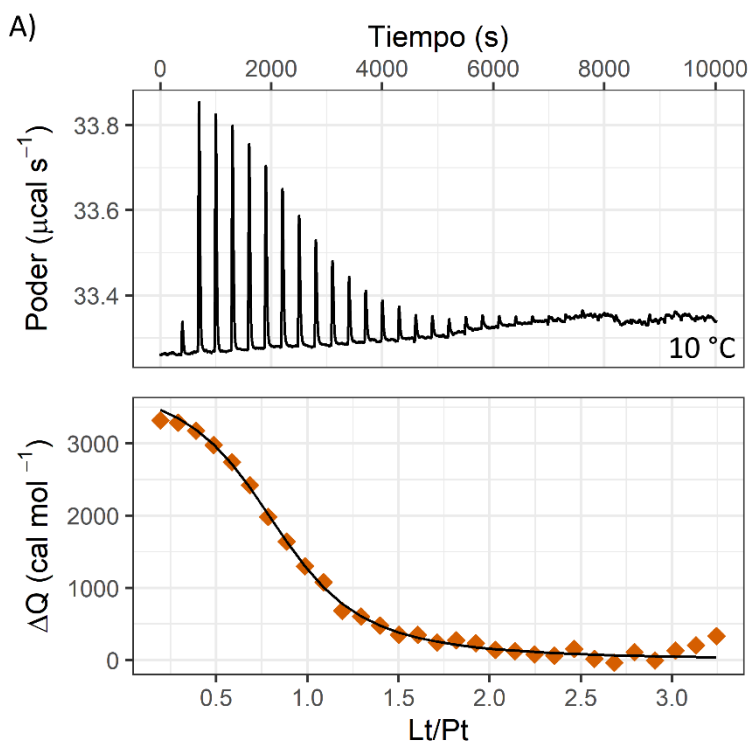
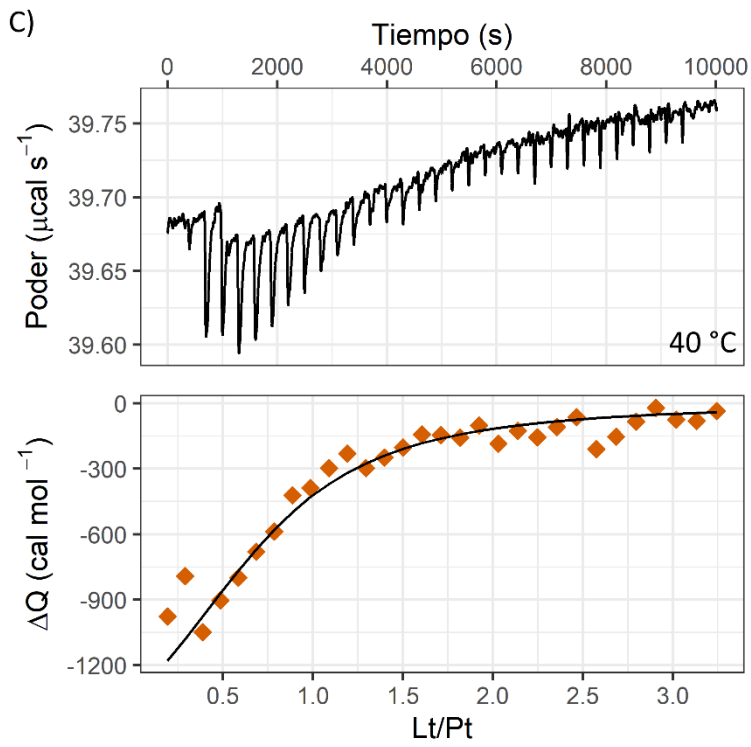
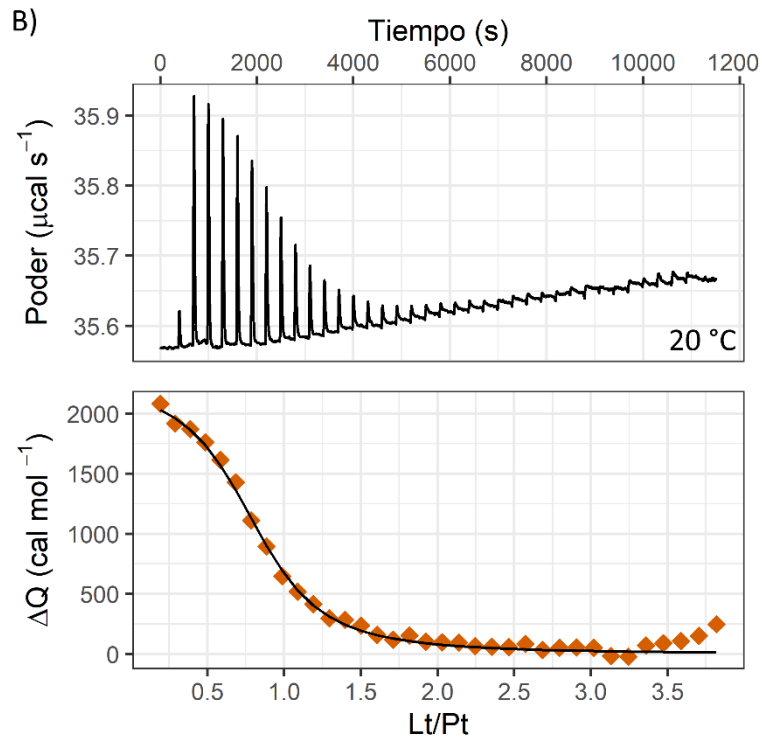


Figura S7. Isotermas de unión de MfoTIM-75 con PGH a varias temperaturas.

Tabla S7. Parámetros de las isotermas de unión de MfoTIM-75 bajo un modelo de unión 1:1

Temperatura (°C)	10	20	40	50
n	0.609	0.757	0.622	0.702
$K_A$ (M <sup>-1</sup> )	$3.14 \times 10^4$	$8.91 \times 10^4$	$5.80 \times 10^4$	$2.80 \times 10^4$
$K_D$ (M)	$3.19 \times 10^{-5}$	$1.12 \times 10^{-5}$	$1.73 \times 10^{-5}$	$3.57 \times 10^{-5}$
$\Delta G$ (cal mol <sup>-1</sup> )	-5825.58	-6639.65	-6824.93	-6576.67
$\Delta H$ (cal mol <sup>-1</sup> )	-2804.80	-3679.7	-6447.7	-10933
-T $\Delta S$ (cal mol <sup>-1</sup> )	-3020.78	-2959.95	-377.23	4356.33
$\Delta S$ (cal mol <sup>-1</sup> K <sup>-1</sup> )	-10.67	-10.10	-1.20	13.48
Valor c	13	37	24	12







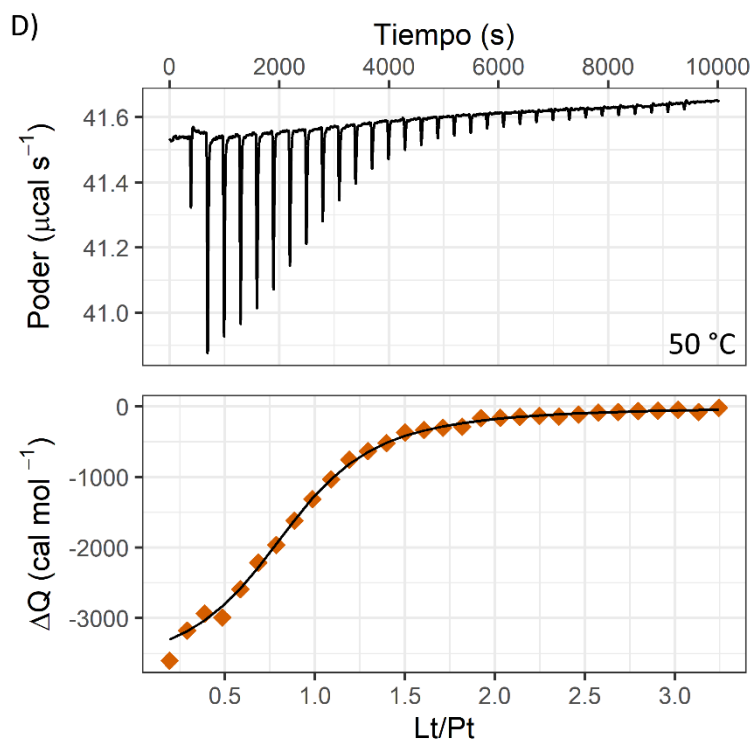


Figura S8. Isothermas de unión de MjaTIM-85 con PGH a varias temperaturas en 0.66 molal de sacarosa.

Tabla S8. Parámetros de las isothermas de unión de MfoTIM-75 con sacarosa bajo un modelo de unión 1:1

Temperatura (°C)	10	20	40	50
<b>n</b>	0.839	0.808	0.684	0.851
<b><math>K_A</math> (M<sup>-1</sup>)</b>	$4.12 \times 10^4$	$4.54 \times 10^4$	$1.13 \times 10^4$	$3.55 \times 10^4$
<b><math>K_D</math> (M)</b>	$2.43 \times 10^{-5}$	$2.20 \times 10^{-5}$	$8.86 \times 10^{-5}$	$2.81 \times 10^{-5}$
<b><math>\Delta G</math> (cal mol<sup>-1</sup>)</b>	-5979.03	-6246.52	-5807.04	-6728.95
<b><math>\Delta H</math> (cal mol<sup>-1</sup>)</b>	3926.4	2286.4	-1847.8	-3790
<b><math>-T\Delta S</math> (cal mol<sup>-1</sup>)</b>	-9905.43	-8532.92	-3959.24	-2938.95
<b><math>\Delta S</math> (cal mol<sup>-1</sup> K<sup>-1</sup>)</b>	34.98	29.11	12.64	9.09
<b>Valor c</b>	13	14	4	11



Thermoplasma volcanium . . . . . MFT . . . . . VLVNKKVAAE . ANGK . . . . . NVEVVFVSKL . PKYDS . . . . . VRLIIAIPAILDLHNA . . . . . HEFR . . . . . NVEFFSQHVDV . . . . . VG . . . . . YG  
Picrophilus torridus . . . . . MKPL . . . . . IFINIKHVEN . ATGR . . . . . NAEELLNFRKLIKQ . . . . . KNVYCYLGP . . . . . GDLLRS . . . . . LLEFP . . . . . DNLNYSQHVDV . . . . . NG . . . . . YG  
Ferroplasma acidamarum . . . . . MEPE . . . . . IFINIKHVN . AVGT . . . . . NSEKLLITDFPKSIKKQ . . . . . NGLYICMAETDLY . . . . . LQ . . . . . KSPF . . . . . EIPVYAGHVDV . . . . . SN . . . . . YG  
Korarchaeum cryptofolium . . . . . MHIFR . . . . . LITIKHME . SSGD . . . . . RAVEISKSARELWRE . . . . . MGLVIGVSNV . . . . . VRLNLSSEFV . . . . . EIPVYAGHVDV . . . . . PT . . . . . AG  
Halobacterium volcanensis . . . . . MKLINP . . . . . LITIKHME . SSGD . . . . . RAVEISKSARELWRE . . . . . MGLVIGVSNV . . . . . VRLNLSSEFV . . . . . EIPVYAGHVDV . . . . . PT . . . . . AG  
Nitrosopumilus maritimus . . . . . M . . . . . FVINIKNVEE . ISGE . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Ca\_Nitrosopumilus\_koreans . . . . . M . . . . . FVINIKNVEE . ISGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Ca\_Nitrososphaeram\_limnia . . . . . M . . . . . FLINIKNVEE . IAGS . . . . . KITPFWKIVKISKK . . . . . YKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Ca\_Nitrosopumilus\_salaria . . . . . M . . . . . FVINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Cenarchaeum synulosum . . . . . MARSV . . . . . FLINIKNVEE . AAGG . . . . . RIDLAAAAGAAAK . . . . . GVRILADLAP . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Ca\_Methanomethylphilus\_a . . . . . M . . . . . FLINIKNVEE . AAGG . . . . . RIDLAAAAGAAAK . . . . . GVRILADLAP . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Ca\_Parvarchaeum\_acidophil . . . . . M . . . . . FLINIKNVEE . APT . . . . . KGIEIAKYSGEISKS . . . . . SGVSVVSPFT . . . . . LKETVAF . . . . . AKTIAQGVNDV . . . . . D . . . . . PG  
Ca\_Parvarchaeum\_acidophil . . . . . MGNKI . . . . . FLINIKNVEE . AFS . . . . . KGIDIAKYAAQV . . . . . SAS . . . . . KNSVIVSPFT . . . . . LKETVAF . . . . . AKTIAQGVNDV . . . . . D . . . . . PG  
Ca\_Caldiarchaeum\_subterra . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Thermofilum carboxyditrop . . . . . M . . . . . FLINIKNVEE . ASGG . . . . . KALQAKVAEKLSKE . . . . . YGVVIAVAP . . . . . QADLRVAQT . . . . . GAETWAQHVDP . . . . . NG . . . . . HG  
Halanaerobacterium\_sulfuri . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Haloflex gibbosii . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Haloterrigena jeotgali . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Halolamina rubra . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Halobellium . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Halopiger salifodinae . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Halapricum salinum . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Halococcus sediminicola . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Thermococcus eurythermalis . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Halorubrum halophilum . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Halorubrum ezemouense . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Haloplanus natans . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Palaeococcus ferrophilus . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Methanococcus pyrrhari . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Thermogladus cellulolyticus . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Methanosarcina thermophila . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Methanosarcina siciliensis . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Geoglobus acetivorans . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Geoglobus ahangari . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Ca\_Nitrosopumilus\_brevi . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Ca\_Acidianus\_cophuensis . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Ca\_Methanoplasma\_terminum . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Methanosarcina lacustris . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Methanosarcina\_hacorobens . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Methanosarcina\_vacuolata . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Thermococcus parvulus . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Methanococcus shengliensis . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Methanobrevibacter\_oralis . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Thermococcus\_nautili . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Methanococcus\_pallidum . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Methanococcus\_thermautotrophicus . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Methanobrevibacter\_ruminantium . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Desulfurococcus\_amoxylophilus . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Ca\_Nitrososphaera\_everglis . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Ca\_Halobacterium\_tyrrellensis . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Halorhabdus\_tiamatea . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Halomicrobium\_katesii . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Methanococcus\_cleefensis . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Natronomonas\_moolapensis . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Natronomonas\_pharaonis . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Thermococcus\_zilligii . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Halogeometricum\_pallidum . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Ca\_Nanopumilus\_sp . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Natronomonas\_pellulirubrum . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Halorhabdus\_utahensis . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Halovivax\_ruber . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Salinararchaeum\_sp . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Halorubrum\_arciis . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Halorubrum\_lipolyticum . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Halococcus\_litorum . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Halococcus\_thailandensis . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Halococcus\_saccharolyticus . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Halococcus\_hamelinensis . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Halobiforma\_lacisalis . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Halorubrum\_japonicum . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Halorubrum\_californicum . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Halorubrum\_argentinensis . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Halorubrum\_amolytica . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Halorubrum\_sinaiensis . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Halorubrum\_vallismortis . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Halorubrum\_coriense . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Halorubrum\_distributum . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Halosimplex\_carlsbadense . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Ca\_Nanopumilus\_acidilobi . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Ca\_Nitrosopumilus\_adriaticum . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Ca\_Nitrosotenuis\_cloacae . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Ca\_Nitrosotenuis\_devanar . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Halobacterium\_huberi . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Thermococcus\_guaymasensis . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Methanocaldococcus\_batho . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Pyrobaculum\_ferrireducens . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Thermococcus\_celericreans . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Thermococcus\_hirodudensis . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Methanocaldococcus\_jejirensis . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Acidiplasma\_cupricumulans . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Acidiplasma\_aeolicum . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Halolamina\_flavescens . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Methanococcus\_sediminis . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Ca\_Nitrososphaera\_gargensis . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Methanobrevibacter\_fillo . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Methanobrevibacter\_curvat . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Methanobrevibacter\_cuticum . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Ca\_Nitrosopumilus\_piranum . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Thermococcus\_chitonophagus . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Ca\_Thalassosphaera\_mari . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Methanobrevibacter\_olleya . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Methanogenium\_cariaci . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Thermococcus\_ptonophilum . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Methanobrevibacter\_milleri . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Pyrodicticum\_occultum . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Ca\_Nitrososphaera\_souiana . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Sulfolobus\_metallicus . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Methanobrevibacter\_arzori . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Halobellium\_limi . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Natronobacterium\_bangense . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Halorubrum\_rudnickae . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Ca\_Nitrososphaera\_gargensis . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Methanocaldococcus\_franklii . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Methanocaldococcus\_blatii . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Methanohalophilus\_euhalob . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Haloterrigena\_hispanica . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Methanocaldococcus\_taiwanensis . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Ca\_Methanosarcina\_sp . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Methanohalophilus\_sp . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Natronobacterium\_chitinivum . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Natronobacterium\_halalkali . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Eury\_archaeon\_TMED285 . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Eury\_archaeon\_TMED248 . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Eury\_archaeon\_TMED117 . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Ca\_Nitrososphaera\_arzori . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Nitrosopumilus\_arzori . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Methanobrevibacter\_gottsc . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Halorubrum\_sp . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Halosimplex\_sp . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Natronomonas\_sp . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Methanohalophilus\_sp\_RSK . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Thauma\_archaeon\_S13 . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Thauma\_archaeon\_S14 . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG  
Nitrosopumilus\_sp\_H13 . . . . . M . . . . . FLINIKNVEE . IAGD . . . . . KIIKFKVTAEKISKK . . . . . FKVKIATICPF . . . . . OHLIGVANS . . . . . SIPVYAGHVDV . . . . . DSK . . . . . VG

Nitrosopumilus\_sp\_H8 .....M.....FVINGKNV...E..ISGD...KMA...GFLKIVH...RVAKK...YKVKI...AVAP...QHLL...GVAAKS...FVIL...LAQHAD...CHM...FG  
 Halorubrum\_sp\_CBA1220 .....MF.....VLVNLKAV...PC.....DP...VAVAA...AARDV...AAE...SGVR...IAVAP...QAAD...LAAV...ADT...GVE...TWAQHV...SPND...YG  
 Halorubrum\_sp\_CBA1229 .....MF.....VLVNLKAV...PC.....DP...VAVAA...AARDV...AAE...SGVR...IAVAP...QAAD...LAAV...ADT...GVE...TWAQHV...SPND...YG  
 Methanohalophilus\_halophi .....MDSPL.....IVVNLKAV...PC.....DP...VAVAA...AARDV...AAE...SGVR...IAVAP...QAAD...LAAV...ADT...GVE...TWAQHV...SPND...YG  
 Methanohalophilus\_pottuca .....MDSPL.....IVVNLKAV...PC.....DP...VAVAA...AARDV...AAE...SGVR...IAVAP...QAAD...LAAV...ADT...GVE...TWAQHV...SPND...YG  
 Halalkalicoccus\_sp\_GSM28 .....MF.....VLVNLKAV...PC.....DP...VAVAA...AARDV...AAE...SGVR...IAVAP...QAAD...LAAV...ADT...GVE...TWAQHV...SPND...YG  
 Halorubrum\_sp\_ZY8 .....MF.....VLVNLKAV...PC.....DP...VAVAA...AARDV...AAE...SGVR...IAVAP...QAAD...LAAV...ADT...GVE...TWAQHV...SPND...YG  
 Ca\_Woesearchaeota\_archaeo .....M.....IVVNLKAV...PC.....DP...VAVAA...AARDV...AAE...SGVR...IAVAP...QAAD...LAAV...ADT...GVE...TWAQHV...SPND...YG  
 Haloplanus\_aerogenes .....MF.....VLVNLKAV...PC.....DP...VAVAA...AARDV...AAE...SGVR...IAVAP...QAAD...LAAV...ADT...GVE...TWAQHV...SPND...YG  
 Halostella\_sp\_LT12 .....MF.....VLVNLKAV...PC.....DP...VAVAA...AARDV...AAE...SGVR...IAVAP...QAAD...LAAV...ADT...GVE...TWAQHV...SPND...YG  
 Halobellus\_sp\_Atliit-31R .....MF.....VLVNLKAV...PC.....DP...VAVAA...AARDV...AAE...SGVR...IAVAP...QAAD...LAAV...ADT...GVE...TWAQHV...SPND...YG  
 Ca\_Korarchaeota\_archaeon .....MILPD.....IVVNLKAV...PC.....DP...VAVAA...AARDV...AAE...SGVR...IAVAP...QAAD...LAAV...ADT...GVE...TWAQHV...SPND...YG  
 Ca\_Thorarchaeota\_archaeon .....MVOIELPV.....VAINLKT...V...PQ...ATGE...RAV...LLA...OHCES...V...AKE...YGV...S...IVVAP...Q...T...D...L...Y...V...S...R...A...V...D...D...I...P...V...F...A...Q...H...V...D...P...V...G...A...P...G  
 Ca\_Thoiarchaeota\_archaeon .....MNRPV.....LGGN...K...M...Q...I...V...E...V...E...S...A...V...K...I...A...K...E...L...S...P...A...L...D...D...L...N...L...K...N...V...E...V...F...I...A...P...Q...T...D...L...Y...V...S...R...A...V...D...D...I...P...V...F...A...Q...H...V...D...P...V...G...A...P...G  
 Ca\_Heimdallarchaeota\_archaeon .....MKS...I...H...Y...P...L...I...V...N...L...K...A...V...P...C...D...P...V...A...V...A...A...A...R...D...V...A...A...E...S...G...V...R...I...A...V...A...P...Q...A...A...D...L...A...A...V...A...D...T...G...V...E...T...W...A...Q...H...V...S...P...N...D...Y...G  
 Ca\_Altiarchaeales\_archaeo .....MEDKI.....I...I...N...L...K...A...V...P...C...D...P...V...A...V...A...A...R...D...V...A...A...E...S...G...V...R...I...A...V...A...P...Q...A...A...D...L...A...A...V...A...D...T...G...V...E...T...W...A...Q...H...V...S...P...N...D...Y...G  
 Ca\_Aenigmarchaeota\_archaeo .....MKPT.....I...I...N...L...K...A...V...P...C...D...P...V...A...V...A...A...R...D...V...A...A...E...S...G...V...R...I...A...V...A...P...Q...A...A...D...L...A...A...V...A...D...T...G...V...E...T...W...A...Q...H...V...S...P...N...D...Y...G  
 Halonotius\_pteroides .....MF.....VLVNLKAV...PC.....DP...VAVAA...AARDV...AAE...SGVR...IAVAP...QAAD...LAAV...ADT...GVE...TWAQHV...SPND...YG  
 Halonotius\_sp\_F13-13 .....MF.....VLVNLKAV...PC.....DP...VAVAA...AARDV...AAE...SGVR...IAVAP...QAAD...LAAV...ADT...GVE...TWAQHV...SPND...YG  
 Halopiger\_aswanensis .....MF.....VLVNLKAV...PC.....DP...VAVAA...AARDV...AAE...SGVR...IAVAP...QAAD...LAAV...ADT...GVE...TWAQHV...SPND...YG  
 Ca\_Nanoclepta\_minutus .....M.....L...L...I...N...L...K...A...V...P...C...D...P...V...A...V...A...A...R...D...V...A...A...E...S...G...V...R...I...A...V...A...P...Q...A...A...D...L...A...A...V...A...D...T...G...V...E...T...W...A...Q...H...V...S...P...N...D...Y...G  
 Natronolimmobius\_sp\_AArc1 .....MF.....L...L...I...N...L...K...A...V...P...C...D...P...V...A...V...A...A...R...D...V...A...A...E...S...G...V...R...I...A...V...A...P...Q...A...A...D...L...A...A...V...A...D...T...G...V...E...T...W...A...Q...H...V...S...P...N...D...Y...G  
 Methanobrevibacter\_woesei .....MSKNSKPT.....V...I...L...N...L...K...A...V...P...C...D...P...V...A...V...A...A...R...D...V...A...A...E...S...G...V...R...I...A...V...A...P...Q...A...A...D...L...A...A...V...A...D...T...G...V...E...T...W...A...Q...H...V...S...P...N...D...Y...G  
 Methanothermobacter\_deflu .....MKDTP...I...V...L...N...L...K...A...V...P...C...D...P...V...A...V...A...A...R...D...V...A...A...E...S...G...V...R...I...A...V...A...P...Q...A...A...D...L...A...A...V...A...D...T...G...V...E...T...W...A...Q...H...V...S...P...N...D...Y...G  
 Halorussus\_rarus .....MK.....V...L...N...L...K...A...V...P...C...D...P...V...A...V...A...A...R...D...V...A...A...E...S...G...V...R...I...A...V...A...P...Q...A...A...D...L...A...A...V...A...D...T...G...V...E...T...W...A...Q...H...V...S...P...N...D...Y...G  
 Halopelagius\_longus .....MF.....L...L...I...N...L...K...A...V...P...C...D...P...V...A...V...A...A...R...D...V...A...A...E...S...G...V...R...I...A...V...A...P...Q...A...A...D...L...A...A...V...A...D...T...G...V...E...T...W...A...Q...H...V...S...P...N...D...Y...G  
 Haloplanus\_salinus .....MF.....VLVNLKAV...PC.....DP...VAVAA...AARDV...AAE...SGVR...IAVAP...QAAD...LAAV...ADT...GVE...TWAQHV...SPND...YG  
 Methanothermobacter\_teneb .....MKYDTP...I...V...L...N...L...K...A...V...P...C...D...P...V...A...V...A...A...R...D...V...A...A...E...S...G...V...R...I...A...V...A...P...Q...A...A...D...L...A...A...V...A...D...T...G...V...E...T...W...A...Q...H...V...S...P...N...D...Y...G  
 Acidianus\_brixleyi .....MKKPI...I...I...N...L...K...A...V...P...C...D...P...V...A...V...A...A...R...D...V...A...A...E...S...G...V...R...I...A...V...A...P...Q...A...A...D...L...A...A...V...A...D...T...G...V...E...T...W...A...Q...H...V...S...P...N...D...Y...G  
 Metallosphaera\_hakonensis .....MKKPI...I...I...N...L...K...A...V...P...C...D...P...V...A...V...A...A...R...D...V...A...A...E...S...G...V...R...I...A...V...A...P...Q...A...A...D...L...A...A...V...A...D...T...G...V...E...T...W...A...Q...H...V...S...P...N...D...Y...G  
 Acidianus\_sulfidivorans\_J .....MKKPI...I...I...N...L...K...A...V...P...C...D...P...V...A...V...A...A...R...D...V...A...A...E...S...G...V...R...I...A...V...A...P...Q...A...A...D...L...A...A...V...A...D...T...G...V...E...T...W...A...Q...H...V...S...P...N...D...Y...G  
 Methanospirillum\_stamsii .....MKPNSLNTPL...V...L...N...L...K...A...V...P...C...D...P...V...A...V...A...A...R...D...V...A...A...E...S...G...V...R...I...A...V...A...P...Q...A...A...D...L...A...A...V...A...D...T...G...V...E...T...W...A...Q...H...V...S...P...N...D...Y...G  
 Methanospirillum\_lacunae .....MAHHTIKSPV...V...L...N...L...K...A...V...P...C...D...P...V...A...V...A...A...R...D...V...A...A...E...S...G...V...R...I...A...V...A...P...Q...A...A...D...L...A...A...V...A...D...T...G...V...E...T...W...A...Q...H...V...S...P...N...D...Y...G  
 Methanospiraera\_cuniculi .....MADSPK...I...I...N...L...K...A...V...P...C...D...P...V...A...V...A...A...R...D...V...A...A...E...S...G...V...R...I...A...V...A...P...Q...A...A...D...L...A...A...V...A...D...T...G...V...E...T...W...A...Q...H...V...S...P...N...D...Y...G  
 Methanobrevibacter\_thauer .....MDTPI...V...L...N...L...K...A...V...P...C...D...P...V...A...V...A...A...R...D...V...A...A...E...S...G...V...R...I...A...V...A...P...Q...A...A...D...L...A...A...V...A...D...T...G...V...E...T...W...A...Q...H...V...S...P...N...D...Y...G  
 Ca\_Nanosidianus\_stetteri .....M.....L...L...I...N...L...K...A...V...P...C...D...P...V...A...V...A...A...R...D...V...A...A...E...S...G...V...R...I...A...V...A...P...Q...A...A...D...L...A...A...V...A...D...T...G...V...E...T...W...A...Q...H...V...S...P...N...D...Y...G  
 Sulfolobus\_islandicus .....MKKPI...I...I...N...L...K...A...V...P...C...D...P...V...A...V...A...A...R...D...V...A...A...E...S...G...V...R...I...A...V...A...P...Q...A...A...D...L...A...A...V...A...D...T...G...V...E...T...W...A...Q...H...V...S...P...N...D...Y...G  
 consensus>70 .....i.v.N.K.y.....a.....v.e.....g.i.vap...d...v.....i.v.aqhvd.v...g



	TT	α3	β4	TT	α4	β5	α5	β6												
	70	80	90	100	110	120	130	140	150											
<i>Methanocaldococcus jannaschii</i>	SHTGHLLA	AKDCCK	GLLHSEK	RMLLAD	LD	AV	LNKCKN	GLTELV	LNLSKA	AAAS	SPDY	IAV	FP	EL	GTG	IP	VS			
<i>Aeropyrum pernix</i>	AHTGYTPVE	AAAEGLR	HSEKRV	TVYRH	QAV	VRKARSL	LGLEVL	AD	PEEAAA	AALR	RPSM	VAIV	FP	EL	GTG	IP	VS			
<i>Aeropyrum camini</i>	AHTGYTPVE	AAAEGLR	HSEKRV	TVYRH	QAV	VRKARSL	LGLEVL	AD	PEEAAA	AALR	RPSM	VAIV	FP	EL	GTG	IP	VS			
<i>Staphylothermus marinus</i>	AKHTYTPVE	AAAEGLR	HSEKRV	TVYRH	QAV	VRKARSL	LGLEVL	AD	PEEAAA	AALR	RPSM	VAIV	FP	EL	GTG	IP	VS			
<i>Aphylothermus hellanicus</i>	AKHTYTPVE	AAAEGLR	HSEKRV	TVYRH	QAV	VRKARSL	LGLEVL	AD	PEEAAA	AALR	RPSM	VAIV	FP	EL	GTG	IP	VS			
<i>Ignicoccus hospitalis</i>	AHTGVSILE	KEI	IDGSI	ISEK	RMRIDE	SLV	VSKRLRAG	KE	SLVAF	PNTAAA	VAA	RPSM	IAV	FP	EL	GTG	IP	VS		
<i>Desulfurococcus kamchatkensis</i>	AHTGVSILE	KEI	IDGSI	ISEK	RMRIDE	SLV	VSKRLRAG	KE	SLVAF	PNTAAA	VAA	RPSM	IAV	FP	EL	GTG	IP	VS		
<i>Desulfurococcus mucosus</i>	AHTGVSILE	KEI	IDGSI	ISEK	RMRIDE	SLV	VSKRLRAG	KE	SLVAF	PNTAAA	VAA	RPSM	IAV	FP	EL	GTG	IP	VS		
<i>Desulfurococcus fermentans</i>	AHTGVSILE	KEI	IDGSI	ISEK	RMRIDE	SLV	VSKRLRAG	KE	SLVAF	PNTAAA	VAA	RPSM	IAV	FP	EL	GTG	IP	VS		
<i>Ignisphaera aggregans</i>	AHTGVSILE	KEI	IDGSI	ISEK	RMRIDE	SLV	VSKRLRAG	KE	SLVAF	PNTAAA	VAA	RPSM	IAV	FP	EL	GTG	IP	VS		
<i>Hyperthermus butylicus</i>	AHTGVSILE	KEI	IDGSI	ISEK	RMRIDE	SLV	VSKRLRAG	KE	SLVAF	PNTAAA	VAA	RPSM	IAV	FP	EL	GTG	IP	VS		
<i>Pyrolobus fumarii</i>	AHTGVSILE	KEI	IDGSI	ISEK	RMRIDE	SLV	VSKRLRAG	KE	SLVAF	PNTAAA	VAA	RPSM	IAV	FP	EL	GTG	IP	VS		
<i>Pyrobaculum aerophilum</i>	AHTGVSILE	KEI	IDGSI	ISEK	RMRIDE	SLV	VSKRLRAG	KE	SLVAF	PNTAAA	VAA	RPSM	IAV	FP	EL	GTG	IP	VS		
<i>Pyrobaculum islandicum</i>	AHTGVSILE	KEI	IDGSI	ISEK	RMRIDE	SLV	VSKRLRAG	KE	SLVAF	PNTAAA	VAA	RPSM	IAV	FP	EL	GTG	IP	VS		
<i>Pyrobaculum arsenicum</i>	AHTGVSILE	KEI	IDGSI	ISEK	RMRIDE	SLV	VSKRLRAG	KE	SLVAF	PNTAAA	VAA	RPSM	IAV	FP	EL	GTG	IP	VS		
<i>Pyrobaculum caldifontis</i>	AHTGVSILE	KEI	IDGSI	ISEK	RMRIDE	SLV	VSKRLRAG	KE	SLVAF	PNTAAA	VAA	RPSM	IAV	FP	EL	GTG	IP	VS		
<i>Thermoproteus neutrophilus</i>	AHTGVSILE	KEI	IDGSI	ISEK	RMRIDE	SLV	VSKRLRAG	KE	SLVAF	PNTAAA	VAA	RPSM	IAV	FP	EL	GTG	IP	VS		
<i>Pyrobaculum oguniense</i>	AHTGVSILE	KEI	IDGSI	ISEK	RMRIDE	SLV	VSKRLRAG	KE	SLVAF	PNTAAA	VAA	RPSM	IAV	FP	EL	GTG	IP	VS		
<i>Thermoproteus tenax</i>	AHTGVSILE	KEI	IDGSI	ISEK	RMRIDE	SLV	VSKRLRAG	KE	SLVAF	PNTAAA	VAA	RPSM	IAV	FP	EL	GTG	IP	VS		
<i>Thermoproteus uzoniensis</i>	AHTGVSILE	KEI	IDGSI	ISEK	RMRIDE	SLV	VSKRLRAG	KE	SLVAF	PNTAAA	VAA	RPSM	IAV	FP	EL	GTG	IP	VS		
<i>Caldivirga maguilingensis</i>	AHTGVSILE	KEI	IDGSI	ISEK	RMRIDE	SLV	VSKRLRAG	KE	SLVAF	PNTAAA	VAA	RPSM	IAV	FP	EL	GTG	IP	VS		
<i>Vulcanisaeta districata</i>	RTHGVPPE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Vulcanisaeta moutnovskia</i>	RTGHVPLE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Thermofilum pendens</i>	SYTHVPLE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Sulfolobus solfataricus</i>	AFTGAILPE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Sulfolobus tokodaii</i>	AHTGAILPE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Sulfolobus acidocaldarius</i>	AHTGAILPE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Halosphaera sedula</i>	AHTGAILPE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Metallosphaera cuprina</i>	AHTGAILPE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Metallosphaera yellowstonensis</i>	AHTGAILPE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Acidianus hospitalis</i>	AHTGAILPE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Acidilobus saccharovorans</i>	AHTGAILPE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Ferrodicrococcus fontis</i>	AHTGAILPE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Caldisphaera laguncularis</i>	AHTGAILPE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Archaeoglobus fulgidus</i>	SHTRINADM	AEYGA	GSLV	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS	
<i>Archaeoglobus profundus</i>	SYTRITAE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Archaeoglobus veneficus</i>	SHTRITAE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Ferroglobus placidus</i>	SHTRITAE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Zhaococcus sulfidocal</i>	SHTRITAE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Halobacterium salinarum</i>	SHTRITAE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Halorcula marismortui</i>	SHTRITAE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Halorcula hispanica</i>	SHTRITAE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Halorubrum lacusprofundi</i>	SHTRITAE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Halomicrobium mukohataei</i>	SHTRITAE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Haloterrigena curkmenica</i>	SHTRITAE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Halalkalicoccus jectajali</i>	SHTRITAE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Halogeometricum borinquense</i>	SHTRITAE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Haladaptatus paucihalophi</i>	SHTRITAE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Halopiger xanaduensis</i>	SHTRITAE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Natronobacterium gregoryi</i>	SHTRITAE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Natronococcus occultus</i>	SHTRITAE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Haloferrax mediterraneus</i>	SHTRITAE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Halostagnicola larsenii</i>	SHTRITAE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Natronorubrum tibetense</i>	SHTRITAE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Natrinema versiforme</i>	SHTRITAE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Natrialba magadii</i>	SHTRITAE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Halorubrum kokorui</i>	SHTRITAE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Scrialobus chahamaensis</i>	SHTRITAE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Natronococcus amylolyticus</i>	SHTRITAE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Natrinema altunense</i>	SHTRITAE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Haloferrax alexandrinus</i>	SHTRITAE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Halorubrum aidingense</i>	SHTRITAE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Haloterrigena thermotolerans</i>	SHTRITAE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Thermococcus maritimus</i>	SHTRITAE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Natronolimobius innermon</i>	SHTRITAE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Haloquadratum walsbyi</i>	SHTRITAE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Haloferrax volcanii</i>	SHTRITAE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Methanosarcina acetivorans</i>	SHTRITAE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA	AAA	SPDY	IAV	FP	EL	GTG	IP	VS
<i>Methanosarcina mazei</i>	SHTRITAE	KEI	IGAS	GVIL	SEKRL	IDE	WLA	VRAARAG	LE	TVVAF	PD	PTSLA								

<i>Thermoplasma volcanium</i>	PYTGHTIAIESIMNYGIIGSLLHSEERLIGEDK	IST	VKKQMIQIGFEIALCCEMEBAKR	YSAKPSPIAYE	PKELTGGNVSVS	
<i>Picrophilus_torridus</i>	PYTGSIISMNFKLSINVNGSLLHSEKRLIEDD	RKT	VEKSKSHDFKLVVLENDIDEAK	YSKRPPIAYE	FRELIGDIDSVS	
<i>Ferroplasma_acidarmanus</i>	PYTGSIISMNFKLSINVNGSLLHSEKRVDSST	RSV	VEKSKSLDGLKIIILVEINDIDEAK	YSKRPPIAYE	FRELIGDIDSVS	
<i>Korarchaeum_cryptofillum</i>	AVTGHVFDLILKAGARSGSLLHSEKQIPHRH	AES	INMRLDILISTVDFSSAG	YAKLIDIAIY	FRELIGSGIFVS	
<i>Nitrososphaera_vulcanensis</i>	STGCFVPELILKKSQVNGSLLHSEKRIEIGV	TKL	VEKSKLKLKTSIVKDVBAEAK	YAKLNPNYIAIE	FRELIGSGKAVS	
<i>Nitrososphaera_maritimus</i>	STGCFVPELILKKSQVNGSLLHSEKRIEIGV	TKL	VEKSKLKLKTSIVKDVBAEAK	YAKLNPNYIAIE	FRELIGSGKAVS	
<i>Ca_Nitrosopumilus_koreans</i>	STGCFVPELILKKSQVNGSLLHSEKRIEIGV	TKL	VEKSKLKLKTSIVKDVBAEAK	YAKLNPNYIAIE	FRELIGSGKAVS	
<i>Ca_Nitrososphaera_limnia</i>	STGCFVPELILKKSQVNGSLLHSEKRIEIGV	TKL	VEKSKLKLKTSIVKDVBAEAK	YAKLNPNYIAIE	FRELIGSGKAVS	
<i>Ca_Nitrosopumilus_salaria</i>	STGCFVPELILKKSQVNGSLLHSEKRIEIGV	TKL	VEKSKLKLKTSIVKDVBAEAK	YAKLNPNYIAIE	FRELIGSGKAVS	
<i>Cenarchaeum_symbiosum</i>	STGCFVPELILKKSQVNGSLLHSEKRIEIGV	TKL	VEKSKLKLKTSIVKDVBAEAK	YAKLNPNYIAIE	FRELIGSGKAVS	
<i>Ca_Methanomethylphilus_a</i>	STGCFVPELILKKSQVNGSLLHSEKRIEIGV	TKL	VEKSKLKLKTSIVKDVBAEAK	YAKLNPNYIAIE	FRELIGSGKAVS	
<i>Ca_Parvarchaenum_acidiphil</i>	AHTAHITWFEIKNTGVSTGLLHSEKRYTCSK	DGPLAYDK	VEA	VKRCMNLGLETIVVQNLDEAR	VLMKPTIAYE	FRELIGGNSIVS
<i>Ca_Parvarchaenum_acidiphil</i>	AHTAHITWFEIKNTGVSTGLLHSEKRYTCSK	DGPLAYDK	VEA	VKRCMNLGLETIVVQNLDEAR	VLMKPTIAYE	FRELIGGNSIVS
<i>Ca_Caldiarchaeum_subterra</i>	AYTGSSVVAEAKESGAGAGSLLHSEKRLKISD	SEA	VNRLRANGLIVSIVADDDHYATA	AAALGPDVAV	FRELIGTGISVS	
<i>Thermofilum_carboxyditrop</i>	STGHWTLLEAKVDAGATITVMVHSEKRVARD	DDV	VKRLRANGLIVSIVADDDHYATA	AAALGPDVAV	FRELIGTGISVS	
<i>Halanaeorchaenum_sulfuri</i>	STGHWTLLEAKVDAGATITVMVHSEKRVARD	DDV	VKRLRANGLIVSIVADDDHYATA	AAALGPDVAV	FRELIGTGISVS	
<i>Halofexax_gibbosii</i>	STGHWTLLEAKVDAGATITVMVHSEKRVARD	DDV	VKRLRANGLIVSIVADDDHYATA	AAALGPDVAV	FRELIGTGISVS	
<i>Haloterrigena_jeoatgii</i>	SNTHALAEHADVAGAVGTLNHSEKRLKIAE	DDG	VAAARADLETIVVANNFAQIGA	AAALGPDVAV	FRELIGTGISVS	
<i>Halolamina_rubra</i>	SNTHALAEHADVAGAVGTLNHSEKRLKIAE	DDG	VAAARADLETIVVANNFAQIGA	AAALGPDVAV	FRELIGTGISVS	
<i>Halobellus_rufus</i>	SNTHALAEHADVAGAVGTLNHSEKRLKIAE	DDG	VAAARADLETIVVANNFAQIGA	AAALGPDVAV	FRELIGTGISVS	
<i>Halopiger_salifodinae</i>	SNTHALAEHADVAGAVGTLNHSEKRLKIAE	DDG	VAAARADLETIVVANNFAQIGA	AAALGPDVAV	FRELIGTGISVS	
<i>Halapiricum_salinum</i>	SNTHALAEHADVAGAVGTLNHSEKRLKIAE	DDG	VAAARADLETIVVANNFAQIGA	AAALGPDVAV	FRELIGTGISVS	
<i>Halococcus_sediminicola</i>	SNTHALAEHADVAGAVGTLNHSEKRLKIAE	DDG	VAAARADLETIVVANNFAQIGA	AAALGPDVAV	FRELIGTGISVS	
<i>Thermococcus_eurythermali</i>	SNTHALAEHADVAGAVGTLNHSEKRLKIAE	DDG	VAAARADLETIVVANNFAQIGA	AAALGPDVAV	FRELIGTGISVS	
<i>Halorubrum_halophilum</i>	SNTHALAEHADVAGAVGTLNHSEKRLKIAE	DDG	VAAARADLETIVVANNFAQIGA	AAALGPDVAV	FRELIGTGISVS	
<i>Halorubrum_ezemouliense</i>	SNTHALAEHADVAGAVGTLNHSEKRLKIAE	DDG	VAAARADLETIVVANNFAQIGA	AAALGPDVAV	FRELIGTGISVS	
<i>Haloplanus_natans</i>	SNTHALAEHADVAGAVGTLNHSEKRLKIAE	DDG	VAAARADLETIVVANNFAQIGA	AAALGPDVAV	FRELIGTGISVS	
<i>Falaecoccus_ferrophilus</i>	SNTHALAEHADVAGAVGTLNHSEKRLKIAE	DDG	VAAARADLETIVVANNFAQIGA	AAALGPDVAV	FRELIGTGISVS	
<i>Aethanohalobium_pavarii</i>	AVTGHVFDLILKAGARSGSLLHSEKQIPHRH	AES	INMRLDILISTVDFSSAG	YAKLIDIAIY	FRELIGSGIFVS	
<i>Methanosarcina_thermophila</i>	SFTGHVFGKAKKEAGAVGTLNHSEKRLTLAE	LEAS	LKSAKEFFGLKAVIINNVPVTTAA	AAALGPDVAV	FRELIGSGIFVS	
<i>Methanosarcina_siciliae</i>	SFTGHVFGKAKKEAGAVGTLNHSEKRLTLAE	LEAS	LKSAKEFFGLKAVIINNVPVTTAA	AAALGPDVAV	FRELIGSGIFVS	
<i>Geoglobus_acetivorans</i>	SFTGHVFGKAKKEAGAVGTLNHSEKRLTLAE	LEAS	LKSAKEFFGLKAVIINNVPVTTAA	AAALGPDVAV	FRELIGSGIFVS	
<i>Geoglobus_ahanqari</i>	SFTGHVFGKAKKEAGAVGTLNHSEKRLTLAE	LEAS	LKSAKEFFGLKAVIINNVPVTTAA	AAALGPDVAV	FRELIGSGIFVS	
<i>Ca_Nitrosopelagicus_brevi</i>	AVTGHVFDLILKAGARSGSLLHSEKQIPHRH	AES	INMRLDILISTVDFSSAG	YAKLIDIAIY	FRELIGSGIFVS	
<i>Ca_Acidianus_cophaneus</i>	AVTGHVFDLILKAGARSGSLLHSEKQIPHRH	AES	INMRLDILISTVDFSSAG	YAKLIDIAIY	FRELIGSGIFVS	
<i>Ca_Methanoplasma_terminum</i>	AVTGHVFDLILKAGARSGSLLHSEKQIPHRH	AES	INMRLDILISTVDFSSAG	YAKLIDIAIY	FRELIGSGIFVS	
<i>Methanosarcina_lacustris</i>	SFTGHVFGKAKKEAGAVGTLNHSEKRLTLAE	LEAS	LKSAKEFFGLKAVIINNVPVTTAA	AAALGPDVAV	FRELIGSGIFVS	
<i>Methanosarcina_horobensis</i>	SFTGHVFGKAKKEAGAVGTLNHSEKRLTLAE	LEAS	LKSAKEFFGLKAVIINNVPVTTAA	AAALGPDVAV	FRELIGSGIFVS	
<i>Methanosarcina_vulcanata</i>	SFTGHVFGKAKKEAGAVGTLNHSEKRLTLAE	LEAS	LKSAKEFFGLKAVIINNVPVTTAA	AAALGPDVAV	FRELIGSGIFVS	
<i>Thermococcus_parrishii</i>	SFTGHVFGKAKKEAGAVGTLNHSEKRLTLAE	LEAS	LKSAKEFFGLKAVIINNVPVTTAA	AAALGPDVAV	FRELIGSGIFVS	
<i>Methermococcus_shengliensis</i>	SFTGHVFGKAKKEAGAVGTLNHSEKRLTLAE	LEAS	LKSAKEFFGLKAVIINNVPVTTAA	AAALGPDVAV	FRELIGSGIFVS	
<i>Methanobrevibacter_oralis</i>	GNTGHVLEPEAKKEAGAVGTLNHSEKRLTLAE	LEAS	LKSAKEFFGLKAVIINNVPVTTAA	AAALGPDVAV	FRELIGSGIFVS	
<i>Thermococcus_nautilii</i>	GNTGHVLEPEAKKEAGAVGTLNHSEKRLTLAE	LEAS	LKSAKEFFGLKAVIINNVPVTTAA	AAALGPDVAV	FRELIGSGIFVS	
<i>Methanocorpusculum_bavari</i>	AFTGHVLEPEAKKEAGAVGTLNHSEKRLTLAE	LEAS	LKSAKEFFGLKAVIINNVPVTTAA	AAALGPDVAV	FRELIGSGIFVS	
<i>Methanocorpusculum_mobile</i>	AFTGHVLEPEAKKEAGAVGTLNHSEKRLTLAE	LEAS	LKSAKEFFGLKAVIINNVPVTTAA	AAALGPDVAV	FRELIGSGIFVS	
<i>Methanobrevibacter_wolinii</i>	AFTGHVLEPEAKKEAGAVGTLNHSEKRLTLAE	LEAS	LKSAKEFFGLKAVIINNVPVTTAA	AAALGPDVAV	FRELIGSGIFVS	
<i>Desulfurococcus_amylolytic</i>	AFTGHVLEPEAKKEAGAVGTLNHSEKRLTLAE	LEAS	LKSAKEFFGLKAVIINNVPVTTAA	AAALGPDVAV	FRELIGSGIFVS	
<i>Ca_Nitrososphaera_everglia</i>	STGCFVPELILKKSQVNGSLLHSEKRIEIGV	TKL	VEKSKLKLKTSIVKDVBAEAK	YAKLNPNYIAIE	FRELIGSGKAVS	
<i>Ca_Halobonum_tyrrellensis</i>	STGCFVPELILKKSQVNGSLLHSEKRIEIGV	TKL	VEKSKLKLKTSIVKDVBAEAK	YAKLNPNYIAIE	FRELIGSGKAVS	
<i>Halorhabdus_tiamatae</i>	RGMGKLLPEAKKEAGAVGTLNHSEKRLTLAE	LEAS	LKSAKEFFGLKAVIINNVPVTTAA	AAALGPDVAV	FRELIGSGIFVS	
<i>Halomicrobium_katesii</i>	RGMGKLLPEAKKEAGAVGTLNHSEKRLTLAE	LEAS	LKSAKEFFGLKAVIINNVPVTTAA	AAALGPDVAV	FRELIGSGIFVS	
<i>Thermococcus_clebensis</i>	SNTHALAEHADVAGAVGTLNHSEKRLKIAE	DDG	VAAARADLETIVVANNFAQIGA	AAALGPDVAV	FRELIGTGISVS	
<i>Natronomonas_moolapensis</i>	SNTHALAEHADVAGAVGTLNHSEKRLKIAE	DDG	VAAARADLETIVVANNFAQIGA	AAALGPDVAV	FRELIGTGISVS	
<i>Natronomonas_pharaonis</i>	SNTHALAEHADVAGAVGTLNHSEKRLKIAE	DDG	VAAARADLETIVVANNFAQIGA	AAALGPDVAV	FRELIGTGISVS	
<i>Thermococcus_zilligii</i>	SNTHALAEHADVAGAVGTLNHSEKRLKIAE	DDG	VAAARADLETIVVANNFAQIGA	AAALGPDVAV	FRELIGTGISVS	
<i>Halococcus_pallidum</i>	SNTHALAEHADVAGAVGTLNHSEKRLKIAE	DDG	VAAARADLETIVVANNFAQIGA	AAALGPDVAV	FRELIGTGISVS	
<i>Ca_Nanopusillus_sp</i>	AVTGHVFDLILKAGARSGSLLHSEKQIPHRH	AES	INMRLDILISTVDFSSAG	YAKLIDIAIY	FRELIGSGIFVS	
<i>Sulfolobus_pellirubrum</i>	RGMGKLLPEAKKEAGAVGTLNHSEKRLTLAE	LEAS	LKSAKEFFGLKAVIINNVPVTTAA	AAALGPDVAV	FRELIGSGIFVS	
<i>Halorhabdus_utahensis</i>	RGMGKLLPEAKKEAGAVGTLNHSEKRLTLAE	LEAS	LKSAKEFFGLKAVIINNVPVTTAA	AAALGPDVAV	FRELIGSGIFVS	
<i>Halovivax_ruber</i>	SNTHALAEHADVAGAVGTLNHSEKRLKIAE	DDG	VAAARADLETIVVANNFAQIGA	AAALGPDVAV	FRELIGTGISVS	
<i>Salinarchaeum_sp</i>	SNTHALAEHADVAGAVGTLNHSEKRLKIAE	DDG	VAAARADLETIVVANNFAQIGA	AAALGPDVAV	FRELIGTGISVS	
<i>Halorubrum_arciis</i>	SNTHALAEHADVAGAVGTLNHSEKRLKIAE	DDG	VAAARADLETIVVANNFAQIGA	AAALGPDVAV	FRELIGTGISVS	
<i>Halorubrum_lipolyticum</i>	SNTHALAEHADVAGAVGTLNHSEKRLKIAE	DDG	VAAARADLETIVVANNFAQIGA	AAALGPDVAV	FRELIGTGISVS	
<i>Halococcus_litorum</i>	SNTHALAEHADVAGAVGTLNHSEKRLKIAE	DDG	VAAARADLETIVVANNFAQIGA	AAALGPDVAV	FRELIGTGISVS	
<i>Halococcus_thailandensis</i>	SNTHALAEHADVAGAVGTLNHSEKRLKIAE	DDG	VAAARADLETIVVANNFAQIGA	AAALGPDVAV	FRELIGTGISVS	
<i>Halococcus_saccharolyticus</i>	SNTHALAEHADVAGAVGTLNHSEKRLKIAE	DDG	VAAARADLETIVVANNFAQIGA	AAALGPDVAV	FRELIGTGISVS	
<i>Halococcus_hamelinensis</i>	SNTHALAEHADVAGAVGTLNHSEKRLKIAE	DDG	VAAARADLETIVVANNFAQIGA	AAALGPDVAV	FRELIGTGISVS	
<i>Halobiforma_lacisalsi</i>	SNTHALAEHADVAGAVGTLNHSEKRLKIAE	DDG	VAAARADLETIVVANNFAQIGA	AAALGPDVAV	FRELIGTGISVS	
<i>Halobacterium_japonicum</i>	SNTHALAEHADVAGAVGTLNHSEKRLKIAE	DDG	VAAARADLETIVVANNFAQIGA	AAALGPDVAV	FRELIGTGISVS	
<i>Halobacterium_californiae</i>	SNTHALAEHADVAGAVGTLNHSEKRLKIAE	DDG	VAAARADLETIVVANNFAQIGA	AAALGPDVAV	FRELIGTGISVS	
<i>Halobacterium_argentinensis</i>	SNTHALAEHADVAGAVGTLNHSEKRLKIAE	DDG	VAAARADLETIVVANNFAQIGA	AAALGPDVAV	FRELIGTGISVS	
<i>Halobacterium_amylolyticum</i>	SNTHALAEHADVAGAVGTLNHSEKRLKIAE	DDG	VAAARADLETIVVANNFAQIGA	AAALGPDVAV	FRELIGTGISVS	
<i>Halobacterium_sinaiensis</i>	SNTHALAEHADVAGAVGTLNHSEKRLKIAE	DDG	VAAARADLETIVVANNFAQIGA	AAALGPDVAV	FRELIGTGISVS	
<i>Halobacterium_vallismortis</i>	SNTHALAEHADVAGAVGTLNHSEKRLKIAE	DDG	VAAARADLETIVVANNFAQIGA	AAALGPDVAV	FRELIGTGISVS	
<i>Halobacterium_coriense</i>	SNTHALAEHADVAGAVGTLNHSEKRLKIAE	DDG	VAAARADLETIVVANNFAQIGA	AAALGPDVAV	FRELIGTGISVS	
<i>Halobacterium_distributum</i>	SNTHALAEHADVAGAVGTLNHSEKRLKIAE	DDG	VAAARADLETIVVANNFAQIGA	AAALGPDVAV	FRELIGTGISVS	
<i>Halosimplex_carlsbadense</i>	SNTHALAEHADVAGAVGTLNHSEKRLKIAE	DDG	VAAARADLETIVVANNFAQIGA	AAALGPDVAV	FRELIGTGISVS	
<i>Ca_Nanopusillus_acidilobi</i>	AVTGHVFDLILKAGARSGSLLHSEKQIPHRH	AES	INMRLDILISTVDFSSAG	YAKLIDIAIY	FRELIGSGIFVS	
<i>Ca_Nitrosopumilus_adriati</i>	STGCFVPELILKKSQVNGSLLHSEKRIEIGV	TKL	VEKSKLKLKTSIVKDVBAEAK	YAKLNPNYIAIE	FRELIGSGKAVS	
<i>Ca_Nitrosotenuis_cloacae</i>	STGCFVPELILKKSQVNGSLLHSEKRIEIGV	TKL	VEKSKLKLKTSIVKDVBAEAK	YAKLNPNYIAIE	FRELIGSGKAVS	
<i>Halobacterium_huberi</i>	SNTHALAEHADVAGAVGTLNHSEKRLKIAE	DDG	VAAARADLETIVVANNFAQIGA	AAALGPDVAV	FRELIGTGISVS	
<i>Thermococcus_guaymasensis</i>	SFTGHVFGKAKKEAGAVGTLNHSEKRLTLAE	LEAS	LKSAKEFFGLKAVIINNVPVTTAA	AAALGPDVAV	FRELIGSGIFVS	
<i>Methanocaldococcus_batho</i>	SFTGHVFGKAKKEAGAVGTLNHSEKRLTLAE	LEAS	LKSAKEFFGLKAVIINNVPVTTAA	AAALGPDVAV	FRELIGSGIFVS	
<i>Pyrobaculum_ferrireducens</i>	AYTGHVLEPEAKKEAGAVGTLNHSEKRLTLAE	LEAS	LKSAKEFFGLKAVIINNVPVTTAA	AAALGPDVAV	FRELIGSGIFVS	
<i>Thermococcus_celericrescens</i>	SFTGHVFGKAKKEAGAVGTLNHSEKRLTLAE	LEAS	LKSAKEFFGLKAVIINNVPVTTAA	AAALGPDVAV	FRELIGSGIFVS	
<i>Thermococcus_theroducens</i>	SFTGHVFGKAKKEAGAVGTLNHSEKRLTLAE	LEAS	LKSAKEFFGLKAVIINNVPVTTAA	AAALGPDVAV	FRELIGSGIFVS	
<i>Producticum_deltam</i>	AVTGHVFDLILKAGARSGSLLHSEKQIPHRH	AES	INMRLDILISTVDFSSAG	YAKLIDIAIY	FRELIGSGIFVS	
<i>Acidiplasma_cupricumans</i>	AVTGHVFDLILKAGARSGSLLHSEKQIPHRH	AES	INMRLDILISTVDFSSAG	YAKLIDIAIY	FRELIGSGIFVS	
<i>Acidiplasma_aeolicum</i>	AVTGHVFDLILKAGARSGSLLHSEKQIPHRH	AES	INMRLDILISTVDFSSAG	YAKLIDIAIY	FRELIGSGIFVS	
<i>Halolamina_pelagica</i>	SNTHALAEHADVAGAVGTLNHSEKRLKIAE	DDG	VAAARADLETIVVANNFAQIGA	AAALGPDVAV	FRELIGTGISVS	
<i>Methanosarcina_flavescens</i>	SFTGHVFGKAKKEAGAVGTLNHSEKRLTLAE	LEAS	LKSAKEFFGLKAVIINNVPVTTAA	AAALGPDVAV	FRELIGSGIFVS	
<i>Methanococcus_sediminis</i>	SFTGHVFGKAKKEAGAVGTLNHSEKRLTLAE	LEAS	LKSAKEFFGLKAVIINNVPVTTAA	AAALGPDVAV	FRELIGSGIFVS	
<i>Ca_Nitrososphaera_gargensis</i>	AVTGHVFDLILKAGARSGSLLHSEKQIPHRH	AES	INMRLDILISTVDFSSAG	YAKLIDIAIY	FRELIGSGIFVS	
<i>Methanobrevibacter_fillo</i>	GHTGNSLIEFTEKESGAGTLNHSEKRLTLAE	LEAS	LKSAKEFFGLKAVIINNVPVTTAA	AAALGPDVAV	FRELIGSGIFVS	
<i>Methanobrevibacter_curvat</i>	GHTGNSLIEFTEKESGAGTLNHSEKRLTLAE	LEAS	LKSAKEFFGLKAVIINNVPVTTAA	AAALGPDVAV	FRELIGSGIFVS	
<i>Methanobrevibacter_cuticu</i>	GHTGNSLIEFTEKESGAGTLNHSEKRLTLAE	LEAS	LKSAKEFFGLKAVIINNVPVTTAA	AAALGPDVAV	FRELIGSGIFVS	
<i>Ca_Nitrosopumilus_piranen</i>	STGCFVPELILKKSQVNGSLLHSEKRIEIGV	TKL	VEKSKLKLKTSIVKDVBAEAK	YAKLNPNYIAIE	FRELIGSGKAVS	
<i>Thermococcus_chitonophagu</i>	SFTGHVFGKAKKEAGAVGTLNHSEKRLTLAE	LEAS	LKSAKEFFGLKAVIINNVPVTTAA	AAALGPDVAV	FRELIGSGIFVS	
<i>Ca_Thalassosphaera_marina</i>	SNTHALAEHADVAGAVGTLNHSEKRLKIAE	DDG	VAAARADLETIVVANNFAQIGA	AAALGPDVAV	FRELIGTGISVS	
<i>Methanobrevibacter_ollega</i>	GHTGNSLIEFTEKESGAGTLNHSEKRLTLAE	LEAS	LKSAKEFFGLKAVIINNVPVTTAA	AAALGPDVAV	FRELIGSGIFVS	
<i>Methanogenium_cariaci</i>	AHTGHVLEPEAKKEAGAVGTLNHSEKRLTLAE	LEAS	LKSAKEFFGLKAVIINNVPVTTAA	AAALGPDVAV	FRELIGSGIFVS	
<i>Thermococcus_ptonophilu</i>	SFTGHVFGKAKKEAGAVGTLNHSEKRLTLAE	LEAS	LKSAKEFFGLKAVIINNVPVTTAA	AAALGPDVAV	FRELIGSGIFVS	
<i>Methanobrevibacter_miller</i>	GHTGNSLIEFTEKESGAGTLNHSEKRLTLAE	LEAS	LKSAKEFFGLKAVIINNVPVTTAA	AAALGPDVAV	FRELIGSGIFVS	
<i>Pyrodicticum_occultum</i>	AHTGHVLEPEAKKEAGAVGTLNHSEKRLTLAE	LEAS	LKSAKEFFGLKAVIINNVPVTTAA	AAALGPDVAV	FRELIGSGIFVS	
<i>Riftia_sedimentis</i>	AVTGHVFDLILKAGARSGSLLHSEKQIPHRH	AES	INMRLDILISTVDFSSAG	YAKLIDIAIY	FRELIGSGIFVS	
<i>Sulfolobus_metallicus</i>	AHTGHVLEPEAKKEAGAVGTLNHSEKRLTLAE	LEAS	LKSAKEFFGLKAVIINNVPVTTAA	AAALGPDVAV	FRELIGSGIFVS	
<i>Methanobrevibacter_arzori</i>	GHTGNSLIEFTEKESGAGTLNHSEKRLTLAE	LEAS	LKSAKEFFGLKAVIINNVPVTTAA	AAALGPDVAV	FRELIGSGIFVS	
<i>Halobellus_limi</i>	AVTGHVFDLILKAGARSGSLLHSEKQIPHRH	AES	INMRLDILISTVDFSSAG	YAKLIDIAIY	FRELIGSGIFVS	
<i>Natronorubrum_bangense</i>	RGMGKLLPEAKKEAGAVGTLNHSEKRLTLAE	LEAS	LKSAKEFFGLKAVIINNVPVTTAA	AAALGPDVAV	FRELIGSGIFVS	
<i>Halorhabdus_rudnickae</i>	RGMGKLLPEAKKEAGAVGTLNHSEKRLTLAE	LEAS	LKSAKEFFGLKAVIINNVPVTTAA	AAALGPDVAV	FRELIGSGIFVS	
<i>Ca_Nitrososphaera_franki</i>	AVTGHVFDLILKAGARSGSLLHSEKQIPHRH	AES	INMRLDILISTVDFSSAG	YAKLIDIAIY	FRELIGSGIFVS	
<i>Methanomicrococcus_biatti</i>	AVTGHVFDLILKAGARSGSLLHSEKQIPHRH	AES	INMRLDILISTVDFSSAG	YAKLIDIAIY	FRELIGSGIFVS	
<i>Methanohalophilus_euhalob</i>	SNTHALAEHADVAGAVGTLNHSEKRLKIAE	DDG	VAAARADLETIVVANNFAQIGA	AAALGPDVAV	FRELIGTGISVS	
<i>Haloterrigena_hispanica</i>	SNTHALAEHADVAGAVGTLNHSEKRLKIAE	DDG	VAAARADLETIVVANNFAQIGA	AAALGPDVAV	FRELIGTGISVS	
<i>Methanococcus_taiwanensis</i>	TGTHVLEPEAKKEAGAVGTLNHSEKRLTLAE	LEAS	LKSAKEFFGLKAVIINNVPVTTAA	AAALGPDVAV	FRELIGSGIFVS	
<i>Ca_Methanosarcinus_sp</i>	AVTGHVFDLILKAGARSGSLLHSEKQIPHRH	AES	INMRLDILISTVDFSSAG	YAKLIDIAIY	FRELIGSGIFVS	
<i>Methanohalophilus_sp</i>	SNTHALAEHADVAGAVGTLNHSEKRLKIAE	DDG	VAAARADLETIVVANNFAQIGA	AAALGPDVAV	FRELIGTGISVS	
<i>Natrarchaebius_chitinivus</i>	RGMGKLLPEAKKEAGAVGTLNHSEKRLTLAE	LEAS	LKSAKEFFGLKAVIINNVPVTTAA	AAALGPDVAV	FRELIGSGIFVS	
<i>Natrarchaebius_halalkali</i>	RGMGKLLPEAKKEAGAVGTLNHSEKRLTLAE	LEAS	LKSAKEFFGLKAVIINNVPVTTAA	AAALGPDVAV	FRELIGSGIFVS	
<i>Eury_archaeon_TMED85</i>	SNTHALAEHADVAGAVGTLNHSEKRLKIAE	DDG	VAAARADLETIVVANNFAQIGA	AAALGPDVAV	FRELIGTGISVS	
<i>Eury_archaeon_TMED248</i>	SNTHALAEHADVAGAVGTLNHSEKRLKIAE	DDG	VAAARADLETIVVANNFAQIGA	AAALGPDVAV	FRELIGTGISVS	
<i>Eury_archaeon_TMED1117</i>	SNTHALAEHADVAGAVGTLNHSEKRLKIAE	DDG	VAAARADLETIVVANNFAQIGA	AAALGPDVAV	FRELIGTGISVS	
<i>Methanoregularia_archaeo</i>	AVTGHVFDLILKAGARSGSLLHSEKQIPHRH	AES	INMRLDILISTVDFSSAG	YAKLIDIAIY	FRELIGSGIFVS	
<i>Nitrosopumilales_archaeo</i>	AVTGHVFDLILKAGARSGSLLHSEKQIPHRH	AES	INMRLDILISTVDFSSAG	YAKLIDIAIY	FRELIGSGIFVS	
<i>Methanobrevibacter_gottsc</i>	GHTGNSLIEFTEKESGAGTLNHSEKRLTLAE	LEAS	LKSAKEFFGLKAVIINNVPVTTAA	AAALGPDVAV	FRELIGSGIFVS	
<i>Halorubrum_sp</i>	SNTHALAEHADVAGAVGTLNHSEKRLKIAE	DDG	VAAARADLETIVVANNFAQIGA	AAALGPDVAV	FRELIGTGISVS	
<i>Natros</i>						

Nitrosopumilus\_sp\_H8 STIGFVVPVPELLKKSQVAGSLINHSEHRIPLGLE...IKKT...VLRRLRKLKMSIVCVKVVAAKAK...YARLGPDYIAIE...PPELIGSGKAVS...  
 Halorarchaeum\_sp\_CBA1220 SHTGSTLAERAAAGAEETLLNHSENRKMLAD...LDAS...VGAARERAGLETVCANNPFOVGA...VAALGPDVAVAE...PPELIGGDVSV...  
 Halorubrum\_sp\_CBA1229 SHTGSTLAERAAAGAEETLLNHSENRKMLAD...VDGS...VRAAERAGLETVCANNPFOVGA...VAALGPDVAVAE...PPELIGGDVSV...  
 Methanohalophilus\_halophi SNTGHVMAISCVQAGAVGSLINHSENRKMLAD...HEAS...VRAAERAGLETVCANNPFOVGA...VAALGPDVAVAE...PPELIGGDVSV...  
 Methanohalophilus\_portuca SNTGHVMAISCVQAGAVGSLINHSENRKMLAD...HEAS...VRAAERAGLETVCANNPFOVGA...VAALGPDVAVAE...PPELIGGDVSV...  
 Halalkalicoccus\_sp\_GSM28 SHTGSALAEISLAAAGATGTMINHSENRKMLAD...LDAG...IGAARERAGLETVCANNPFOVGA...VAALGPDVAVAE...PPELIGGDVSV...  
 Halorubrum\_sp\_ZY8 SHTGSTLAERAAAGAEETLLNHSENRKMLAD...VDGS...VRAAERAGLETVCANNPFOVGA...VAALGPDVAVAE...PPELIGGDVSV...  
 Ca\_Woesearchaeota\_archaeo AHTGALIPQALDLSGANALINHSERPVPLSN...LEEL...VKMFKLLGLISIVCAGSARAESR...LLGSPDLLAIE...PPELIGGDVSV...  
 Haloplanus\_aerogenes SHTGHTLAEAVADAGAEETLLNHSENRKMLAD...LDAS...VGAARERAGLETVCANNPFOVGA...VAALGPDVAVAE...PPELIGGDVSV...  
 Halostella\_sp\_LTI2 SHTGSALAEAAADAGAEETLLNHSENRKMLAD...VDAS...VGAARERAGLETVCANNPFOVGA...VAALGPDVAVAE...PPELIGGDVSV...  
 Halobellus\_sp\_Atlit-31R SHTGSTLAERAAAGAEETLLNHSENRKMLAD...VDAA...VRAAERAGLETVCANNPFOVGA...VAALGPDVAVAE...PPELIGGDVSV...  
 Ca\_Korarchaeota\_archaeon ARTGHVSPRLREYGLKSLINHSENRKMLAD...VRAA...VQMLRGDGLASIVCAGDEEESIA...AAALRPTAVAE...PPELIGTGISVS...  
 Ca\_Thorarchaeota\_archaeon RHTHGVLGETVVEAGCSVTLNHSENRKMLAD...LEEA...VRAAERAGLETVCANNPFOVGA...VAALGPDVAVAE...PPELIGGDVSV...  
 Ca\_Lokiarchaeota\_archaeon APTGEIVLSDLLNAGCEVYLLHSENRKMLAD...TNAQINQMKKALEKGLKPVLCGETAQRERDGMEEVNRILQLEESLADIA...EQLSQTIVAYE...FVWAINNPLNPNV...  
 Ca\_Heimdallarchaeota\_arch SHTGHLLLETVKEAGATGTLVNHSENRKMLAD...TALI...IQRAENHLYTVCANNPFOVGA...VAALGPDVAVAE...PPELIGGDVSV...  
 Ca\_Altiarchaeales\_archaeo SNTGHLLLETVKEAGATGTLVNHSENRKMLAD...TALI...IQRAENHLYTVCANNPFOVGA...VAALGPDVAVAE...PPELIGGDVSV...  
 Ca\_Aenigmarchaeota\_archae SNTGHDMPECKENAGAVGSLINHSENRKMLAD...IKKT...IFCKEAGKHTVCEHTAEGKD...VVKMLPDVIAIE...PPELIGTGISVS...  
 Halonotius\_pteroides SFTGSTLAERAAAGAEETLLNHSENRKMLAD...LDAA...VRAAERAGLETVCANNPFOVGA...VAALGPDVAVAE...PPELIGGDVSV...  
 Halonotius\_sp\_F13-13 SFTGSTLAERAAAGAEETLLNHSENRKMLAD...LDAA...VRAAERAGLETVCANNPFOVGA...VAALGPDVAVAE...PPELIGGDVSV...  
 Halopiger\_aswanensis SNTGHTLAEETVADAGAEETLLNHSENRKMLAD...LDGA...VQAAQRADLETVCANNPFOVGA...VAALGPDVAVAE...PPELIGTGTFS...  
 Ca\_Nanoclepta\_minutus ASHTGHILPEAVKEAGAVTLLNHSENRKMLAD...LGEA...TEIARKIGLITVVCANNPFOVGA...VAALGPDVAVAE...PPELIGTGTFS...  
 Natronolimnolobus\_sp\_AArc1 SNTGHTLAEETVADAGAEETLLNHSENRKMLAD...LDGS...VRAAERAGLETVCANNPFOVGA...VAALGPDVAVAE...PPELIGTGTFS...  
 Methanobrevibacter\_woesei GHFGHDLIDTLEAGVSGSLINHSENRKMLAD...LDEV...VKLTRENHLLSCVCCANNPFOVGA...VAALGPDVAVAE...PPELIGTGTFS...  
 Methanothermobacter\_deflu GHFGSGLAECAARDAGAEETLLNHSENRKMLAD...LDAS...VRAAERAGLETVCANNPFOVGA...VAALGPDVAVAE...PPELIGTGTFS...  
 Halorussus\_rarus SHTGSTLAERAAADAGAEETLLNHSENRKMLAD...LDAS...VRAAERAGLETVCANNPFOVGA...VAALGPDVAVAE...PPELIGTGTFS...  
 Halopelagius\_longus SHTGSTLAERAAADAGAEETLLNHSENRKMLAD...LDAS...VRAAERAGLETVCANNPFOVGA...VAALGPDVAVAE...PPELIGTGTFS...  
 Haloplanus\_salinus SHTGHTLAEARAAAGATGTLLNHSENRKMLAD...VDGA...VGAARERAGLETVCANNPFOVGA...VAALGPDVAVAE...PPELIGGDVSV...  
 Methanothermobacter\_teneb GHFGSILLETCAKEAGATGTLLNHSENRKMLAD...LQRV...IEKVEKADLISVVCANNPFOVGA...VAALGPDVAVAE...PPELIGSGIFVS...  
 Acidianus\_brixleyi ARTGHTVTVEHMKDAGAKGSLINHSENRKMLAD...LDDA...IKKLNKLEGLLESVVCANNPFOVGA...VAALGPDVAVAE...PPELIGSGIFVS...  
 Metallosphaera\_hakonensis AHTGAVTPEHMKDAGAKGSLINHSENRKMLAD...LDDA...IKKLNKLEGLLESVVCANNPFOVGA...VAALGPDVAVAE...PPELIGSGIFVS...  
 Acidianus\_sulfidivorans\_J AHTGSIITAEHMKDAGAKGSLINHSENRKMLAD...LHDS...LTKMKKLGLESVVCANNPFOVGA...VAALGPDVAVAE...PPELIGSGIFVS...  
 Methanospirillum\_stamsii AHTGRMPLLAVKSAAGISSLVNHSEYRLNIAD...LEKN...VTALRRDAGMLSCVCCANNPFOVGA...VAALGPDVAVAE...PPELIGSGIFVS...  
 Methanospirillum\_lacunae AHTGRIPPLTAIKTAGAVGSLINHSEYRLNIAD...LETN...VQALRNEGMLSCVCCANNPFOVGA...VAALGPDVAVAE...PPELIGSGIFVS...  
 Methanosphaera\_cuniculi SNTGSLVLEALMEAGINLTLLNHSENRKMLAD...LDEI...IEITKRYDTSLSNNPFOVGA...VAALGPDVAVAE...PPELIGSGIFVS...  
 Methanobrevibacter\_thaueri GHFGSILNTLEAGVSGSLINHSENRKMLAD...LDEV...IKKLNKLEGLLESVVCANNPFOVGA...VAALGPDVAVAE...PPELIGSGIFVS...  
 Ca\_Nanosidianus\_stetteri AHTGVYILPEAVKEAGAVTLLNHSENRKMLAD...LEKA...INMAKRLGLTTIVCANNPFOVGA...VAALGPDVAVAE...PPELIGSGIFVS...  
 Sulfolobus\_islandicus AYTGVAVLPEHMKDAGAKGSLINHSENRKMLAD...LDV...LKRTRKLLGLSLLCVDVYELVYP...FSLRPPDAILIE...PPELIGTGTFS...  
 consensus>70 ..t.g...e...e.g.a.g.l.nHs#.r.l.l.d...id...v...l...vvc.nn...a...a.l.pd.vavE.ppeIlg.g...vs...



Methanocaldococcus jannas



Methanocaldococcus jannas	KAN	PEV	VEG	TV	RA	VK	EIN	KDV	KVLC	GAGIS	KGEDV	KA	LDL	GAE	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I					
Aeropyrum pernix	OAK	PEV	ITR	GE	VA	EA	RVA	PGV	VAVL	GAGIT	AGEDA	RR	AVEL	GAO	GVL	LAG	AVM	KA	D	PHG	KML	LELA	EAM	ARK				
Aeropyrum canini	RAR	PEV	ITR	GE	VA	EA	SVD	RGV	VAVL	GAGIT	GEDA	RR	AVEL	GAR	GVL	LAG	AVM	KA	D	PHG	KML	LELA	EAM	ARK				
Staphylothermus marinus	KAR	PEV	ITR	GE	VA	EA	KVN	KDV	VILV	GAGIS	RGEV	VY	AV	KLGT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I				
Staphylothermus hollancticus	KAR	PEV	ITR	GE	VA	EA	KVN	KDV	VILV	GAGIS	RGEV	VY	AV	KLGT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I				
Ignicoccus hospitalis	RAR	PEV	ITR	GE	VA	EA	CTG	FAG	VILV	GAGIS	SSEDV	KA	IEL	GAD	GVL	LAG	AVM	KA	D	PHG	KML	LELA	EAM	VR				
Desulfurococcus kamchatke	KAR	PEV	ITR	GE	VA	EA	RVN	KDV	VILV	GAGIT	GEDA	Y	AV	KLGT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I				
Desulfurococcus mucosus	RAR	PEV	ITR	GE	VA	EA	RVN	KEV	ILIT	V	GAGIS	NGEDA	Y	AV	KLGT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I			
Desulfurococcus fermentans	KAR	PEV	ITR	GE	VA	EA	RVN	KDV	VILV	GAGIT	GEDA	Y	AV	KLGT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I				
Thermoplasma aggregans	KAR	PEV	ITR	GE	VA	EA	RVN	KEV	ILIT	V	GAGIS	TEDE	V	AV	KLGT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I			
Thermoplasma volcanium	KAR	PEV	ITR	GE	VA	EA	RVN	KEV	ILIT	V	GAGIS	TEDE	V	AV	KLGT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I			
Hyperthermus butylicus	RAR	PEV	ITR	GE	VA	EA	RVN	KEV	ILIT	V	GAGIS	TEDE	V	AV	KLGT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I			
Pyrolobus fumarii	RAR	PEV	ITR	GE	VA	EA	RVS	PNV	VILV	GAGIS	VEDA	RR	S	VEL	GAR	GVL	LAG	AVM	KA	D	PHG	KML	LELA	EAM	VR			
Pyrobaculum aerophilum	KYK	PEV	ITR	GE	VA	EA	RHF	PDV	VILV	GAGIS	EGED	EA	AL	KLGT	R	GVL	LAG	AAV	KA	D	PHG	KML	LELA	EAM	VR			
Pyrobaculum islandicum	KYK	PEV	ITR	GE	VA	EA	RHF	PDV	VILV	GAGIS	EGED	EA	AL	KLGT	R	GVL	LAG	AAV	KA	D	PHG	KML	LELA	EAM	VR			
Pyrobaculum arsenaticum	RVR	PEV	ITR	GE	VA	EA	RHF	ADV	VILV	GAGIS	EGED	EA	AL	KLGT	R	GVL	LAG	AAV	KA	D	PHG	KML	LELA	EAM	VR			
Pyrobaculum caldifontis	KYK	PEV	ITR	GE	VA	EA	RHF	PDV	VILV	GAGIS	EGED	EA	AL	KLGT	R	GVL	LAG	AAV	KA	D	PHG	KML	LELA	EAM	VR			
Thermoproteus neutrophilus	RVR	PEV	ITR	GE	VA	EA	RHF	ADV	VILV	GAGIS	EGED	EA	AL	KLGT	R	GVL	LAG	AAV	KA	D	PHG	KML	LELA	EAM	VR			
Pyrobaculum oguniense	KYK	PEV	ITR	GE	VA	EA	RHF	PDV	VILV	GAGIS	EGED	EA	AL	KLGT	R	GVL	LAG	AAV	KA	D	PHG	KML	LELA	EAM	VR			
Thermoproteus tenax	RVR	PEV	ITR	GE	VA	EA	RHF	ADV	VILV	GAGIS	EGED	EA	AL	KLGT	R	GVL	LAG	AAV	KA	D	PHG	KML	LELA	EAM	VR			
Thermoproteus uzoniensis	RVR	PEV	ITR	GE	VA	EA	RHF	ADV	VILV	GAGIS	EGED	EA	AL	KLGT	R	GVL	LAG	AAV	KA	D	PHG	KML	LELA	EAM	VR			
Caldivirga maguilingsensis	RTR	PEV	ITR	GE	VA	EA	RVN	GSV	VILV	GAGIS	EDM	RR	KAL	EL	GT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I			
Vulcaniseta districata	REK	PEV	ITR	GE	VA	EA	RVN	PSV	VILV	GAGIS	YEDV	RR	AV	EL	GT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I			
Vulcaniseta moutnovskia	REK	PEV	ITR	GE	VA	EA	RVN	PSV	VILV	GAGIS	YEDV	RR	AV	EL	GT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I			
Thermofilum pendens	KAR	PEV	ITR	GE	VA	EA	RVN	PNV	VILV	GAGIT	VGED	Y	AA	LR	LT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I			
Sulfolobus sulfataricus	KAR	PEV	ITR	GE	VA	EA	RVN	PNV	VILV	GAGIT	VGED	Y	AA	LR	LT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I			
Sulfolobus tokodaii	KEK	PEV	ITR	GE	VA	EA	RVN	GVY	LIA	GAGIT	TGED	V	Y	AA	LR	LT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I		
Sulfolobus acidocaldarius	KAR	PEV	ITR	GE	VA	EA	RVN	KTK	K	VILV	GAGIT	TGED	V	Y	AA	LR	LT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I	
Metallophaea sedula	RAR	PEV	ITR	GE	VA	EA	RVN	GVY	LIA	GAGIT	TGED	V	Y	AA	LR	LT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I		
Metallophaea cuprina	RAR	PEV	ITR	GE	VA	EA	RVN	GVY	LIA	GAGIT	TGED	V	Y	AA	LR	LT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I		
Metallophaea yellowston	RAR	PEV	ITR	GE	VA	EA	RVN	G	LK	L	V	GAGIS	TGED	V	Y	AA	LR	LT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I
Acidianus hospitalis	KAR	PEV	ITR	GE	VA	EA	RVN	GVY	LIA	GAGIS	TGED	V	Y	AA	LR	LT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I		
Acidobolus saccharovorans	KAR	PEV	ITR	GE	VA	EA	RVN	ADV	VILV	GAGIT	YRED	V	RR	AV	EL	GT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I		
Feravidicoccus fontis	KSR	PEV	ITR	GE	VA	EA	RVN	KKI	VILV	GAGIT	YRED	V	RR	AV	EL	GT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I		
Caldisphaera lagunesis	KAR	PEV	ITR	GE	VA	EA	RVN	SSV	VILV	GAGIS	TGED	V	Y	AA	LR	LT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I		
Archaeoglobus fulgidus	KAR	PEV	ITR	GE	VA	EA	RVN	KS	VILV	GAGIT	TGED	V	Y	AA	LR	LT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I		
Archaeoglobus profundus	KAR	PEV	ITR	GE	VA	EA	RVN	PK	VILV	GAGIT	TGED	V	Y	AA	LR	LT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I		
Archaeoglobus veneficus	KAR	PEV	ITR	GE	VA	EA	RVN	PK	VILV	GAGIT	TGED	V	Y	AA	LR	LT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I		
Ferroglobus placidus	KAR	PEV	ITR	GE	VA	EA	RVN	EEV	VILV	GAGIT	YRED	V	RR	AV	EL	GT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I		
Archaeoglobus sulcatifall	KAD	PEV	ITR	GE	VA	EA	RVN	SNV	VILV	GAGIS	KHE	V	Y	AA	LR	LT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I		
Haloarcula salinarum	KAD	PEV	ITR	GE	VA	EA	RVN	DDV	VILV	GAGIS	TGED	V	Y	AA	LR	LT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I		
Haloarcula marismortui	KAD	PEV	ITR	GE	VA	EA	RVN	DDV	VILV	GAGIS	TGED	V	Y	AA	LR	LT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I		
Haloarcula hispanica	KAD	PEV	ITR	GE	VA	EA	RVN	DDV	VILV	GAGIS	TGED	V	Y	AA	LR	LT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I		
Haloarcula lacusprofundi	TAD	PEV	ITR	GE	VA	EA	RVN	PAV	VILV	GAGIS	TGED	V	Y	AA	LR	LT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I		
Halomicrobium mukohataei	KAD	PEV	ITR	GE	VA	EA	RVN	ESV	VILV	GAGIS	TGED	V	Y	AA	LR	LT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I		
Haloterrigena curkmenica	QAD	PEV	ITR	GE	VA	EA	RVN	SEV	VILV	GAGIS	TGED	V	Y	AA	LR	LT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I		
Halalkalicoccus jectabilis	QAD	PEV	ITR	GE	VA	EA	RVN	DSV	VILV	GAGIS	TGED	V	Y	AA	LR	LT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I		
Halogeometricum borinquense	TAD	PEV	ITR	GE	VA	EA	RVN	ESV	VILV	GAGIS	TGED	V	Y	AA	LR	LT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I		
Haladaptatus paucihalophi	KAD	PEV	ITR	GE	VA	EA	RVN	EDV	VILV	GAGIS	TGED	V	Y	AA	LR	LT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I		
Halopiger xanaduensis	QAD	PEV	ITR	GE	VA	EA	RVN	DDV	VILV	GAGIS	TGED	V	Y	AA	LR	LT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I		
Natronobacterium gregoryi	QAD	PEV	ITR	GE	VA	EA	RVN	DDV	VILV	GAGIS	TGED	V	Y	AA	LR	LT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I		
Natronococcus occultus	QAD	PEV	ITR	GE	VA	EA	RVN	DDV	VILV	GAGIS	TGED	V	Y	AA	LR	LT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I		
Haloferax mediterraneus	TAD	PEV	ITR	GE	VA	EA	RVN	DDV	VILV	GAGIS	TGED	V	Y	AA	LR	LT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I		
Haloferax volcanii	QAD	PEV	ITR	GE	VA	EA	RVN	DDV	VILV	GAGIS	TGED	V	Y	AA	LR	LT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I		
Natronobacterium tibetense	QAD	PEV	ITR	GE	VA	EA	RVN	DDV	VILV	GAGIS	TGED	V	Y	AA	LR	LT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I		
Natronella versiforme	QAD	PEV	ITR	GE	VA	EA	RVN	DDV	VILV	GAGIS	TGED	V	Y	AA	LR	LT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I		
Natrialba magadii	QAD	PEV	ITR	GE	VA	EA	RVN	DDV	VILV	GAGIS	TGED	V	Y	AA	LR	LT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I		
Haloarubrum kokorui	QAD	PEV	ITR	GE	VA	EA	RVN	DDV	VILV	GAGIS	TGED	V	Y	AA	LR	LT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I		
Natrialba chaharsensis	QAD	PEV	ITR	GE	VA	EA	RVN	DDV	VILV	GAGIS	TGED	V	Y	AA	LR	LT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I		
Natronococcus amlyolyticus	QAD	PEV	ITR	GE	VA	EA	RVN	DDV	VILV	GAGIS	TGED	V	Y	AA	LR	LT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I		
Natrinella altunense	QAD	PEV	ITR	GE	VA	EA	RVN	DDV	VILV	GAGIS	TGED	V	Y	AA	LR	LT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I		
Haloferax alexandrinus	TAD	PEV	ITR	GE	VA	EA	RVN	DDV	VILV	GAGIS	TGED	V	Y	AA	LR	LT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I		
Haloarubrum aidingense	TAD	PEV	ITR	GE	VA	EA	RVN	DDV	VILV	GAGIS	TGED	V	Y	AA	LR	LT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I		
Haloterrigena thermotolerans	QAD	PEV	ITR	GE	VA	EA	RVN	DDV	VILV	GAGIS	TGED	V	Y	AA	LR	LT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I		
Halococcus moribundus	QAD	PEV	ITR	GE	VA	EA	RVN	DDV	VILV	GAGIS	TGED	V	Y	AA	LR	LT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I		
Natronolimobius innermon	QAD	PEV	ITR	GE	VA	EA	RVN	DDV	VILV	GAGIS	TGED	V	Y	AA	LR	LT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I		
Haloquadratum walsbyi	TAD	PEV	ITR	GE	VA	EA	RVN	DDV	VILV	GAGIS	TGED	V	Y	AA	LR	LT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I		
Haloferax volcanii	TAD	PEV	ITR	GE	VA	EA	RVN	DDV	VILV	GAGIS	TGED	V	Y	AA	LR	LT	I	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I		
Methanosarcina acetivorans	KAD	PEV	ITR	GE	VA	EA	RVN	SGV	VILV	GAGIS	KGED	RA	LDL	GSE	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I	I				
Methanosarcina mazei	KAD	PEV	ITR	GE	VA	EA	RVN	SGV	VILV	GAGIS	KGED	RA	LDL	GSE	GVL	LAG	SVV	KA	NVE	EA	REL	IKF	I	I				
Methanosarcina barkeri	KAD	PEV	ITR	GE	VA																							





Nitrosopumilus\_sp\_H8 .....RH[P]DLIITKAA[AAV]...N[FG]...NK[TR]LL[CGAGITSGEDV]ASAVRLG[SR]GILVASGIVK[SGNWEKAI]SEFA[SSM]..V.....  
Halarchaeum\_sp\_CBA1220 .....SAD[P]DIVTDAV[AAAE]..AVD...DAV[DV]V[CGAGVSTGDD]TAADDL[GAS]GVLLASGVAKA[DDP]RAALE[DLV]APL...R.....  
Halorubrum\_sp\_CBA1229 .....TAD[P]GIVEDAV[AAAE]..AVD...PAV[DV]V[CGAGVSTGDD]TAADDL[GAS]GVLLASGVAKA[DDP]RAALE[DLV]SGL.....  
Methanohalophilus\_halophi .....QAD[P]DVVKG[SV]EAVK..RIN...PDV[VL]C[AGISRGEDL]TAAMD[LGSV]GVLLASGIVKAE[DF]RSAL[EDLV]SP.....  
Methanohalophilus\_pottuca .....QAD[P]DVVKG[SV]EAVK..RIN...PDV[VL]C[AGISRGEDL]TAAMD[LGSV]GVLLASGIVKAE[DF]RSAL[EDLV]SQ.....  
Halalkalicoccus\_sp\_GSM28 .....QAD[P]DVVTDV[AAAE]..AVD...DSV[EL]C[AGISTGED]TAEEEL[GAE]GVLLASGVAKA[DDP]QAA[EDLI]KPL.....  
Halorubrum\_sp\_ZY8 .....TAD[P]GIVEDAV[AAAE]..DVD...PAV[DV]V[CGAGVSTGDD]TAADDL[GAS]GVLLASGVAKA[DDP]RAALE[DLV]SGL.....  
Ca\_Woesearchaeota\_archaeo .....KAR[P]DLIEK[SV]KRA...G[V]S[V]L[CGAGVHTEKDV]SAA[KL]GAR[GVLLASGVAKA]RNP[E]QAL[QSL]ARGE...HKARE.....  
Haloplanus\_aerogenes .....TAD[P]GIVEGAV[AAAS]..AVD...PDV[DV]V[CGAGVSTGED]LAAADL[GAE]GVLLASGVAKA[DDP]RAALE[DLV]SGV.....  
Halostella\_sp\_IT12 .....KAD[P]DIVSDAV[AAAG]..AVD...EDV[DV]V[CGAGISTGED]LVAAREL[GAD]GVLLASGVAKA[DDP]RAALE[DLV]EP.....  
Halobellus\_sp\_Atliit-31R .....TAD[P]GIVEDAV[AAAG]..SVD...DDV[DV]V[CGAGISSGDDV]TAAREL[GAS]GILLASGVAKA[DDP]RAALE[ESLV]EP.....  
Ca\_Korarchaeota\_archaeon .....KAR[P]VAVTR[TV]NAIR..SKF...EDV[AV]I[CGAGVSTGEDV]RRAPEL[GVD]AVLLASAVAKA[DF]RPEK[LL]EL[EL]ARGA...LAPED.....  
Ca\_Thorarchaeota\_archaeon .....QA[P]EIIITGTVEKIR..AVN...PDV[TL]C[AGISTGEDV]RAALDL[GAG]GVLLASAVAKS[DF]PVE[VL]K[SLV]EP...RDC.....  
Ca\_Lokiarchaeota\_archaeon EIKTATPAQ[AE]DHNFR[ISWLNK]KFGEEYAKNIP[LY]G[SGMKPANCK]GLELQ[NIN]GGLIGGASLSA[E]YFIFIVKSAI[EL]L...DF.....  
Ca\_Heimdallarchaeota\_arch Ca\_Altiarchaeales\_archaeo .....KAR[P]EITK[SV]KIE..KFN...PAV[AP]I[CGAGITNSQDV]KALEL[GT]K[GILLASGVAKA]TDPK[IL]EKMA[QV]L...LKFQEK.....  
Ca\_Aenigmarchaeota\_archae .....ANAP[ESV]KGFSDMVDSPNRE..GKEV[P]L[CGAGISTGDDV]RKSITEL[GMH]GVLLASAITKAENP[ESV]R[EP]VA.....  
Halonotius\_pteroides .....TAD[P]DIVTDV[AAAS]..GVD...EDV[DV]V[CGAGISTGDDV]TAEEEL[GAS]GILLASGVAKA[DDP]RAALE[DLV]SGL.....  
Halonotius\_sp\_F13-13 .....TAD[P]DIVTDV[AAAS]..GVD...EDV[DV]V[CGAGISTGDDV]TAEEEL[GAS]GILLASGVAKA[DDP]RAALE[DLV]SGL.....  
Halopiger\_aswanensis .....QAD[P]DIVEDAV[AAAE]..NVD...DDV[DL]C[AGISTGDDV]VAAGDL[GAE]GVLLASGVAKA[DDP]RAALE[DLV]EP.....  
Ca\_Nanoclepta\_minutus .....KAK[P]EIVKNSV[KSV]K..SVN...EKV[KV]L[VGAGISNKEDV]KALEL[GAD]GVLLASAVTKA[N]FYEKI[RELA]EGF...FYK.....  
Natronolimmobius\_sp\_AArc1 .....QAD[P]EVEG[V]DAA[R]..NVD...EDV[AV]L[CGAGISTGDDV]VAADDL[GSE]GVLLASGVAKA[DDP]KAA[EDLV]DP.....  
Methanobrevibacter\_woesei .....QAQ[P]EVEEDT[KG]V..EHD...SS[KV]L[CGAGISNGDDM]KAIDL[GAE]GVLLASGITKAENP[KD]AL[EDLV]SK.....  
Methanothermobacter\_deflu .....KAE[P]EVI[TS]G[V]DAV[K]..KVN...PEV[S]L[CGAGISTGDDM]KAAVDL[GAE]GVLLASGITL[LD]S[PRD]AL[LDLV]SKV.....  
Halorussus\_rarus .....KAD[P]DVVTDV[AAAE]..AVD...DSV[EV]V[CGAGISTGED]LTAAREL[GAE]GVLLASGVAKA[DDP]KAA[ADLV]APL.....  
Halopelagius\_longus .....TAD[P]DIVEDAV[AAAE]..SVD...EDV[EV]V[CGAGISSGDD]EAAAGDL[GAT]GILLASGVAKA[DDP]RAALE[DLV]EP.....  
Haloplanus\_salinus .....TAD[P]GIVEGAV[AAAG]..AVD...PEV[DV]V[CGAGVSTGDD]LAAADL[GAA]GVLLASGVAKA[DDP]RAALE[DLV]SGA.....  
Methanothermobacter\_teneb .....KAE[P]EVI[TS]G[V]DAV[K]..MNN...PNV[KV]L[CGAGISTGED]LFAALDL[GSE]GVLLASGVILADN[PRE]AL[DLV]SK.....  
Acidianus\_brixleyi .....KAK[P]EVI[TS]G[V]EEN..KMK...GVD[L]I[AGAGISSGDDV]TAIKL[GAG]GIGVASAVMK[SN]E[QV]E[DF]KSA...TKAMSEK.....  
Metallosphaera\_hakonensis .....KAK[P]EVI[TS]G[V]EEN..KMK...GVD[L]I[AGAGISSGDDV]TAIKL[GAG]GIGVASAVMK[SN]E[QV]E[DF]KSA...TKAMSEK.....  
Acidianus\_sulfidivorans\_J .....KAK[P]EVI[TS]G[V]EEN..KMK...GVD[L]I[AGAGISSGDDV]TAIKL[GAG]GIGVASAVMK[SN]E[QV]E[DF]KSA...TKAMSEK.....  
Methanospirillum\_stamsii .....EAN[P]GIITGSVD[IV]K..KIN...PAV[KV]L[TAGIHSKCV]KTAIDL[GT]DGVLLASVVKAE[DP]AAV[LDV]SL.....  
Methanospirillum\_lacunae .....EAN[P]EIIITGSV[QAA]L..KAN...PDV[KV]L[TAGIHSKCV]KTAIDL[GT]DGVLLASVVKAE[DP]AAV[LDV]SL.....  
Methanospira\_cuniculi .....OC[P]EIVSNT[AAVH]..EIN...DS[PV]L[CGAGITNGDDV]KAALDL[GAE]GVLLASGVAKA[ED]SKAV[EDLV]SK.....  
Methanobrevibacter\_thauer .....QAQ[P]EVEEDT[KG]V..KIN...KN[KV]L[CGAGITGDDM]KAAMD[LGAD]GVLLASGITKAENP[KD]AL[EDLV]SK.....  
Ca\_Nanosidianus\_stetteri .....TAK[P]EIVINS[KM]I..K..EVN...PNV[KV]L[VGAGISNKEDV]KKSLEY[GAE]GVLLASAVTKA[N]FYEKI[RELA]EGF...IR.....  
Sulfolobus\_islandicus .....KAK[P]EVI[TS]G[V]EEN..KMK...GVD[L]I[AGAGISSGDDV]TAIKL[GAG]GIGVASAVMK[SN]E[QV]E[DF]KSA...TKAMSEK.....  
consensus>70 .....a.peiv...v...vn...v.vl.Gagi.ged...a.elg..gvllasgv.ka.dp...div.....