



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS
ESCUELA DE PREGRADO

“Caracterización bioquímica de la 4-amino-5-
hidroximetil-2-metilpirimidina quinasa de
Salmonella typhimurium y *Thermus
thermophilus*”

Seminario de Título entregado a la Universidad de Chile en cumplimiento
parcial de los requisitos para optar al Título de Ingeniero en Biotecnología
Molecular.

PABLO ANTONIO CEA MEDINA

Director del seminario de título:
Dr. Víctor Castro-Fernández

Co-director del seminario
Dra. Victoria Guixé Leguía

Enero, 2019
Santiago - Chile



INFORME DE APROBACIÓN SEMINARIO DE TÍTULO

Se informa a la Escuela de Pregrado de la Facultad de Ciencias, de la Universidad de Chile que el Seminario de Título, presentado por el **Sr Pablo Antonio Cea Medina**

“Caracterización bioquímica de la 4-amino-5-hidroximetil-2-metilpirimidina quinasa de *Salmonella typhimurium* y *Thermus thermophilus*”

Ha sido aprobado por la Comisión de Evaluación, en cumplimiento parcial de los requisitos para optar al Título de Ingeniero en Biotecnología Molecular.

Dr. Víctor Castro-Fernández

Director Seminario de Título:

Dra. Victoria Guixé Leguía

Co-directora Seminario de título

Comisión Revisora y Evaluadora

Presidente Comisión

Evaluador

Agradecimientos

Primero que todo, debo agradecer a mi familia, en especial a mi madre y mi padre por su apoyo incondicional durante toda mi carrera.

Agradezco a la Dra. Victoria Guixé por haberme recibido en el Laboratorio de Bioquímica y Biología Molecular, por su guía certera, su paciencia y por sobre todo, por darme la oportunidad de enfrentar todos mis desafíos científicos con total libertad, amparado sólo en su confianza. También le debo un sincero y especial agradecimiento a Profesor Tutor Dr. Víctor “El Toro” Castro-Fernández, quien ha sido un mentor excelente y un gran amigo durante todo mi proceso formativo.

Debo dar las gracias también al Dr. Jans Alzate Morales y Rodrigo Recabarren, del Centro de Bioinformática y Simulación Molecular de la Universidad de Talca, quienes me recibieron amablemente e hicieron posible incorporar la metodología que más me apasiona en este trabajo.

Las páginas en este documento se harían cortas para expresar todo lo que realmente le debo a mis compañeros del Mandrilab, quienes a pesar de mi relativa juventud, me brindaron la más cálida bienvenida y me hicieron sentir parte de su grupo y de su historia desde el día uno. Gracias a Pablo Villalobos, Exequiel Medina y Alonso Carvajal del grupo del Dr. Jorge Babul, quienes siempre estuvieron dispuestos a brindar su apoyo y conocimientos de biofísica y biología computacional. A Pablo Maturana y Juan Carlos Cáceres del grupo del Dr. Ricardo Cabrera por las discusiones enriquecedoras de mis resultados. A mis compañeros de grupo; Dr(c) Gabriel Vallejos por su apoyo en experimentos de cinética enzimática y HPLC. A Gissela Araya por su ayuda en los experimentos de especificidad por sustrato. A Felipe “Felipín” González, quien fuese mi primer guía y me dio la bienvenida al mundo de las proteínas. Y con especial cariño, agradezco a Alejandra Herrera y Ricardo Zamora, quienes han sido mis hermanos mayores

Gracias a todos por estas amistades fraguadas al calor de las pipetas, licores, alegrías y desazones. Gracias por esas largas conversaciones de ciencia, fútbol y amores. Y por sobretodo, gracias por darme el coraje de seguir con la insensatez que es hacer ciencia en este país.

Índice temático

Agradecimientos	ii
Índice temático	iii
Índice de tablas	vi
Índice de figuras	vii
Lista de abreviaturas	ix
Resumen	xi
Abstract	xiv
1. Introducción	1
1.1 Tiamina y su biosíntesis	1
1.2 Enzimas bifuncionales HMPK/PLK	4
1.3 Enzimas con actividad HMPK.	5
1.4 Objetivo general y objetivos específicos	11
2. Materiales y Métodos	12
2.1 Manejo de cepas y vectores	12
2.1.1 Cepas	12
2.1.2 Vectores	12
2.1.3 Medios de cultivo	13
2.1.4 Transformación de <i>E. coli</i> quimiocompetentes	13
2.1.5 Aislamiento de plásmidos	13
2.2 Expresión y purificación de proteínas	14
2.2.1 Purificación de StHMPK y TtHMPK	14

2.2.2 Electroforesis denaturante de proteínas en geles de poliacrilamida (SDS-PAGE)	15
2.3 Estequiometría de la reacción catalizada por <i>St</i> HMPK y <i>Tt</i> HMPK	15
2.4 Separación de productos de reacción mediante HPLC	16
2.5 Especificidad por sustrato de <i>St</i> HMPK y <i>Tt</i> HMPK	17
2.6 Reconstrucción filogenética entre HMPKs y PLKs	17
2.7 Dinámica molecular del complejo ternario ATP-HMP de <i>St</i> HMPK	18
2.8 Determinación de las constantes cinéticas de <i>St</i> HMPK Y <i>Tt</i> HMPK	20
2.9 Cromatografía de exclusión molecular	21
2.10 Determinación de la estabilidad de <i>St</i> HMPK y <i>Tt</i> HMPK	22
2.11 Comparación estructural entre <i>St</i> HMPK y <i>Tt</i> HMPK	24
2.11.1 Dinámica molecular de <i>St</i> HMPK y <i>Tt</i> HMPK	24
2.11.2 Análisis de las trayectorias	25
2.11.3 Composición aminoacídica del núcleo	27
2.11.4 Análisis de interacciones no covalente	27
2.12 Anexo	28
2.12.1 Secuencia de <i>St</i> HMPK	28
2.12.2 Secuencia de <i>Tt</i> HMPK	28
3. Resultados	29
3.1 Expresión heteróloga y purificación de <i>St</i> HMPK y <i>Tt</i> HMPK	29
3.2 Análisis de la fosforilación doble de HMP por <i>St</i> HMPK y <i>Tt</i> HMPK	30
3.3 Especificidad de sustrato en HMPKs	34
3.4 Caracterización cinéticas de <i>St</i> HMPK y <i>Tt</i> HMPK	44

3.5 Caracterización biofísica de <i>S</i> tHMPK y <i>T</i> tHMPK	46
3.6 Análisis estructural de <i>S</i> tHMPK y <i>T</i> tHMPK	51
4. Discusión	62
4.1 Doble fosforilación consecutiva y especificidad por sustrato	62
4.2 Biofísica y biología estructural de las HMPKs	67
5. Conclusiones	71
6. Referencias	73

Índice de tablas

Tabla I. Estequiometría del consumo de NADH en <i>S</i> tHMPK y TtHMPK	32
Tabla II. Parámetros cinéticos para HMP de <i>S</i> tHMPK y <i>T</i> tHMPK	45
Tabla III. Tiempos de retención y masas moleculares estimadas de <i>S</i> tHMPK y <i>T</i> tHMPK mediante filtración molecular.	47
Tabla IV. Valores de T_m para <i>S</i> tHMPK y <i>T</i> tHMPK	49
Tabla V. Valores de C_m para <i>S</i> tHMPK y <i>T</i> tHMPK	51
Tabla VI. Interacciones por puentes de hidrógeno en StHMPK y TtHMPK.	61

Índice de figuras

Figura 1. Estructuras químicas de la tiamina	1
Figura 2. Resumen de la vía metabólica de biosíntesis bacteriana de tiamina	2
Figura 3. Esquema de la reacción catalizada por la HMPK	6
Figura 4. Estructura de las HMPK	7
Figura 5. Mecanismo químico de la conversión de HMP a HMP-P	9
Figura 6. Cromatogramas representativos de las purificaciones	29
Figura 7. SDS-PAGE de la purificación de <i>St</i> HMPK y <i>Tt</i> HMPK	30
Figura 8. Curvas de progreso representativas de la titulación de NADH para <i>St</i> HMPK y <i>Tt</i> HMPK	31
Figura 9. Análisis de los productos de la reacción catalizada por <i>St</i> HMPK mediante cromatografía de intercambio aniónico	33
Figura 10. Especificidad por sustrato en <i>St</i> HMPK y <i>Tt</i> HMPK	35
Figura 11. Árbol filogenético de máxima verosimilitud para secuencias de proteínas PLK y HMPK	37
Figura 12. La simulación del complejo ternario de <i>St</i> HMPK es estable durante la ventana de tiempo estudiada y tiene coherencia con la reacción catalizada	39
Figura 13. Diagrama 2D de los residuos posicionados a 5 Å del sustrato aceptor del grupo fosfato en <i>St</i> HMPK y <i>Sa</i> PLK	42
Figura 14. Alineamiento de secuencias múltiples entre ThiD, PdxK y PdxY	43
Figura 15. Curvas de saturación de HMP para <i>St</i> HMPK y <i>Tt</i> HMPK	45
Figura 16. Curvas de saturación para ATP de <i>St</i> HMPK y <i>Tt</i> HMPK	46

Figura 17. Calibración de la columna de exclusión molecular	47
Figura 18. Estabilidad térmica de <i>St</i> HMPK y <i>Tt</i> HMPK determinadas por CD	48
Figura 19. Desplegamiento de <i>St</i> HMPK y <i>Tt</i> HMPK frente a GndCl monitoreada por fluorescencia intrínseca de triptófanos	50
Figura 20. RMSD y RMSF, para <i>St</i> HMPK y <i>Tt</i> HMPK resultantes de 50 ns de simulación de dinámica molecular	53
Figura 21. Diferencias en las fluctuaciones estructurales de <i>St</i> HMPK y <i>Tt</i> HMPK, mapeadas en las estructuras de mínima energía	54
Figura 22. Matrices de correlación cruzada de residuos de <i>St</i> HMPK y <i>Tt</i> HMPK	56
Figura 23. Diferencias en el núcleo hidrofóbico de <i>St</i> HMPK y <i>Tt</i> HMK	58
Figura 24. Diferencias en las interacciones electrostáticas entre <i>St</i> HMPK y <i>Tt</i> HMK	60

Lista de abreviaturas

ADP	Adenosín difosfato
ATP	Adenosín trifosfato
C α	Carbono alfa
CD	Dicroísmo circular
Cm	Concentración media de desplegamiento por guanidinio
GndCl	Cloruro de guanidinio
HEPES	Acido 4-(2-hidroxietil)-1-piperazinetanosulfónico
HMP	4-amino-5-hidroximetil-2-metilpirimidina
HMP-P	4-amino-5-hidroximetil-2-metilpirimidina fosfato
HMP-PP	4-amino-5-hidroximetil-2-metilpirimidina pirofosfato
HMPK	4-amino-5-hidroximetil-2-metilpirimidina quinasa
HMPK/PLK	Enzima piridoxal quinasa bifuncional que cataliza con HMP y PL
IPTG	Isopropil β -D-1-tiogalactopiranosido
kDa	Kilo Dalton
K _M	Constante de Michaelis-Menten
k_{cat}	Constante catalítica
LB	Luria-Bertrani
LDH	Lactato deshidrogenasa
NAD ⁺	Nicotín adenin dinucleótido oxidado
NADH	Nicotín adenin dinucleótido reducida
NPT	Ensamble con temperatura, presión y N° de partículas constantes.
NVT	Ensamble con temperatura, volumen y N° de partículas constantes.
PAGE	Electroforesis en gel de poliacrilamida
<i>pdxK</i>	Gen codificante de la piridoxal quinasa
PdxK	Producto del gen <i>pdxK</i>
PEP	Fosfoenolpiruvato
PK	Piruvato quinasa

PL	Piridoxal
PLK	Enzima con actividad piridoxal quinasa
PM	Piridoxamina
PMSF	Fluoruro de fenilmetilsulfonilo
PN	Piridoxina
RMSD	Desviación cuadrática media
RMSF	Fluctuación cuadrática media
<i>Sa</i> PLK	Piridoxal quinasa de <i>Staphylococcus aureus</i>
SDS	Dodecilsulfato de sodio
<i>St</i> HMPK	HMPK de <i>Salmonella typhimurium</i>
<i>thiD</i>	Gen codificante de la HMPK
ThiD	Producto del gen <i>thiD</i>
TRIS	Tris(hidroximetil)aminometano
<i>Tt</i> HMPK	HMPK de <i>Thermus thermophilus</i>

Resumen

La 4-amino-5-hidroximetil-2-metilpirimidina quinasa (HMPK, EC 2.7.1.49) es una enzima perteneciente a la superfamilia riboquinasa y participa en la biosíntesis de tiamina (vitamina B₁) en bacterias. Se ha descrito que esta enzima es capaz de catalizar dos fosforilaciones consecutivas dependientes de ATP altamente específicas sobre el sustrato hidroximetil pirimidina (HMP), generando como producto hidroximetil pirimidina pirofosfato. Esto contrasta notablemente con lo que se ha observado en las piridoxal quinazas de bacterias Gram positivas (HMPK/PLK, EC 2.7.1.35), un grupo de enzimas homólogas cercanas capaces de fosforilar hidroximetil pirimidina, piridoxal, piridoxina y piridoxamina, pero incapaces de catalizar dos fosforilaciones consecutivas, por lo que sólo producen hidroximetil pirimidina fosfato. Las HMPKs no han sido estudiadas tan exhaustivamente como las HMPK/PLKs y sólo hay dos caracterizaciones breves disponibles en la literatura; la de la HMPK de *Escherichia coli* y la de *Bacillus subtilis*. Por lo tanto, aún no se conoce si las propiedades observadas en las enzimas descritas son ubicuas para linajes bacterianos distintos, especialmente aquellos filogenéticamente distantes y que han sido sometido a presiones selectivas fuertes, como los extremófilos. Por esta razón, en este trabajo se realizó la caracterización bioquímica de la HMPK de la enterobacteria *Salmonella typhimurium* (StHMPK) y de la bacteria termófila *Thermus thermophilus* (TtHMPK).

A través de experimentos de estequiometría de reacción y análisis de generación de productos, se demostró que ambas enzimas son capaces de catalizar dos

fosforilaciones consecutivas. Experimentos de especificidad de sustrato revelaron que ambas enzimas son altamente específicas por hidroximetil pirimidina. Análisis filogenéticos mostraron que estas enzimas están estrechamente relacionadas con las HMPKs/PLK de organismos gram positivos, y estas últimas parecen ser descendientes directos de las HMPKs. Por lo tanto, para estudiar cómo estos grupos de enzimas han divergido en términos de sus actividades catalíticas, se realizaron simulaciones de dinámica molecular del complejo ternario (Mg·ATP - HMP) de *StHMPK*, para analizar el sitio de unión a sustrato y compararlo con el de la HMPK/PLK de *Staphylococcus aureus* (*SaPLK*). Los resultados mostraron que existe un alto grado de conservación entre ambos sitios, existiendo sólo unas pocas diferencias que podrían explicar la divergencia funcional observada, principalmente la presencia de una treonina adyacente a la base catalítica en *StHMPK*, que es reemplazada por una alanina en *SaPLK*, y la presencia de una glutamina en *StHMPK* que forma puentes de hidrógeno con el HMP.

La caracterización cinética de *StHMPK* y *TtHMPK* mostró que ambas enzimas poseen una K_M similar para HMP (cerca a $30\mu\text{M}$) y que la V_{max} para *TtHMPK* es un orden de magnitud menor que para *StHMPK* a $37\text{ }^\circ\text{C}$. Sin embargo, estos parámetros fueron obtenidos para las curvas de saturación de HMP, las cuales mostraban un comportamiento del tipo Michaelis-Menten, mientras que las curvas de saturación para ATP mostraron una clara desviación de este modelo y por lo tanto, no se pudieron determinar parámetros cinéticos.

Finalmente, se realizó una caracterización estructural y biofísica para evaluar diferencias de estabilidad. Ambas enzimas parecen ser monómeros en las condiciones estudiadas, a diferencia de lo reportado para la enzima de *E. coli* que forma un

tetrámero. Experimentos de desplegamiento por temperatura y agentes químicos mostraron que *TtHMPK* es significativamente más estable que *StHMPK*. Las bases estructurales de estas diferencias fueron analizadas mediante simulaciones de dinámica molecular, las que revelaron que la proteína termoestable es más rígida, tiene un menor contenido de residuos polares en el núcleo y tiene mayor cantidad de interacciones electrostáticas que su homólogo mesoestable.

Abstract

4-amino-5-hydroxymethyl-2-methylpyrimidine kinase (HMPK, EC 2.7.1.49) is a bacterial enzyme that belongs to the ribokinase superfamily and participates in the thiamine (vitamine B₁) biosynthetic pathway. It has been described that this enzyme is capable to catalyze two consecutive highly specific ATP dependent phosphorylations on the substrate hydroxymethyl pyrimidine, yielding hydroxymethyl pyrimidine pyrophosphate. This contrast notoriously with what has been observed for the closely related homologous enzymes pyridoxal kinases from Gram positive bacteria (HMPK/PLK, EC 2.7.1.35), which can phosphorylate hydroxymethyl pyrimidine, pyridoxal, pyridoxine and pyridoxamine, but are unable to catalyze two consecutive phosphorylations, thus only produce hydroxymethyl pyrimidine phosphate. HMPKs have not been as extensively studied as HMPKs/PLK, and only two brief biochemical characterizations are available on the literature; the characterization of the HMPK from *Escherichia coli* and from *Bacillus subtilis*. Therefore, it is still unknown whether the properties observed in the described enzymes are ubiquitous among different bacterial lineages, especially those that come from a very distinct phylogenetic background and have been subject to strong selective pressures, as the enzymes from extremophilic organisms. For this reason, in this work we address the biochemical characterization of the HMPK from the enterobacteria *Salmonella typhimurium* (StHMPK) and the thermophilic bacteria *Thermus thermophilus* (TtHMPK).

Through stoichiometric experiments and product generation analysis, it was established that both enzymes are able to perform two consecutive phosphorylations. Substrate specificity experiments revealed that both enzymes are highly specific for hydroxymethyl pyrimidine. Phylogenetic analysis of these enzymes showed that are closely related to HMPKs/PLK from Gram positive organisms, being the later a direct descendant from HMPKs. Therefore, to study how these two groups of enzymes have diverged so much in terms of their catalytic activities, we analysed the substrate binding site of *St*HMPK by molecular dynamics simulations of the ternary complex (Mg•ATP - HMP) and compared it to the binding site of the PLK from *Staphylococcus aureus* (*Sa*PLK). The results showed that there is an overall great conservation among the active sites, with just a few differences that could be responsible for the functional divergences observed, mainly the presence of a threonine residue adjacent to the catalytic base in *St*HMPK which is replaced by an alanine in *Sa*PLK, and the presence of a glutamine that forms hydrogen bonds with the HMP in *St*HMPK.

Kinetic characterization of *St*HMPK and *Tt*HMPK showed that both enzymes have a similar K_M for HMP (around 30 μ M) while the V_{max} for *Tt*HMPK is one order of magnitude lower than the V_{max} for *St*HMPK. However, these parameters were obtained only for HMP saturation curves, which showed a Michaelis-Menten behaviour, whereas ATP saturation curves displayed a clear deviation from a Michaelis-Menten model and therefore, no kinetic parameters could be deduced from these experiments.

Finally, a biophysical and structural characterization to assess stability differences was performed. Both enzymes seem to be in monomeric state under the conditions assayed, in contrast with what was reported for the enzyme from *E. coli*,

which forms a tetramer. Thermal and chemical unfolding experiments showed that *Tt*HMPK is significantly more stable than *St*HMPK. The structural basis for these differences were investigated through molecular dynamics simulations, which revealed that the thermostable protein is more rigid, has a reduced content of polar amino acids in its core, and has more electrostatic interactions than its mesostable homologous.

1. Introducción

1.1 Tiamina y su biosíntesis

La tiamina o vitamina B₁, es un cofactor enzimático que participa principalmente en procesos oxidativos del metabolismo energético y está presente en organismos de los 3 dominios de la vida. La tiamina está constituida por un anillo de metilpirimidina y un anillo tiazólico (Figura 1A). Esta estructura corresponde a una forma sólo de transporte no metabólicamente activa del cofactor. Las formas metabólicamente activas corresponden a la tiamina pirofosfato (Figura 1B) y la tiamina trifosfato (Figura 1C), siendo la tiamina pirofosfato la forma más caracterizada a nivel funcional, la cual participa como estabilizador de grupos acil-carbaniones en reacciones que forman parte del metabolismo oxidativo de hidratos de carbono y en el metabolismo de aminoácidos de cadena ramificada (Schowen, 1998).

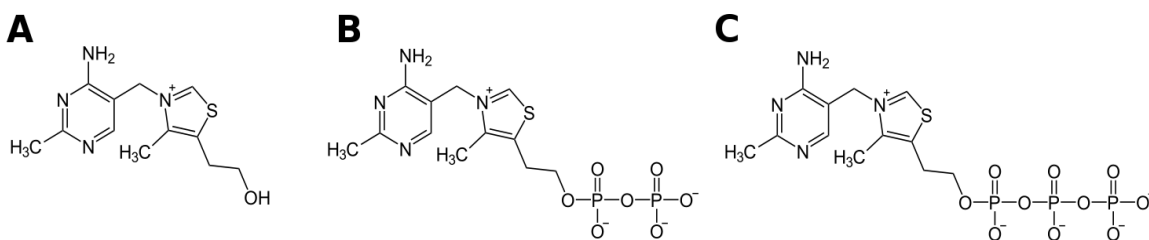


Figura 1. Estructuras químicas de la tiamina. A) Tiamina. B) Tiamina pirofosfato. C) Tiamina trifosfato.

metilpirimidina pirofosfato (HMP-PP) (Figura 2). Los genes implicados en la biosíntesis de tiamina en bacterias están organizados en dos operones denominados *thiCEFSGH* y *thiMD*, además de 4 loci independientes (*thiI*, *thiL*, *thiK* y *dsx*) (Miranda-Ríos y cols., 2002). Además, se ha demostrado que algunos de estos genes son regulados a nivel traduccional, a través del bloqueo del sitio de unión a ribosoma mediante *riboswitches* dependientes de tiamina (Winkler y cols., 2002).

Interesantemente, dentro de esta ruta biosintética, participan múltiples enzimas con estructuras muy similares y que pertenecen a la misma familia de proteínas; la familia de quinazas de vitamina dependientes de ATP. Esta familia forma parte de la superfamilia riboquinasa, donde también se agrupa la familia de quinazas de azúcares dependientes de ATP y la familia quinazas de azúcares dependientes de ADP. Las enzimas de la familia de quinazas de vitaminas que participan en la vía biosintética de tiamina son la tiazol quinasa (ThiK), la piridoxal quinasa (PdxK o PLK) y la hidroximetilpirimidina quinasa (HMPK o ThiD). La presencia de enzimas estructuralmente similares y filogenéticamente relacionadas dentro de la misma ruta biosintética sugiere un origen evolutivo común, donde eventos de duplicación y selección han generado enzimas específicas para cada paso de conversión metabólica. Sin embargo, al observar la ruta de fosforilación de la rama biosintética de la 4-amino-5-hidroximetil-2-metilpirimidina fosfato (HMP-P, compuesto n° 11 en fig. 2), queda en evidencia que existe redundancia en las enzimas capaces de realizar estos procesos. El paso de 4-amino-5-hidroximetil-2-metilpirimidina (HMP) a hidroximetil pirimidina fosfato (HMP-P), puede ser catalizado por dos productos de genes distintos, *pdxK*

(codificante de la piridoxal quinasa) y *thiD* (codificante de la hidroximetil pirimidina quinasa), respectivamente.

1.2 Enzimas bifuncionales HMPK/PLK

La primera descripción de la existencia de redundancia funcional en la biosíntesis de tiamina fue realizada en el año 2004, por Park y colaboradores (Park y cols., 2004). El objetivo inicial de este trabajo era clonar y caracterizar la proteína homóloga al gen *thiD* de *Escherichia coli* en *Bacillus subtilis*. Sin embargo, los resultados obtenidos referentes a la actividad catalítica de esta enzima mostraron que es capaz de utilizar HMP y piridoxal (PL) como aceptores de fosfato (siendo preferente por éste último), generando hidroximetilpirimidina fosfato (HMP-P) y piridoxal fosfato, respectivamente. Por lo que el producto del gen aislado de *B. subtilis* posee dos actividades metabólicamente relevantes, es decir, es una enzima bifuncional (HMPK/PLK, EC 2.7.1.35) que participa tanto en la ruta de biosíntesis de piridoxal fosfato (vitamina B₆) como de tiamina, razón por la cual debió ser recategorizado como un homólogo al gen *pxk* (gen de la piridoxal quinasa) en lugar de *thiD*. Posteriormente, otros estudios determinaron la presencia de enzimas HMPK/PLK en distintos organismos Gram positivos, como *Staphylococcus aureus* y *Clostridium difficile*, de los cuales se han determinado estructuras cristalográficas y los mecanismos químicos por los cuales transcurre el proceso catalítico (Nodwell y cols., 2014; Newman y cols., 2006).

Análisis filogenéticos basados en estructuras disponibles en el PDB de estas enzimas, muestran que son más cercanas a las hidroximetilpirimidina quinasa

codificadas en el gen *thiD* (HMPKs, EC 2.7.1.35), que a las piridoxal quinasas presentes en organismos Gram negativos y eucariontes, y que han sido las canónicamente estudiadas, como la piridoxal quinasa de *Homo sapiens* y *Escherichia coli*. Un estudio de la historia evolutiva de la bifuncionalidad en HMPKs/PLK mediante reconstrucción de secuencias ancestrales reveló que el ancestro común hipotético entre HMPKs y HMPK/PLKs era específico por HMP, por lo que el rol de esta enzima en la vía de biosíntesis de vitamina B₆ (piridoxal) correspondería a una novedad adaptativa (Castro-Fernandez y cols., 2014).

1.3. Enzimas con actividad HMPK.

A diferencia de las piridoxal quinasas, las HMPKs bacterianas no han sido estudiadas tan exhaustivamente. A la fecha, existe una única caracterización bioquímica de esta clase de enzimas, que corresponde al producto del gen *thiD* de *Escherichia coli* (Mizote y cols., 1990). En este estudio se demostró que esta enzima oligomeriza en forma de tetrámero y que es capaz de realizar dos fosforilaciones consecutivas a expensas de ATP, por lo que aparte de catalizar la conversión de HMP a HMP-P, cataliza la conversión de HMP-P a HMP-PP (Figura 3), por lo que también es una enzima con dos actividades metabólicamente relevantes: hidroximetil pirimidina quinasa y fosfometilpirimidina quinasa.

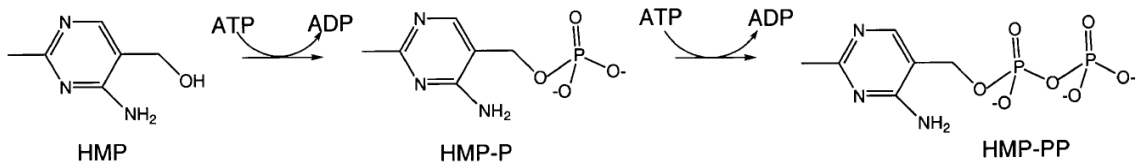


Figura 3. Esquema de la reacción catalizada por la HMPK. Extraído de Cheng y cols., 2002.

Pese a que HMPK/PLKs y HMPKs son enzimas homólogas, las propiedades catalíticas de esta última son sustancialmente distintas. Si bien las HMPK/PLKs son capaces de utilizar dos sustratos distintos como aceptores de fosforilo, el fenómeno de doble fosforilación observado en HMPKs es único dentro de la superfamilia riboquinasa, ya que en la segunda reacción se utiliza un grupo fosfato como aceptor de un nuevo grupo fosfato, en lugar de un hidroxilo como se observa comúnmente, lo cual resulta sorprendente cuando se consideran las grandes diferencias estereoquímicas y de densidad de carga que existen entre ambos grupos químicos y hasta ahora, no existe claridad acerca de los mecanismos estructurales que permiten esta segunda fosforilación. Además, se reportó que la HMPK de *E. coli* es altamente específica por HMP, y no es capaz de fosforilar piridoxal, piridoxina ni piridoxamina, pese a la alta similitud estructural entre estos sustratos.

Si bien hay un único reporte bioquímico de HMPKs (Mizote y cols., 1990), existen dos estructuras cristalográficas disponibles en el *Protein Data Bank*; la estructura de la HMPK de *Salmonella typhimurium* (1JXI y 1JXH) y la de *Thermus thermophilus* (1UB0), de las cuales sólo la primera posee una publicación de asociada (Cheng y cols.,

2002), de la que desprenden las características estructurales que se detallan a continuación.

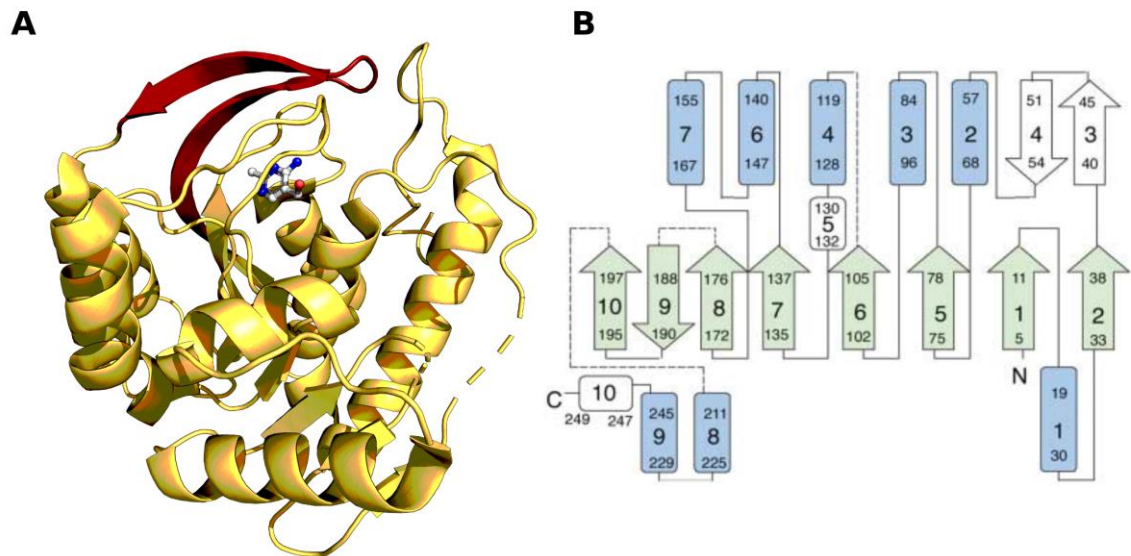


Figura 4. Estructura de las HMPK. A) Estructura cristalográfica de la HMPK de *S. typhimurium* (PDB: 1JXI) unida a HMP, en amarillo se muestra el dominio *rossmann*, en rojo el hairpin- β y en código CPK el ligando HMP. B) Diagrama topológico de la HMPK, las hebras 3 y 4 constituyen el hairpin- β ilustrado en rojo en A, las líneas punteadas representan las conexiones no resueltas a partir del mapa de densidad electrónica. Modificado de Cheng y cols., 2002

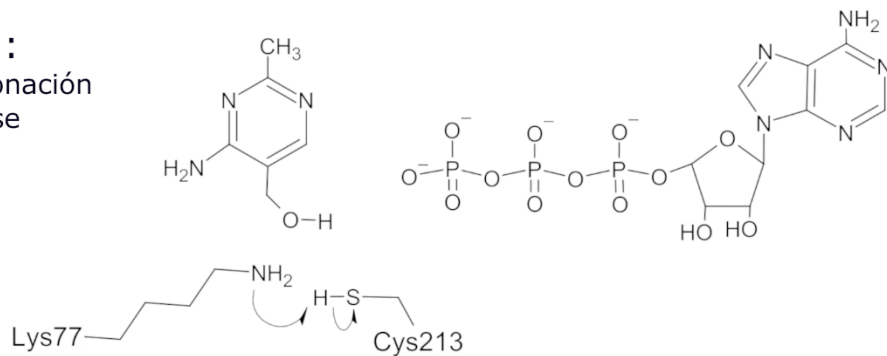
Las enzimas HMPK forman parte de la superfamilia riboquinasa. Al igual que todas las enzimas de esta superfamilia, las HMPKs poseen un dominio mayor con un plegamiento tipo *Rossmann* constituido por 8 hélices α , 2 hélices 3_{10} y 8 hebras β y un *hairpin*- β , equivalente al dominio menor de los otros miembros de esta superfamilia, que actúa como “tapa” excluyendo al solvente del sitio activo (Cheng y cols., 2002) (Figura 4A). El mecanismo químico por el cual transcurre la catálisis de la formación del primer producto es similar al de otros miembros de la superfamilia, con la diferencia de que un

residuo de cisteína actúa como base catalítica, en lugar de un residuo de aspártico (Figura 5). El primer paso consiste en la desprotonación de la base catalítica por acción de un residuo de lisina, luego la cisteína sustrae un protón desde el hidroxilo formando un ion alcóxido el cual realiza un ataque nucleofílico sobre el fosfato γ del complejo Mg-ATP, produciendo una sustitución nucleofílica de tipo SN_2 . Sin embargo, el mecanismo por el cual transcurre la segunda reacción de fosforilación y el detalle exacto de cómo se une HMP-P junto a ATP-Mg en el sitio activo aún son desconocidos.

A partir de los datos extraídos de esta estructura cristalográfica, Cheng y cols., proponen un mecanismo para explicar cómo estas enzimas podrían catalizar el proceso de doble fosforilación, el cual se basa en la reorientación del grupo fosfato del HMP-P hacia un bolsillo capaz de estabilizar aniones, pero esta hipótesis no es acompañada por datos experimentales o teóricos que puedan sustentarla.

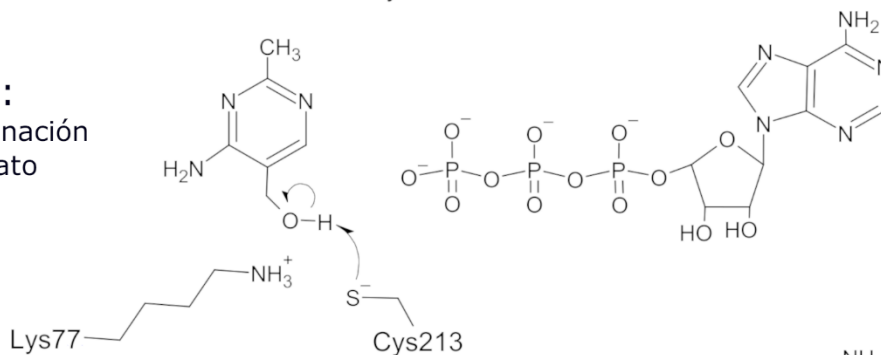
Paso 1:

Desprotonación de la base catalítica



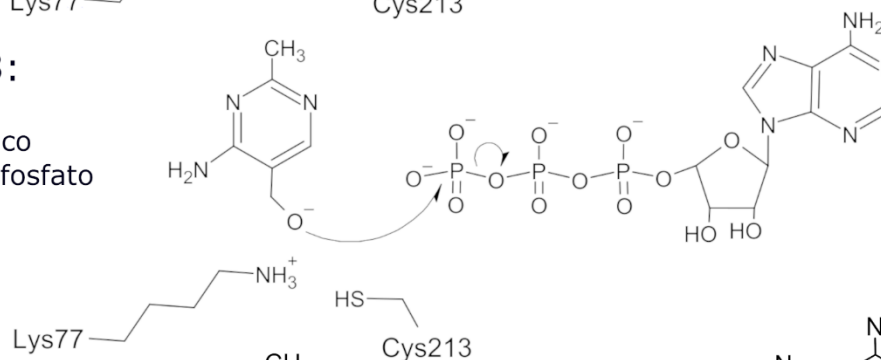
Paso 2:

Desprotonación del sustrato aceptor



Paso 3:

Ataque nucleofílico sobre el fosfato



Productos

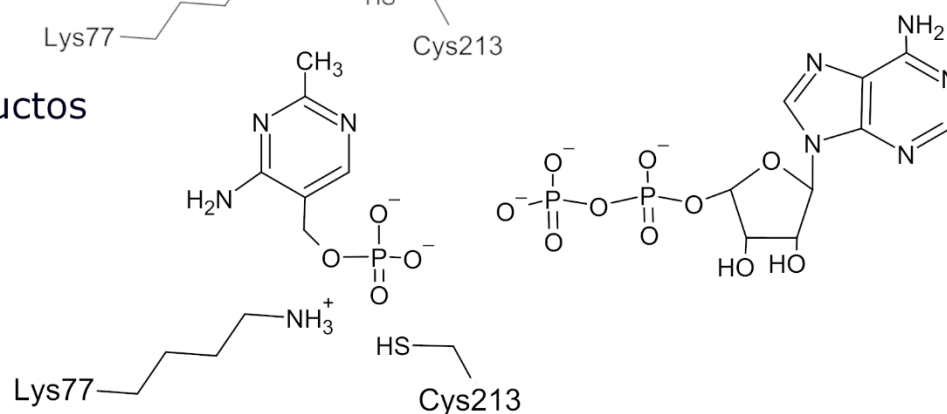


Figura 5. Mecanismo químico de la conversión de HMP a HMP-P. Se indican los residuos según la numeración en la secuencia de la HMPK de *S. typhimurium*.

Dado que a la fecha existe un único reporte donde se demuestra experimentalmente la capacidad de la HMPK de *Escherichia coli* de catalizar dos fosforilaciones consecutivas (Mizote y cols., 1990), queda abierta la interrogante acerca la ubicuidad de esta propiedad funcional en ortólogos de otras especies de bacterias. Este detalle no es trivial, sobre todo si se considera que organismos distantes evolutivamente y que han sido sometidos a presiones selectivas fuertes, como por ejemplo, temperatura o pH extremo, pueden adquirir adaptaciones que se traduzcan en diferencias significativas en el funcionamiento de su maquinaria molecular. En vista de que existen dos estructuras conocidas de HMPKs, provenientes de organismos que habitan nichos ecológicos con temperaturas sumamente distintas (*S. typhimurium* y *T. thermophilus*) y ninguna de ellas ha sido caracterizada, en este seminario de título se realizó la caracterización bioquímica de estas enzimas (*St*HMPK y *Tt*HMPK, respectivamente) en términos de sus actividades catalíticas y propiedades biofísicas. Además, mediante simulaciones de dinámica molecular, se describen las características estructurales que pudiesen estar asociadas a las diferencias de termoestabilidad que se espera para estas proteínas.

Los resultados obtenidos en este seminario podrían ser de aporte en diversas áreas. Las enzimas pertenecientes a la ruta de biosíntesis de tiamina han sido propuestas como blancos atractivos para el desarrollo de nuevos antibióticos contra *enterobacterias* patogénicas, siendo la HMPK un candidato sobresaliente debido a su ausencia total de homólogos eucariontes. Por lo que conocer las propiedades catalíticas básica de esta enzima aportaría datos para conocer que andamiaje de ligandos son capaces de unirse a la enzima, y por ende, ser usados como base para el desarrollo de inhibidores.

Por otra parte, dado que este trabajo tiene un carácter comparativo, se muestran datos referentes a los mecanismos estructurales que explican las diferencias de estabilidad entre ambas enzimas, lo cual podría tener implicancias para mejorar las aproximaciones de diseño racional de biocatalizadores termoestables.

Objetivo general: Caracterizar las propiedades bioquímicas, cinéticas y de estabilidad de las HMPKs de *S. typhimurium* y *T. Thermophilus*.

Objetivos específicos:

- Expresar heterológamente las HMPKs de *S. typhimurium* y *T. thermophilus* en *E. coli*.
- Purificar ambas enzimas mediante cromatografía de afinidad en columna de níquel.
- Demostrar la presencia de HMP-P y HMP-PP en los productos de reacción a través de cromatografía de intercambio aniónico.
- Determinar la especificidad por sustrato de *St*HMPK y *Tt*HMPK frente a distintas formas de vitamina B₆ (piridoxal, piridoxina y piridoxamina).
- Determinar los parámetros cinéticos de ambas enzimas (K_M , V_{max} , k_{cat}).
- Determinar el estado de oligomerización de ambas enzimas por cromatografía de exclusión molecular.
- Determinar la temperatura media de desplegamiento (T_m) y la concentración media desnaturante de cloruro de guanidinio (GndCl), mediante dicroísmo circular y fluorescencia intrínseca.
- Evaluar las diferencias estructurales entre ambas enzimas.

2. Materiales y Métodos

2.1 Manejo de cepas y vectores

2.1.1 Cepas

Se empleó la cepa de *Escherichia coli* DH5 α (genotipo: F-80dlacZ M15 (lacZYA-argF) U169 recA1 endA1hsdR17(rk-, mk+) phoA^{sup}E44 -thi-1 gyrA96 relA1) para la producción y purificación de plásmidos. Para la sobreexpresión de proteínas se utilizó la cepa de *E. coli* BL21(D3), cuyo genotipo es fhuA2 [lon] ompT gal (λ DE3) [dcm] Δ hsdS λ DE3 = λ sBamHIo Δ EcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δ nin5.

2.1.2 Vectores

Las secuencias codificantes de la 4-amino-5-hidroximetil-2-metil-pirimidina quinazas de *S. typhimurium* y *T. thermophilus* fueron sintetizadas por la compañía comercial GenScript (GenScript USA Inc., Piscataway, NJ, Estados Unidos) y recibidas en el vector de expresión pET-TEV-15b (derivado plásmido pET-15b), el cual incluye un *cassette* de resistencia a ampicilina. La región codificante de las HMPKs se posicionó entre los sitios de restricción *Nco*I y *Bam*HI, incluyendo un *tag* de polihistidina en el extremo amino terminal y un sitio de corte reconocido por la proteasa del virus del tabaco (TEV) entre el *tag* y el primer aminoácido de la secuencia. Las secuencias empleadas se muestran en formato FASTA en el anexo de este capítulo.

2.1.3 Medios de Cultivo

El medio Luria Bertani (LB) se preparó utilizando 10 g/L de Bacto-triptona, 5 g/L de extracto de levadura y 5 g/L de NaCl y se esterilizó mediante autoclave.

El medio *Terrific broth* (TB) se preparó utilizando 24 g/L de extracto de levadura, 12 g/L de triptona, 4 mL/L glicerol y se esterilizó mediante autoclave. Separadamente, se preparó y una solución de 100 mL de tampón fosfato (0.17 M KH_2PO_4 y 0.72 M K_2HPO_4). Ambas partes se mezclan en razón 1:9 inmediatamente antes del uso del medio.

2.1.4 Transformación de E. coli quimiocompetentes

Las células de *E. coli* quimiocompetentes fueron preparadas y transformadas siguiendo el protocolo estándar indicado en el kit *Mix & Go!*, utilizando 2 μL de plásmido purificado (ZymoResearch, Irvine, Estados Unidos).

2.1.5 Aislamiento de plásmidos.

Los vectores obtenidos desde GenScript fueron empleados para transformar *E. coli* DH5 α . Para la purificación de plásmidos se empleó el kit comercial E.Z.N.A. Plasmid Miniprep Kit II (Omega Bio-Tek, Norcross, GA, Estados Unidos) de acuerdo a las instrucciones del fabricante. Los plásmidos aislados fueron utilizados para las transformaciones subsecuentes de *E. coli* BL21(DE3).

2.2 Expresión y purificación de proteínas

2.2.1 Purificación de StHMPK y TtHMPK.

Las bacterias transformadas con el plásmido codificante de *StHMPK* fueron crecidas en medio LB con 100 µg/mL de ampicilina a 37 °C hasta alcanzar OD₆₀₀ de ~ 0.6. En el caso de *TtHMPK*, las bacterias fueron crecidas en medio TB con la misma temperatura y concentración de antibiótico, hasta alcanzar un OD₆₀₀ de ~ 0.8. La expresión de proteínas fue inducida por 0.5 mM de Isopropil-β-D-tiogalactopiranosido (IPTG) durante toda la noche a 30 °C para *StHMPK* y a 37 °C para *TtHMPK*. Las bacterias fueron centrifugadas por 15 minutos a 3.210 g y resuspendidas en Buffer de unión (Tris 50 mM pH 7,8 NaCl 0,5 mM, PMSF 1 mM, imidazol 20 mM). Posteriormente, las bacterias resuspendidas se sonicaron en hielo, con 10 pulsos de 20 segundos a 35% de amplitud con 59 segundos de descanso (Digital Sonifier 450, Branson). El lisado celular resultante se centrifugó a 18514 g por 30 minutos y la fracción soluble se filtró empleando filtros de 0.45 µm (Jet Biofil). En el caso de *TtHMPK*, la fracción soluble fue sometida a un shock térmico de 70 °C por 15 minutos previo a la filtración y se centrifugó nuevamente a 16514 g por 10 minutos. La fracción soluble se cargó en una columna HisTrap 5 mL (GE Healthcare) previamente equilibrada con buffer de unión empleando una bomba peristáltica BioRad con un flujo de 2 mL/minuto y se lavó con 50 mL de buffer de unión. *StHMPK* fue eluida con un gradiente de 100 mL de 20 a 250 mM de imidazol y *TtHMPK* con un gradiente de 20 a 500 mM de imidazol, usando un flujo de 2 mL/min y colectando fracciones de 2 mL en un sistema de purificación AKTA Prime (GE Healthcare). La pureza de las proteínas

obtenidas fue evaluada mediante SDS-PAGE. Las fracciones con actividad fueron mezcladas y dializadas contra 2 L de un buffer 50 mM Tris-HCl pH 7.5, 5 mM 2-mercaptoetanol y 10% glicerol para *StHMPK* y 50 mM Tris-HCl pH 7.5, 5 mM 2-mercaptoetanol, 300 mM NaCl y 10% glicerol para *TtHMPK*. En el caso particular de *StHMPK*, se agregaron 100 mM de arginina y se llevó la concentración de glicerol a 20% antes del almacenado. Las proteínas fueron alicuotadas en volúmenes de 400 μ L y almacenadas a - 80 °C hasta su uso posterior. La concentración de proteína fue determinada por absorbancia a 280 nm, asumiendo un coeficiente de extinción molar de 31970 M⁻¹ cm⁻¹ para *StHMPK* y 11460 M⁻¹ cm⁻¹ para *TtHMPK*, los que fueron determinados mediante el servidor web ProtParam de Expasy (Gasteiger y cols., 2005)

2.2.2 Electroforesis desnaturante de proteínas en geles de poliacrilamida (SDS-PAGE)

Se agregaron 20 μ g de proteína en una solución de Tris-HCl pH 7.6 60 mM, SDS al 10%, glicerol al 10%, β -mercaptoetanol al 5% y azul de bromofenol al 0.02% en un volumen final de 20 μ L y se calentó a ~100 °C por 5 minutos. La electroforesis se realizó en geles de poliacrilamida al 12% en presencia de solución TG-SDS 1X (3.02 g/L Tris, 14.4 g/L glicina, 1 g/L SDS a pH 8.4), a un amperaje constante de 30 mA. Una vez finalizada la electroforesis, los geles fueron teñidos con azul de Coomassie al 0.1% y desteñidos con una mezcla de metanol 40%, ácido acético 10% y agua.

2.3 Estequiometría de la reacción catalizada por StHMPK y TtHMPK

La actividad enzimática de *StHMPK* y *TtHMPK* se midió mediante el ensayo acoplado piruvato quinasa / lactato deshidrogenasa (Castro-Fernandez, 2014).

Brevemente, este ensayo se basa en el uso del ADP liberado de la fosforilación ATP-dependiente. El ADP es usado como sustrato por la piruvato quinasa en presencia de fosfoenolpiruvato (PEP) para generar ATP y piruvato, el piruvato es reducido a lactato en presencia de NADH por la lactato deshidrogenasa. La oxidación de NADH a NAD⁺ fue monitoreada espectrofotométricamente siguiendo la pérdida de absorbancia a 340 nm. Los ensayos estequiométricos se realizaron en un volumen final de 0.5 mL de una solución con 50 mM HEPES pH 7.0, 125 mM KCl, 0.3 mM PEP, 0.1 mM NADH, 0.05 mM HMP, 1 mM ATP, 5 mM MgCl₂, 1.6 U/mL de piruvato quinasa, 4.8 U/mL de lactato deshidrogenasa y 0.5 μM de *S*tHMPK o 10 μM de *T*tHMPK y asumiendo un coeficiente de extinción molar de 6.22 mM⁻¹ cm⁻¹.

2.4 Separación de productos de reacción mediante HPLC

Los productos de una mezcla inicial de sustratos (0.07 mg/mL de *S*tHMPK, 50 mM HEPES pH 7.0, 0.5 mM HMP, 0.1 mM ATP, 5 mM MgCl₂) fueron separados por cromatografía líquida de alto rendimiento (HPLC) utilizando una columna de intercambio aniónico (MonoQ HR 5/5, Pharmacia Biotechnology, pre equilibrada con buffer 25 mM Tris-HCl pH 7.5) con un gradiente de 0 a 35% de un buffer 1 M NaCl, 25 mM Tris-HCl pH 7.5 a flujo 1 mL/minuto, monitoreando la absorbancia a 235 nm. Al mix de reacción se le adiciono 300 mM KCl, 0.3 mM PEP y 1.6 U/mL de piruvato quinasa, para remover el ADP del medio de reacción y prevenir la co-migración con HMP-PP. La proteína se removió del medio de reacción transcurrido distintos tiempos de incubación por ultrafiltración con tubos Amicon (Sigma-Aldrich) de 10 KDa y 25 μL de muestra fueron cargados en la columna.

2.5 Especificidad por sustrato de *St*HMPK y *Tt*HMPK

Primero, se determinó la actividad específica de *St*HMPK y *Tt*HMPK en presencia de 100 μ M de HMP y 1 mM de ATP, en un buffer HEPES 50 mM pH 8.0, 125 mM KCl, 0.3 mM PEP, 0.1 mM NADH, 6 mM MgCl₂, 1.6 U/mL de piruvato quinasa, 4.8 U/mL de lactato deshidrogenasa, siguiendo la ecuación 1

$$A. E = \frac{Abs/s \times 60 \times V_T}{6.22 \times V_E \times C} \quad \text{Ecuación 1}$$

Donde A.E es la actividad específica medida en unidades de actividad enzimática (U, definida como μ M min⁻¹ de ADP producido) por mg, Abs/s corresponde al cambio de absorbancia por segundo, V_T es el volumen en mL en que se realiza la reacción, V_E es el volumen en mL de enzima HMPK usado, C es la concentración en mg/mL de la enzima HMPK usada y 6.22 representa el coeficiente de extinción molar del NADH, medido en mM⁻¹ cm⁻¹. Luego, se determinó la actividad en presencia de 100 μ M de piridoxal, piridoxina o piridoxamina.

2.6 Reconstrucción filogenética entre HMPKs y PLKs.

Se extrajeron secuencias del producto del gen *thiD*, *pdxK* y *pdxY* desde la base de datos UniProtKB mediante búsqueda directa. Se realizó un alineamiento de secuencias múltiple con restricciones estructurales derivadas de las estructuras cristalográficas de *St*HMPK (PDB 1JXI), *Tt*HMPK (PDB 1UB0), la PLK de *S. aureus* (PDB 4C5J) y la PLK de *B. subtilis* (PDB 2I5B), mediante el servidor web PROMALS-

3D (Pey y cols., 2008). Se eligió el modelo evolutivo que explica mejor el set de datos utilizando el programa ProtTest 3.0 y discriminando según el criterio de información de Akaike corregido (Darriba y cols., 2011). El árbol filogenético fue construido por método de máxima verosimilitud empleando el programa PhyML (Guindon y cols., 2010) utilizando un modelo de sustitución aminoacídica LG con corrección por sitios invariantes y con variación entre sitios de acuerdo a una distribución gamma, la robustez de los nodos se evaluó mediante el test de razón de verosimilitudes.

2.7 Dinámica molecular del complejo ternario ATP-HMP de StHMPK

El complejo ternario ATP·Mg - HMP de *StHMPK* fue modelado utilizando Modeller v9.15 (Fiser y cols., 2000), empleando como plantados estructurales la enzima *StHMPK* en presencia de HMP (PDB 1JXI) y la PLK de *S. aureus* (PDB 4C5J) en presencia de un análogo no hidrolizable de ATP. Los ligandos fueron incorporados mediante la función *env.io.hetatm* y se utilizó el alineamiento generado durante la inferencia filogenética para la construcción de los modelos. Se generaron 30 modelos, de los cuales se seleccionó el mejor mediante la función energética basada en potenciales estadísticos DOPE Score (Shen & Sali, 2006).

Para ATP y Mg, se utilizaron los parámetros disponibles en la literatura (Allner y cols., 2012; Meagher y cols., 2003) y para la proteína se utilizó el campo de fuerza *ff99SB*. Para el caso de HMP, se utilizó como base el campo de fuerza general de amber GAFF (Wang y cols., 2004). Los parámetros faltantes y la topología de la molécula fueron generados utilizando el programa *antechamber* incorporado en AmberTools (Wang y cols., 2001). El sistema fue solvatado en una caja de modelo de agua TIP3P

extendida a 10 Å por sobre los átomos de proteína y neutralizado con iones monovalentes (Na⁺/Cl⁻) utilizando el módulo *tleap* de AMBER (Case y cols, 2005). El sistema se llevó a una conformación de mínima energía mediante un proceso iterativo escalonado. Primero se minimizaron las estructuras con una restricción posicional armónica sobre todos los átomos pesados (C, O, N, S) de 500 kcal/mol x Å², empleando 1000 pasos de gradiente conjugado y 1000 pasos de máxima pendiente, luego se repitió el proceso de minimización usando los mismos pasos, pero sólo con restricciones posicionales a los C α , disminuyendo la magnitud de la restricción en cada iteración. Se realizaron en total 5 minimizaciones extra, restringiendo el sistema a 50, 40, 30, 20 y a 5 kcal/mol x Å². Los sistemas de mínima energía resultantes del último proceso de minimización fueron equilibrados en una simulación con condiciones de límite periódicas en el ensamble canónico (NVT) aumentando la temperatura desde 100 K a 300 K en una ventana de 200 ps, empleando un paso de tiempo de 2 fs y 100000 pasos totales de simulación. Después se realizaron simulaciones de producción con condiciones de límite periódicas en el ensamble isotérmico e isobárico (NPT, a 300 K y 1 atm) de 50 ns, empleando un paso de tiempo de 2 fs y 25000000 pasos totales de simulación, y controlando la temperatura mediante termostato Langeviano. Se utilizó un *cut-off* de 10 Å para el cálculo de interacciones electrostáticas coulombicas de corto alcance y de van der Waals y se utilizó el método de grillas de partículas de Ewald (PME) para evaluar las interacciones electrostáticas de largo alcance. La desviación cuadrática media (RMSD) de los átomos del sistema fue monitoreado empleando *RMSD Trajectory Tool* de VMD, según la ecuación 2 (Humphrey y cols., 1996)

$$RMSD = \sqrt{\frac{1}{N} \sum_{j=0}^T \sum_{i=1}^N (r_i(t_j) - r_i(t_0))^2} \quad \text{Ecuación 2}$$

Donde N corresponde al número total de átomos, T corresponde al tiempo total de simulación, $r_i(t_j)$ es la posición del i -ésimo átomo en el tiempo t_j y $r_i(t_0)$ corresponde a la posición inicial del i -ésimo átomo.

2.8 Determinación de las constantes cinéticas de StHMPK Y TtHMPK

Las preparaciones enzimáticas se mezclaron con una solución con 50 mM HEPES pH 7,0, 125 mM KCl, 0.3 mM PEP, 0.1 mM NADH, 5 mM MgCl₂, 1.6 U/mL de piruvato quinasa, 4.8 U/mL de lactato deshidrogenas y se variaron las concentraciones de uno de los co-sustratos mientras el otro se mantuvo saturante. Se registró la pendiente de la región lineal de las curvas de progreso y se transformó la tasa de consumo sustratos a actividad específica (U/mg) empleando la ecuación 1.

Se ajustó un modelo Michaelis-Menten (ecuación 3) a la curva de HMP obtenida para *StHMPK* y *TtHMPK*, mediante una regresión no lineal empleando GnuPlot (Racine, 2006).

$$v_0 = \frac{V_{max}}{K_M + [S]} \quad \text{Ecuación 3}$$

Donde v_0 es la velocidad de reacción, V_{\max} es la velocidad máxima del sistema, K_m es la constante Michaelis-Menten y $[S]$ es la concentración de sustrato. La constante catalítica k_{cat} (s^{-1}) se calculó de acuerdo a la ecuación 4.

$$k_{cat} = \frac{V_{max} \times PM}{1000 \times 60} \quad \text{Ecuación 4}$$

Donde V_{\max} es la velocidad máxima en U/mg, PM corresponde al peso molecular en g/mol, determinado a través del servidor web ProtParam de ExPASy (Gasteiger y cols., 2005). Los valores empleados fueron 30.94 KDa para *St*HMPK y 29.18 KDa para *Tt*HMPK.

2.9 Cromatografía de exclusión molecular

Se cargaron 20 μ L de proteína a 1 μ M de concentración (*St*HMPK o *Tt*HMPK) en una columna Superdex-75g (GE Healthcare) y se eluyeron con un gradiente isocrático de buffer 25 mM Tris-HCl pH 7.5, 300 mM NaCl y 5 mM 2-mercaptoetanol a un flujo de 0.8 mL minuto. Los tiempos de retención se determinaron monitoreando la absorbancia a 280 nm. La columna se calibró con un estándar molecular de peso molecular conocido (BioRad). Se ajustó un modelo de decaimiento exponencial a los datos obtenidos usando el programa GraphPad Prism.

2.10 Determinación de la estabilidad de StHMPK y TtHMPK

La temperatura media de desplegamiento (T_m) de cada proteína se determinó a través de dicroísmo circular (CD) monitoreando la señal de elipticidad a 222 nm de una muestra con 0.2 mg/mL de proteína en un buffer 50 mM $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$ pH 7.8 en un volumen final de 0.3 mL, sometida a rampas de temperatura de gradiente 1 y 3 °C/minuto. Los datos de elipticidad obtenidos se transformaron a fracción nativa (f_N) siguiendo la ecuación 5.

$$f_N = 1 - f_U \quad \text{Ecuación 5}$$

Donde f_U corresponde a la fracción desplegada, la cual se calcula mediante la ecuación 6.

$$f_U = \frac{\theta - \theta_N}{\theta_U - \theta_N} = \frac{\theta - (m_N T + C_N)}{(m_U T + C_U) - (m_N T + C_N)} \quad \text{Ecuación 6}$$

Donde θ es la elipticidad medida, θ_N es la señal del estado nativo, θ_U es la señal del estado desplegado, m_N es la pendiente del ajuste lineal de la región pre-exponencial, C_N es la constante del ajuste de la región pre-exponencial, m_U es la pendiente del ajuste lineal de la región post-exponencial, C_U es la constante del ajuste de la región post-exponencial y T es la temperatura. Se ajustó una distribución de Boltzmann a los datos para obtener el valor exacto de T_m de acuerdo a la ecuación 7.

$$f_N = I - \frac{I}{1 + e^{(Tm - \frac{T}{m})}} \quad \text{Ecuación 7}$$

Donde e es el número de euler, T es la temperatura, Tm es la temperatura media de desplegamiento y m es la pendiente de la región exponencial.

La concentración media desnaturante de cloruro de guanidinio (C_m) se determinó monitoreando la fluorescencia intrínseca de triptófanos en un fluorímetro JASCO FP8300, frente a concentraciones crecientes del desnaturante (GndCl). Las reacciones de desplegamiento se realizaron utilizando una concentración final de proteína de 1 μ M, en un buffer con 50 mM HEPES pH 7,0, 100 mM arginina, 5% glicerol y 1 mM TCEP para *St*HMPK y 50 mM HEPES pH 7,0, 300 mM NaCl y 1 mM TCEP para *Tt*HMPK. Las muestras fueron equilibradas toda la noche a 25 °C antes de realizar las mediciones. Las muestras fueron excitadas a 295 nm y se registró el espectro de emisión entre 305 y 450 nm. Se determinó la fracción nativa utilizando la ecuación 6, pero reemplazando la observable experimental θ por la fluorescencia a 320 nm para *St*HMPK y la intensidad máxima de fluorescencia para *Tt*HMPK. Los datos obtenidos fueron ajustados a un modelo idéntico que el mostrado en la ecuación 7, pero que considera el valor de C_m en lugar de T_m y la concentración de GndCl en lugar de la temperatura, tomando los datos desde el punto de máxima intensidad de fluorescencia.

2.11 Comparación estructural entre StHMPK y TtHMPK

2.11.1 Dinámica molecular de StHMPK y TtHMPK

Se utilizaron como base las estructuras cristalográficas en conformación apo de *StHMPK* y *TtHMPK* (PDB: 1JXH y 1UB0, respectivamente). Los *loops* faltantes en ambas estructuras fueron adicionados empleando Modeller v9.15 (Fiser y cols., 2000). Las cargas formales en los residuos protonables fueron asignadas usando Propka (Li y cols., 2005), incorporado en el conjunto de *Protein Preparation Wizard* del software Maestro (Schrödinger, 2018). Las estructuras resultantes fueron solvatadas en cajas de modelo de agua TIP3P extendida a 10 Å por sobre los átomos de proteína y neutralizadas con iones monovalentes (Na⁺ y Cl⁻, según corresponde), utilizando el módulo *tleap* de AMBER (Case y cols, 2005). Las estructuras fueron minimizadas empleando el campo de fuerza *ff99SB* de AMBER (Wang y cols., 2005), mediante un proceso iterativo escalonado. Primero se minimizaron las estructuras con una restricción posicional armónica sobre todos los átomos pesados (C, O, N, S) de 500 kcal/mol x Å², empleando 1000 pasos de gradiente conjugado y 1000 pasos de máxima pendiente, luego se repitió el proceso de minimización usando los mismos pasos, pero sólo con restricciones posicionales a los C α , disminuyendo la magnitud de la restricción en cada iteración. Se realizaron en total 5 minimizaciones extra, restringiendo el sistema a 50, 40, 30, 20 y a 5 kcal/mol x Å². Los sistemas de mínima energía resultantes del último proceso de minimización fueron equilibrados en una simulación con condiciones de límite periódicas en el ensamble canónico (NVT) aumentando la temperatura desde 100 K a 300 K en una ventana de 200 ps, empleando un paso de tiempo de 2 fs y 100000 pasos

totales de simulación. Después se realizaron simulaciones de producción con condiciones de límite periódicas en el ensamble isotérmico e isobárico (NPT, a 300 K y 1 atm) de 50 ns, empleando un paso de tiempo de 2 fs y 25000000 pasos totales de simulación, y controlando la temperatura mediante termostato Langeviano. Se utilizó un *cut-off* de 10 Å para el cálculo de interacciones electrostáticas coulómbicas de corto alcance y de van der Waals y se utilizó el método de grillas de partículas de Ewald (PME) para evaluar las interacciones electrostáticas de largo alcance.

2.11.2 Análisis de las trayectorias

El RSMD de los C α de cada sistema se determinó utilizando la extensión *RMSD Trajectory Tool* de VMD (ver ecuación 2) (Humphrey y cols., 1996). El RMSF de los C α se determinó usando la función *atomicfluct* del software *cpptraj*, según la ecuación 8 (Roe & Cheatham, 2013).

$$RMSF_i = \sqrt{\frac{1}{T} \sum_{t_i=1}^T |r_i(t_i) - r_i^{REF}|^2} \quad \text{Ecuación 7}$$

Donde $RMSF_i$ corresponde a la fluctuación del i -ésimo átomo, T es la cantidad total de pasos de simulación analizados, $r_i(t_i)$ representa la posición del i -ésimo átomo en el i -ésimo paso de simulación y r_i^{REF} corresponde a la posición del i -ésimo átomo en la estructura de referencia. Para los cálculos de RMSF la estructura de referencia utilizada fue la conformación derivada del análisis de *clusters* detallado más adelante.

Las matrices de correlación cruzada de residuos fueron calculadas empleando el paquete Bio3D para el software de estadística R, siguiendo la ecuación 8 para correlación de elementos en el espacio (Grant y cols., 2006).

$$C_{ij} = \frac{\Delta t}{T} \left[\sum_{t=0}^{T-\Delta t} r_i(t)r_j(t) - \frac{\Delta t}{T} (\sum_{t=0}^{T-\Delta t} r_i(t))(\sum_{t=0}^T r_j(t)) \right] \quad \text{Ecuación 8}$$

Donde C_{ij} es el índice de correlación entre los átomos i y j , Δt corresponde al paso de tiempo de la simulación, T corresponde al tiempo total de simulación analizado, $r_i(t)$ es la posición del átomo i en el tiempo t , y $r_j(t)$ es la posición del átomo j en el tiempo t . Posteriormente, los índices de correlación son normalizados para que fluctúen entre -1 y 1 mediante la ecuación 9.

$$C_{ij} = \frac{C_{ij}}{\sqrt{C_{ii}}\sqrt{C_{jj}}} \quad \text{Ecuación 9}$$

Las estructuras representativas de los ensambles conformacionales explorados durante la simulación fueron determinadas mediante análisis de *cluster* utilizando la función *cluster* de *cpptraj*. Este método, se basa en la agrupación de datos en *clusters*, de manera que los datos pertenecientes a un *cluster* son más similares entre sí, que los datos fuera del cluster (Roe y cols., 2014). Las variaciones de RMSD mayores a 1 Å dentro de la trayectoria fueron consideradas significativas y por ende, agrupadas en ensambles

distintos, y además, una cantidad mínima de 25 estructuras (equivalentes a mantener el sistema dentro del ensamble por 125 ps) para considerar al ensamble como relevante.

2.11.3 Composición aminoacídica del núcleo de StHMPK y TtHMPK

En base a la estructura representativa del ensamble conformacional más visitado durante el tiempo de simulación, se determinó la distribución de aminoácidos en el núcleo y la superficie de la proteína. Para esto, se calculó la superficie molecular de cada residuo de la proteína mediante el servidor web WHAT-IF (Vriend, 1990) y se categorizaron de acuerdo a esta. Si la superficie molecular del residuo es mayor a 5 \AA^2 , se consideró como residuo expuesto a la superficie, de lo contrario se consideró como parte del núcleo. Luego, se analizó la composición porcentual para cada residuo en el núcleo de la proteína.

2.11.4 Análisis de interacciones no covalentes en StHMPK y TtHMPK

Las interacciones iónicas presentes en las estructuras representativas de la dinámica fueron calculadas utilizando el servidor web *Protein Interaction Calculator* (PIC) (Tina y cols., 2007), utilizando una distancia máxima de 3.5 \AA . Para determinar la existencia de pares iónicos a lo largo de la trayectoria de dinámica molecular, se utilizó la extensión *Salt Bridges* de VMD, sólo se consideraron como interacciones significativas aquellas que persisten por más del 5% de los 40 ns de simulación analizados.

Las interacciones por puente de hidrógeno se calcularon utilizando la función *hbond* de *cpptraj*, utilizando como criterios de corte una distancia máxima de 3.5 Å entre el hidrógeno del dador y el átomo aceptor, y un ángulo entre el átomo dador de puente de hidrogeno, el hidrógeno y el átomo aceptor de al menos 120° (de un máximo de 180°). Para el conteo de interacciones, sólo se consideraron aquellos puentes que se mantienen en configuraciones congruentes con la definición dada por al menos 5 % del tiempo de simulación.

2.12 Anexo

2.12.1 Secuencia de *StHMPK*

>*StHMPK*

MGSSHHHHHSSGENLYFQGHMQRINALTIAGTDPSGGAGIQADLKTFSALGAYGCSVI
TALVAQNTCGVQSVYRIEPDFVAAQLDSVFSVDRIDTTKIGMLAETDIVEAVAERLQRH
HVRNVVLDTVMLAKSGDPLLSPSAIETLRVRLLPQVSLITPNLPEAAALLDAPHARTEQE
MLAQGRALLAMGCEAVLMKGGHLEDAQSPDWLFTREGEQRFSAPRVNTKNTHGTGC
TLAALAALRPRHRSWGETVNEAKAWLSAALAQAQTLEVGKGIGPVHHFHAWW

2.12.2 Secuencia de *TtHMPK*

>*TtHMPK*

MGSSHHHHHSSGENLYFQGHMRVALTIAGSDSGGGAGVQADLKVFFRFGVYGTSAL
TLVTAQNTLGVQRVHLLPPEVVYAQIESVAQDFPLHAAKTGALGDAAIVEAVAEAVRR
FGVRPLVVDPMVAKSGDPLLAKEAAAALKERLFPLADLVTPNRLEAEALLGRPIRTLK
EAEEAAKALLALGPKAVLLKGGHLEGEEAVDLLATRGGVLRFSAPRVHTRNTHGTGCT
LSAAIAALLAKGRPLAEAVAEAKAYLTRALKTAPSLGHGHGPLDHW

3. Resultados

3.1 Expresión heteróloga y purificación de *StHMPK* y *TtHMPK*

Ambas enzimas fueron expresadas utilizando el vector pET-TEV-15b en *E. coli* BL21(DE3) y purificadas mediante un único paso cromatográfico empleando una columna de afinidad de níquel, eluyendo con un gradiente de imidazol. En el caso de *TtHMPK* se realizó un paso previo a la cromatografía, que consistió en calentar el lisado celular a 70°C por 15 minutos y descartar la fracción insoluble mediante centrifugación. El resultado de la purificación de ambas enzimas se muestra en la figura 6. En ambos casos, la proteína eluye en un único pico entre los 35 y 50 mL, cuando se alcanza un ~35 % de la solución de elución.

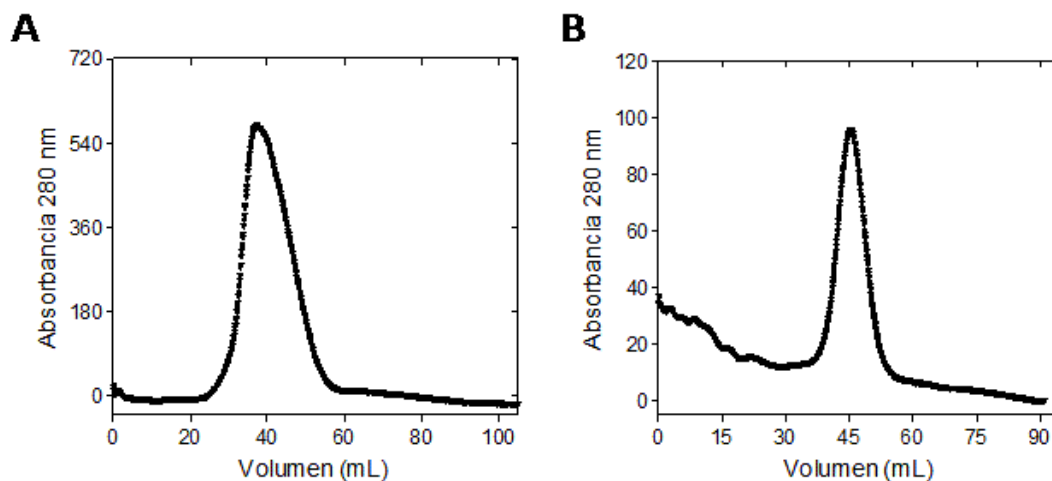


Figura 6. Cromatogramas representativos de las eluciones de *StHMPK* (A) y de *TtHMPK* (B) en columnas de afinidad de níquel.

La pureza de las proteínas fue analizada por SDS-PAGE. A modo de control, se cargaron 20 μ g de proteína de cada etapa del proceso de purificación (figura 7).

En ambos casos, los resultados indican que tras un único paso cromatográfico se obtiene una proteína altamente pura (cerca al 95%). Para *St*HMPK se obtienen alrededor de 15 mg por litro de cultivo, mientras que el rendimiento para *Tt*HMPK es de ~5 mg por litro de cultivo.

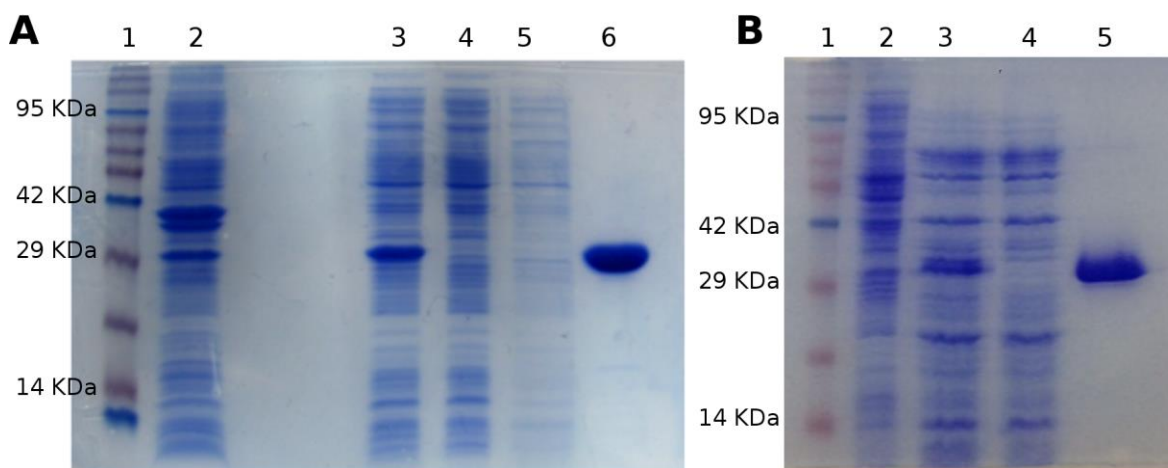


Figura 7. SDS-PAGE de la purificación de *St*HMPK (A) y *Tt*HMPK (B). A) 1. Estándares de peso molecular, 2. Pella, 3. Lisado celular, 4. Frente de la columna, 5. Lavado, 6. Proteína pura. B) 1. Estándar de peso molecular, 2. Pella, 3. Lisado celular, 4. Frente de la columna, 5. Proteína pura.

3.2 Análisis de la fosforilación doble de HMP por *St*HMPK y *Tt*HMPK

Para demostrar que las enzimas purificadas son funcionales y que poseen la capacidad de catalizar dos fosforilaciones dependientes de ATP, se realizaron análisis estequiométricos entre el consumo de NADH (asociado a la aparición de ADP) y la disponibilidad de HMP en el medio de reacción. Se realizaron experimentos de

titulación de ADP empleando el ensayo acoplado PK/LDH y utilizando una concentración inicial de 50 μM de HMP (Figura 8).

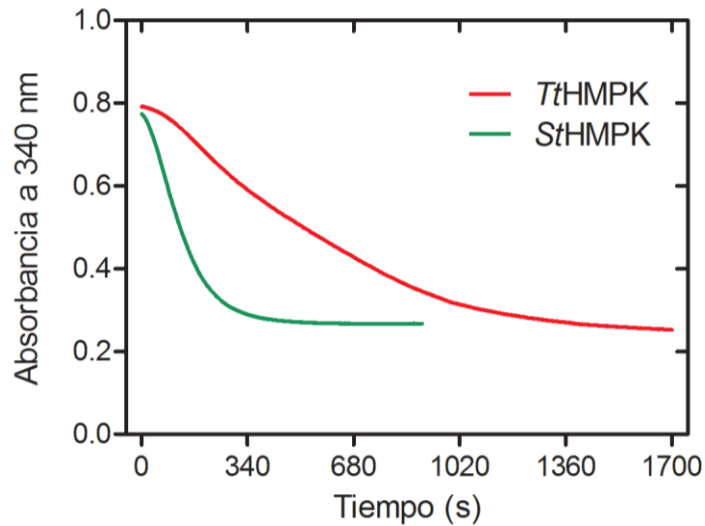


Figura 8. Curvas de progreso representativas de la titulación de ADP para *StHMPK* (verde) y *TtHMPK* (rojo).

En base a las diferencias entre la absorbancia inicial y final, es posible establecer la cantidad de moles de ADP producidos (equivalentes a los moles de ATP consumidos) por mol de HMP, considerando que por cada mol de ADP producido, se oxida 1 mol de NADH y que este tiene un coeficiente de extinción molar de $6.22 \text{ mM}^{-1} \text{ cm}^{-1}$ (Tabla I.)

Tabla I. Estequiometría del consumo de NADH por producción de ADP, empleando 50 μM de HMP. Se indican los valores promedio de 2 réplicas biológicas y el error estándar.

Proteína	Abs. inicial (340 nm)	Abs. final (340 nm)	Δabs	NADH consumido (μM)	[NADH]/[HMP]
<i>St</i> HMPK	0.81 ± 0.03	0.21 ± 0.06	0.60 ± 0.09	96.65 ± 15.24	1.93 ± 0.30
<i>Tt</i> HMPK	0.77 ± 0.02	0.24 ± 0.02	0.54 ± 0.002	86.58 ± 0.24	1.73 ± 0.005

Estos resultados indican que por cada mol de HMP que se consume, se producen 1.93 mol de ADP para *St*HMPK y 1.73 mol de ADP para *Tt*HMPK, lo que evidencia que ambas enzimas pueden fosforilar dos veces el sustrato HMP.

Para obtener una confirmación más analítica de la doble fosforilación realizada por las HMPK, se realizó un análisis de la formación de productos mediante HPLC. Para esto, se utilizó una columna de intercambio aniónico para separar los productos de reacción según su grado de fosforilación, los que se eluyeron aumentando la fuerza iónica del medio.

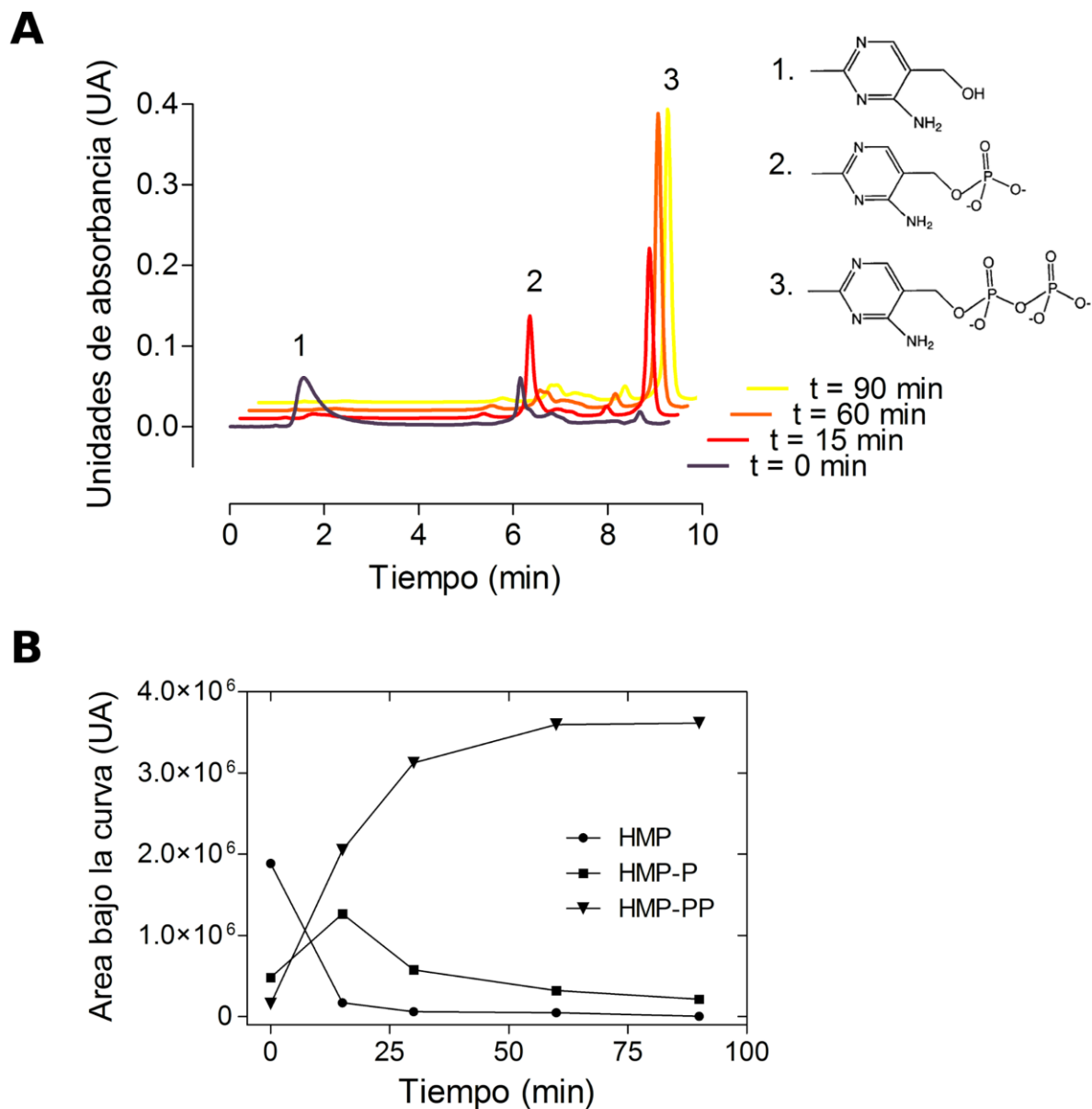


Figura 9. Análisis de los productos de la reacción catalizada por *Sr*HMPK mediante cromatografía de intercambio aniónico. A) Cromatogramas resultantes de la separación de 4 muestras distintas, correspondientes a diferentes tiempos de incubación con la enzima, monitoreadas a 235 nm. Cada compuesto dentro de la mezcla (1. HMP, 2.HMP-P, 3. HMP-PP), fue asignado según los tiempos de retención esperados de acuerdo a los grados de fosforilación de cada uno. B) Curvas de progreso para HMP, HMP-P y HMP-PP, resultantes de la integración del área bajo la curva de los cromatogramas mostrados en A.

Del cromatograma resultante y de las curvas de progreso inferidas a partir de éste (Figura 9A y 9B, respectivamente), se observa la desaparición tiempo-dependiente del primer pico y la aparición consecutiva de dos picos con tiempos de retención 6.1 y 8.6 minutos respectivamente. Sin embargo, se observa que el segundo pico (HMP-P) aparece de manera transitoria y desaparece tras tiempos largos de reacción.

3.3 Especificidad por sustrato en HMPKs

La HMPK de *E. coli* ha sido reportada como una enzima altamente específica por HMP, siendo incapaz de catalizar la fosforilación de sustratos muy similares, como piridoxal (PL), piridoxina (PN) y piridoxamina (PM) (Mizote y cols., 1998). Por otro lado, la primera descripción de un homólogo a HMPK en un organismo Gram positivo (*B. subtilis*) reveló que éste cataliza la fosforilación de piridoxal (PL) como sustrato principal y una única fosforilación con menor afinidad sobre HMP, lo que forzó la recategorización de este gen de *thiD* a *pdxK* (Park y cols., 2004). Estudios posteriores corroboraron que PdxKs de otros organismos Gram positivos también fosforilan tanto PL como HMP (Nodwell y cols., 2014). Con el objetivo de evaluar la especificidad por sustrato de *StHMPK* y *TtHMPK*, se realizaron ensayos de actividad en presencia de 100 μ M de PL, PM o PN y 1 mM de ATP. Los resultados muestran que ambas enzimas son altamente específicas por HMP (Figura 10).

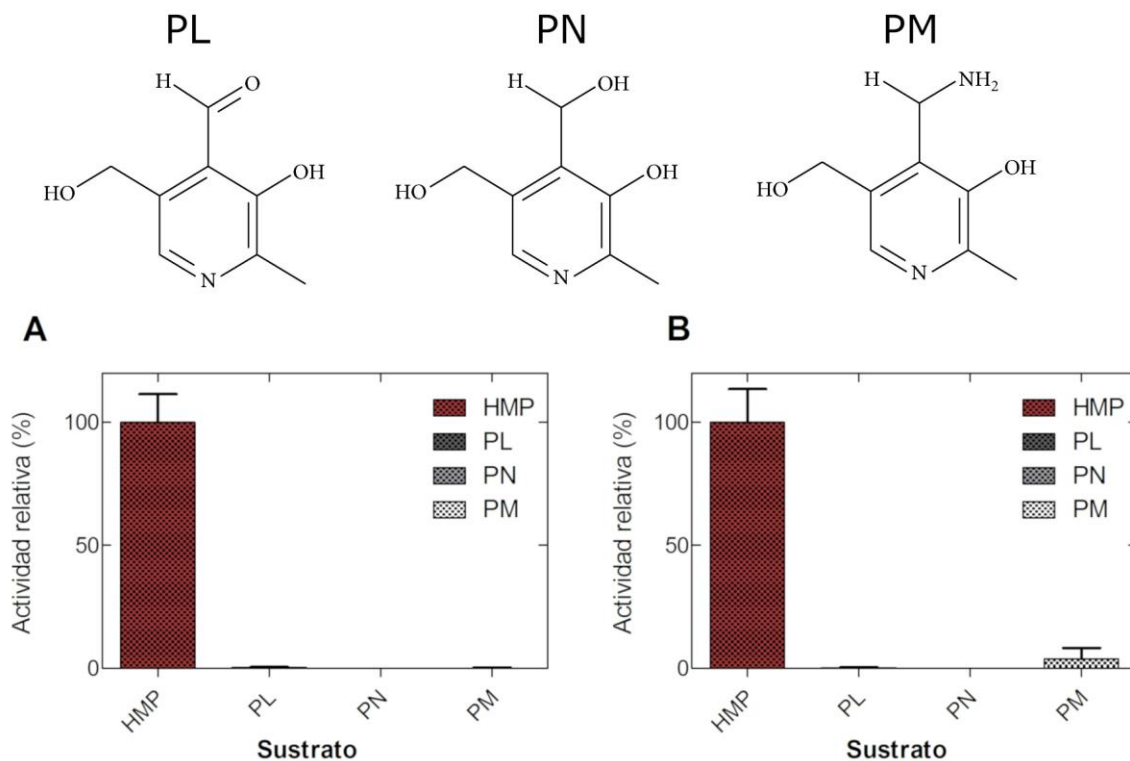


Figura 10. Especificidad de sustrato en (A) *SthHMPK* y (B) *TthHMPK*. En ambos casos, la enzima prácticamente no muestra actividad en presencia de piridoxal (PL), piridoxina (PN) y piridoxamina (PM). La actividad se normalizó en relación a la actividad específica frente a HMP, la altura de la barra representa el promedio de tres réplicas y la barra de error representa la desviación estándar.

La alta especificidad por HMP y la capacidad de catalizar dos fosforilaciones consecutivas observadas en las HMPKs contrasta notablemente con lo reportado para las enzimas con actividad HMPK/PLK codificadas en el gen *pdxK* presentes en organismos Gram positivos, que catalizan la fosforilación de múltiples sustratos, pero sólo son capaces de fosforilar sobre hidroxilos por lo que sólo fosforilan una vez. Si bien estudios anteriores muestran que las PLK/HMPK están más emparentadas con las HMPK que con las PLKs canónicamente estudiadas, presentes en organismos Gram negativos (Yang y cols., 1996), estos resultados sólo se basan en alineamientos estructurales de proteínas,

y no han sido exhaustivos en recopilar todas las secuencias parálogas presentes en cada organismo. Para determinar la distancia evolutiva entre ambos grupos de enzimas, se realizó un alineamiento de secuencias múltiples para construir un árbol filogenético por máxima verosimilitud, incluyendo secuencias aminoacídicas codificadas por genes *thiD* (HMPK), por genes anotados como *pdxK* (PLK o HMPK/PLK) y secuencias de la piridoxal quinasa codificada en el gen *pdxY*, otro miembro de esta familia de quinastas que es específica por piridoxal (Yang y cols., 1998) (Figura 11).

El árbol filogenético construido muestra una divergencia evolutiva clara entre las PdxK incorporadas. Existe un grupo parafilético de PdxK (destacado en rojo en fig.11), compuesto por secuencias provenientes de organismos Gram negativos y algunos organismos Gram positivos pertenecientes a los phylums como *Firmicutes*, *Lactobacillales* y *Clostridiales* y se encuentran más emparentadas con las piridoxal quinastas codificadas en el gen *pdxY*, las cuales corresponden a otro subgrupo perteneciente a la superfamilia riboquinasa (Safo y cols., 2004). Todas las ThiD analizadas se agrupan en un único grupo (destacado en azul), del cual emerge otro grupo monofilético de PdxK, donde se encuentran las secuencias de PdxK de organismos Gram positivos, reportadas como enzimas bifuncionales, como la HMPK/PLK de *S. aureus* y *B. subtilis*.

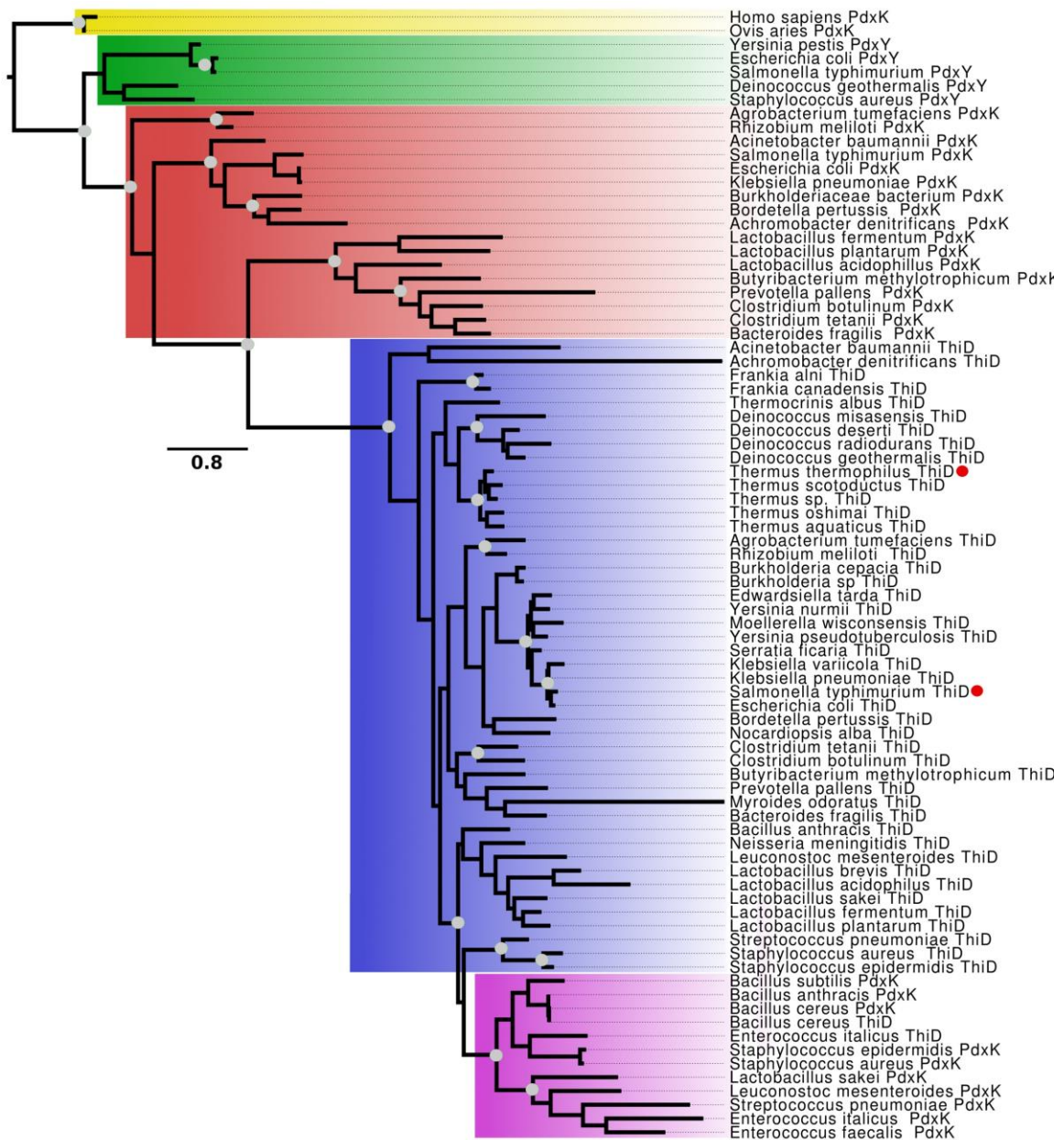


Figura 11. Árbol filogenético de máxima verosimilitud para secuencias de proteínas PLK y HMPK. El árbol fue enraizado con secuencias de piridoxal quinasa de organismos eucariontes. Las secuencias de HMPKs de *Enterobacteriales* se destacan en verde, las secuencias de *Thermales* se destacan en rojo, las secuencias de PdxK Gram positivos se muestran en cyan y las PdxK de Gram negativos se muestran en dorado. La escala indica el número de sustituciones por sitio. Los nodos destacados en gris tienen una probabilidad > 0.95 de acuerdo al test de razón de verosimilitudes. Los puntos rojos destacan las proteínas caracterizadas en este seminario de título.

Pese que las HMPK y las HMPK/PLK (PdxK) de Gram positivos están emparentados, han divergido significativamente en sus propiedades funcionales. Por lo tanto, si se desea conocer las bases estructurales de la especificidad por sustrato, es necesario observar en detalle el sitio de interacción entre la proteína y el sustrato. Sin embargo, sólo la HMPK/PLK de *S. aureus* posee una estructura cristalográfica del complejo ternario con un análogo no hidrolizable de ATP (PDB 4C5N), mientras que para HMPK solo *S. typhimurium* cuenta con una estructura depositada en presencia de HMP (PDB 1JXI).

Por esta razón, para obtener una visión detallada del sitio activo, se modeló el complejo ternario HMP - ATP·Mg de *StHMPK* empleando Modeller v9.15, utilizando como templados estructurales de *S. aureus* en presencia de un análogo no hidrolizable de ATP (PDB 4C5M) y la estructura con HMP de *StHMPK* (PDB 1JXI) y se realizaron simulaciones de dinámica molecular. El sistema fue simulado por 50 ns en el ensamble isotérmico - isobárico a 300 K y 1 atm, considerando los 10 primeros nanosegundos como fase de equilibrado y los 40 nanosegundos posteriores como fase productiva.

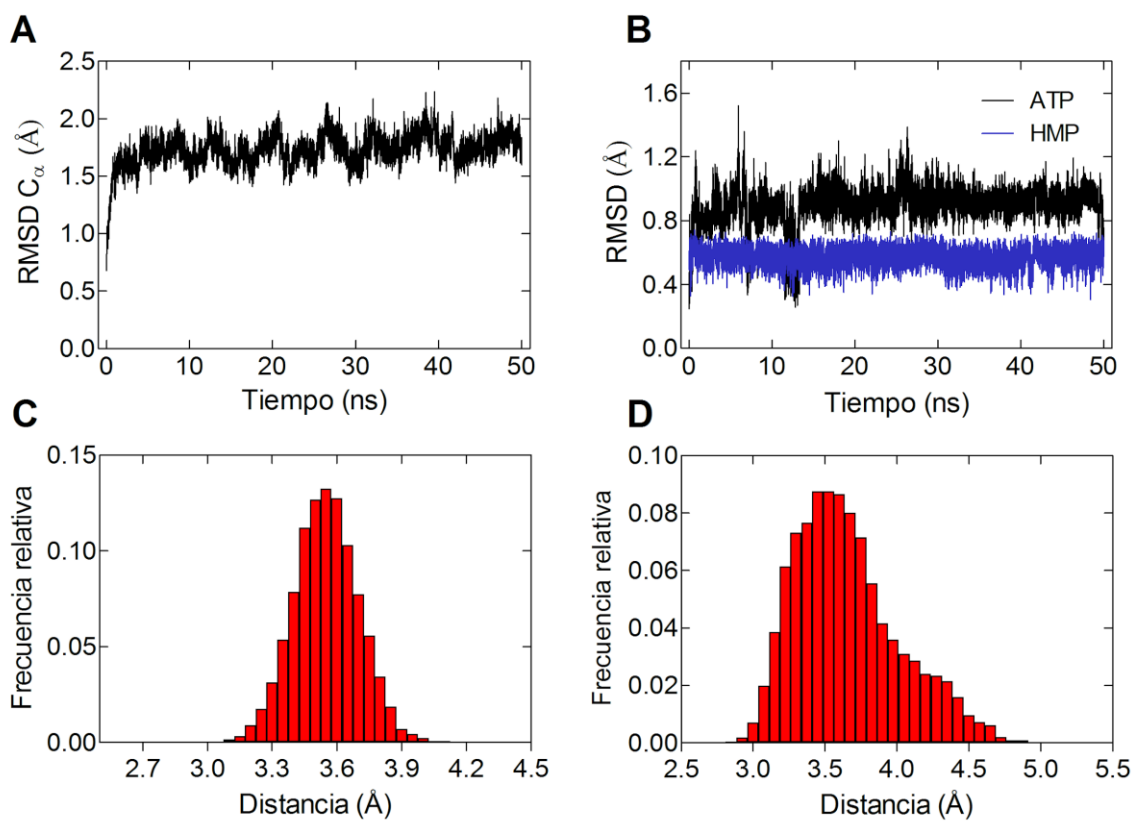


Figura 12. La simulación del complejo ternario de *StHMPK* es estable durante la ventana de tiempo estudiada y tiene coherencia con la reacción catalizada. A) RMSD de los carbonos alfa de la proteína. B) RMSD de los átomos pesados de los sustratos (ATP en negro y HMP en azul). C) Histograma de distribución de la distancia entre el fosfato γ del ATP y el hidroxilo libre del HMP. D) Histograma de distribución de la distancia entre el grupo tiol de la cisteína catalítica C213 y el hidroxilo del HMP.

El RMSD de los carbonos alfa de la proteína y el RMSD de los átomos pesados de los ligandos (Figura 12A-B) indican que el sistema converge y es estable dentro del pozo energético muestreado durante en tiempo de simulación. Si el sistema simulado tiene sentido catalítico, es de esperarse que las distancias entre el fosfato γ del ATP y el hidroxilo aceptor del HMP se mantengan a una distancia menor a 4.5 Å y permanezcan en una disposición alineada para que la transferencia de fosfatos sea viable. Al analizar

la distribución de distancias entre estos dos átomos a lo largo de todos los pasos de la simulación (Figura 12C), se observa que la distancia promedio entre ambos es de $3.55 \pm 0.15 \text{ \AA}$, lo que concuerda con la distancia esperada para una reacción de transferencia química. Además, se ha demostrado que la catálisis en la familia de quinasas de cofactores dependientes de ATP es mediada por un residuo de cisteína que actúa como base catalítica y desprotona el hidroxilo aceptor del fosfato, generando una especie nucleofílica que ataca el fosfato γ del ATP, por lo que la transferencia transcurriría por un mecanismo químico SN_2 (Dyguda y cols., 2004). La distancia promedio observada en la simulación entre el grupo tiol del residuo catalítico Cys213 y el hidroxilo del HMP es de $3.65 \pm 0.36 \text{ \AA}$, lo que también corrobora que el resultado de la simulación es concordante con lo esperado para esta reacción. También se observa que el átomo de Mg^{+2} está siendo coordinado por los oxígenos presentes en los fosfatos γ y β del ATP junto con 4 moléculas de agua, completando la geometría octaédrica de la capa de solvatación de este metal. A su vez, las aguas de la capa de solvatación están siendo coordinadas por los residuos Asn139 y Gln142, que conforman el motivo NXXE, el cual está altamente conservado en la superfamilia riboquinasa (Abarca-Laguna y cols., 2015). Por ende, los resultados apuntan a que los parámetros empleados en el campo de fuerza están generando una descripción con sentido fisicoquímico del complejo enzima - sustrato.

Para analizar las diferencias estructurales que pudiesen explicar la diferencia de especificidad por sustrato observado en HMPK/PLKs y HMPK, se comparó la capa de coordinación del sustrato aceptor de fosfato utilizando el último cuadro producido

durante la simulación de *St*HMPK y la estructura cristalográfica del complejo ternario piridoxal - ATP de la PLK de *S. aureus* (PDB 4C5N, *Sa*HMPK/PLK) (Figura 13).

Los resultados indican que hay una alta conservación de los residuos del sitio activo. Sin embargo, es posible identificar algunas diferencias que podrían ser claves para explicar por qué *St*HMPK es incapaz de utilizar sustratos derivados de piridoxal. En *St*HMPK, hay un residuo de glutamina (Glu44) ubicado en el *hairpin*- β que forma una interacción polar con el grupo amino y un nitrógeno del anillo pirimidínico, mientras que en *Sa*HMPK/PLK la glutamina es reemplazada por un residuo de metionina. Además, en *Sa*PLK/HMPK el aldehído del piridoxal forma un grupo hemitioacetal con un residuo de cisteína (Cys110) que está ausente en *St*HMPK. La formación de este hemitioacetal es indispensable para que *Sa*HMPK/PLK catalice la fosforilación de PL, pero no afecta la catálisis con PN y PM (Nodwell y cols., 2014).

Sin embargo, se observa que en *St*HMPK hay un mayor número de residuos posicionados a una distancia menor a 5 Å que en *Sa*HMPK/PLK (Figura 13), además, el PL posicionado en el sitio activo de *Sa*HMPK/PLK posee superficies accesibles al solvente, algo que no es esperado para esta clase de sistemas, ya que normalmente las enzimas de la superfamilia riboquinasa excluyen el solvente del sitio activo para prevenir la hidrólisis prematura del sustrato (Sigrell y cols., 1998). Esto se debe a que la estructura cristalográfica de *Sa*HMPK/PLK contiene regiones cercanas al sitio activo que no pudieron ser determinadas en base a la densidad electrónica. Por lo tanto, se realizó una comparación a nivel de secuencia primaria entre HMPKs, PdxK de organismos Gram positivos, PdxK de organismos Gram negativos y PdxY, para intentar determinar otras diferencias en el sitio activo (Figura 14).

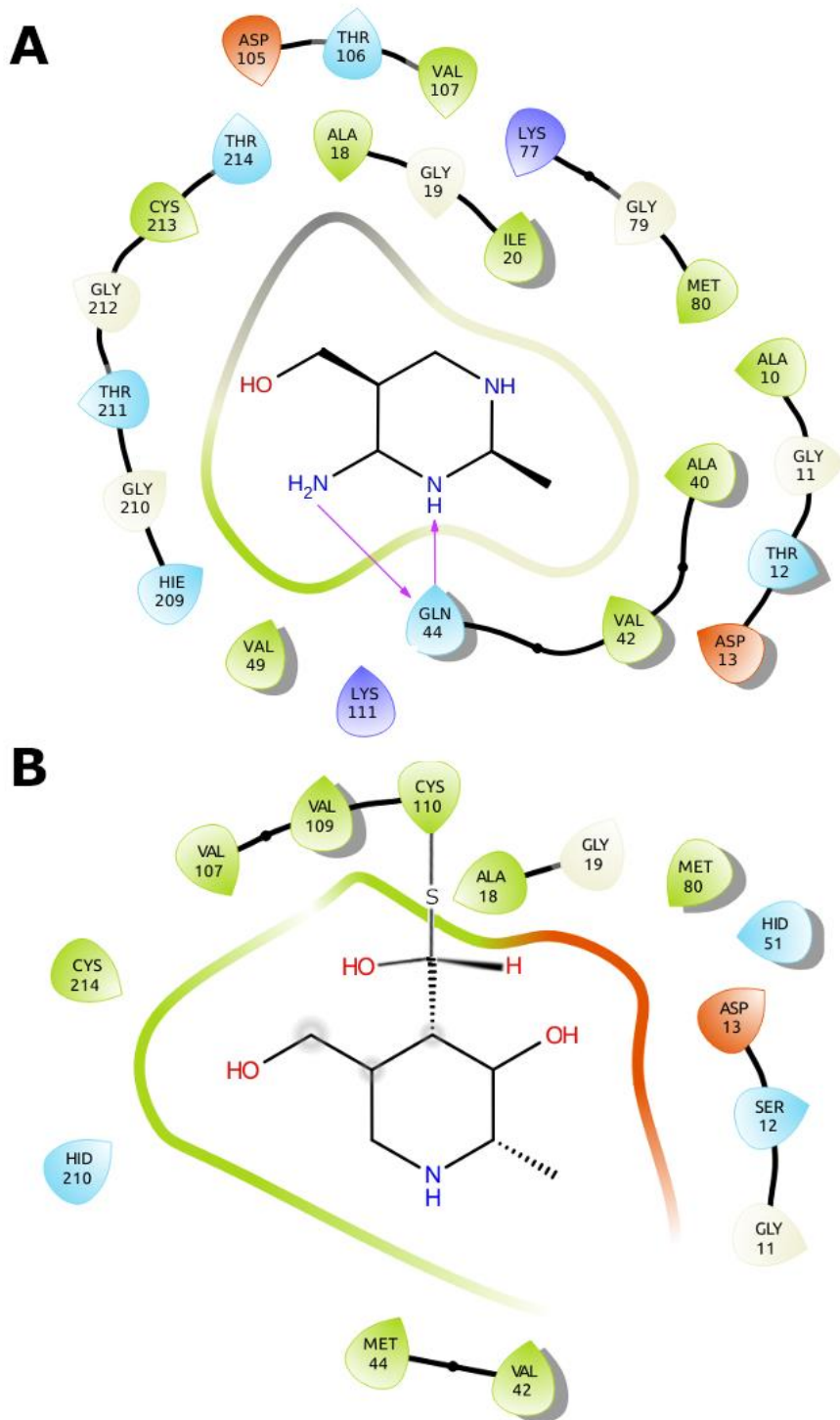


Figura 13. Diagrama 2D de los residuos posicionados a 5 Å del sustrato aceptor del grupo fosfato. A) Sitio activo de *St*HMPK, las flechas moradas representan interacciones polares de tipo puente de hidrógeno. B) Sitio activo de *Sa*HMPK/PLK. Las circunferencias grises representan regiones accesibles al solvente.

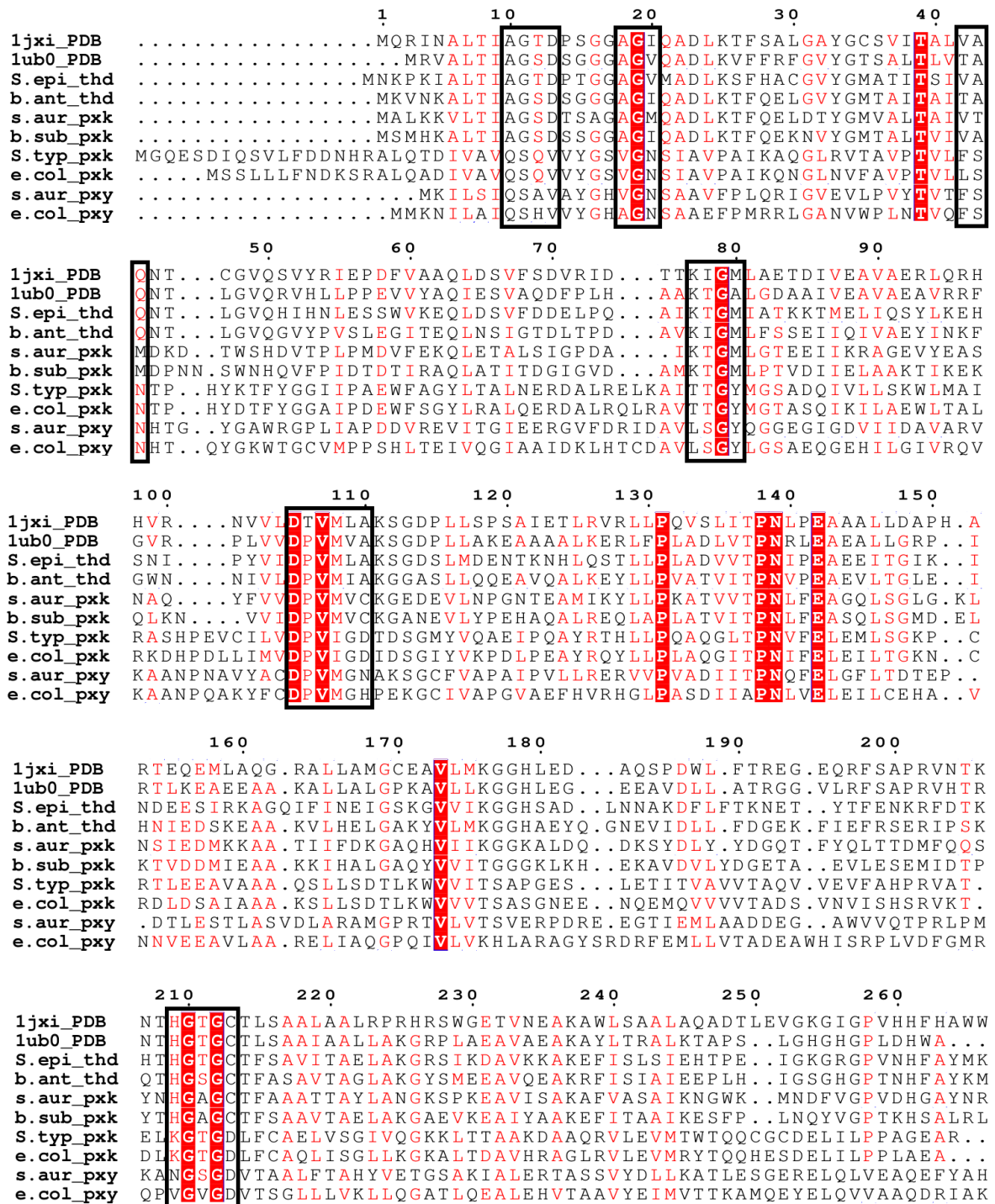


Figura 14. Alineamiento de secuencias múltiples entre las ThiD estudiadas (1jxi_PDB y lub0_PDB), ThiD de organismos Gram positivos (s.epi_thd y b.ant_thd), PdxK de organismos Gram positivos (s.aur_pxx y b.sub_pxx), PdxK de organismos Gram negativos (S.typ_pxx y e.col_pxx), PdxY de un organismo Gram positivo (s.aur_pxy) y PdxY de un organismo Gram negativo (e.col_pxy). Los residuos implicados en la unión del sustrato se muestran en recuadros negros.

La conservación del sitio activo entre las enzimas comparadas ilustra el alto grado de similitud que existe entre las ThiD y las PdxK de organismos Gram positivos (PLK/HMPK), en relación al resto de enzimas con actividad piridoxal quinasa. Pese a esto, el análisis de secuencia primaria revela una segunda diferencia que podría ser relevante en la unión del sustrato y es la presencia de un residuo de treonina ubicado en la posición 211 del alineamiento, justo antes de la cisteína catalítica, la cual es reemplazado por una alanina en las PdxK de organismos Gram positivos e interactúa a través de su cadena lateral con el hidroxilo y el fosfato.

Se requerirán futuros estudios por mutagénesis sitio dirigida para establecer si realmente los residuos propuestos son los responsables de las diferencias en especificidad por sustrato o si otras mutaciones epistáticas también participan en este fenómeno.

3.4 Caracterización cinética de StHMPK y TtHMPK

Para determinar los parámetros cinéticos de estado estacionario, se realizaron medidas de actividad enzimática a distintas concentraciones de sustrato, manteniendo el co-sustrato a concentración fija y saturante. En la figura 15 se muestran las curvas de saturación para HMP de *StHMPK* y *TtHMPK* y en la tabla II, los parámetros cinéticos determinados ajustando los datos a un modelo de Michaelis-Menten.

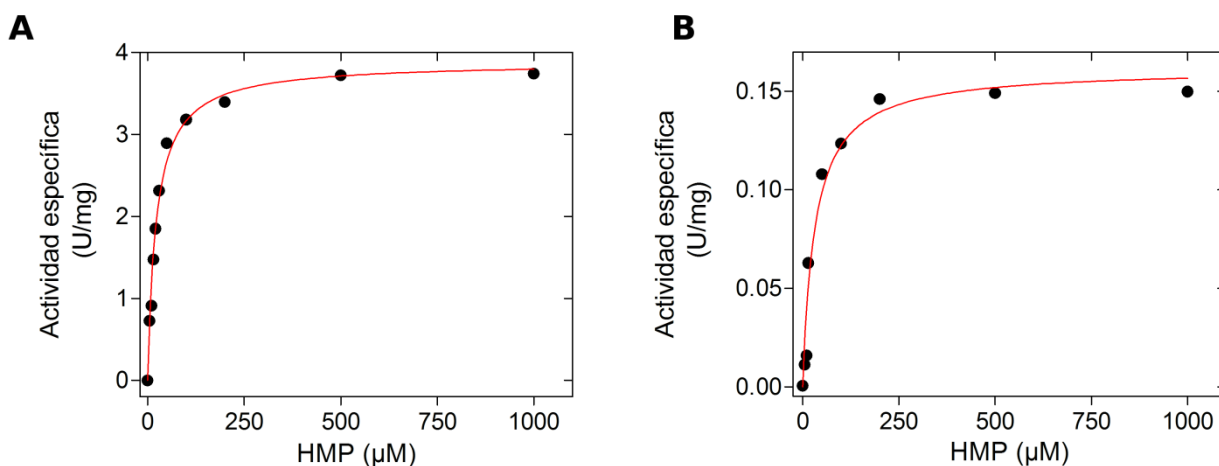


Figura 15. Curvas de saturación de HMP para *StHMPK* (A) y *TtHMPK* (B), empleando 5 mM de ATP-Mg a 37 °C y pH 7. En rojo se muestra el ajuste no lineal del cual se infirieron los parámetros cinéticos.

Tabla II. Parámetros cinéticos para HMP de *StHMPK* y *TtHMPK* y el error estándar del ajuste.

Proteína	K_M (μM)	V_{max} (U/mg)	k_{cat} (s^{-1})	k_{cat} / K_M
<i>StHMPK</i>	23 ± 2	3.9 ± 0.1	0.46 ± 0.04	$2.03 \times 10^4 \pm 2.6 \times 10^3$
<i>TtHMPK</i>	32 ± 7	0.16 ± 0.01	0.078 ± 0.003	$2.43 \times 10^3 \pm 5.4 \times 10^2$

En ambos casos se observa que la K_M para HMP es similar, mientras que la k_{cat} de *StHMPK* es un orden de magnitud mayor que la de *TtHMPK*, a la temperatura a la que se realizaron los experimentos (37 °C).

Por otro lado, las curvas de saturación para ATP de ambas enzimas muestran una clara desviación del modelo Michaelis-Menten (Figura 16), presentando una conducta que se asemeja la superposición de dos hipérbolas. Esto probablemente se deba las dos reacciones de fosforilación que ocurren en el ensayo, si se considera que la enzima

tuviere parámetros cinéticos muy diferentes para las dos reacciones. Los resultados obtenidos no permiten ajustar un modelo del que se puedan determinar constantes cinéticas para ATP en estas enzimas.

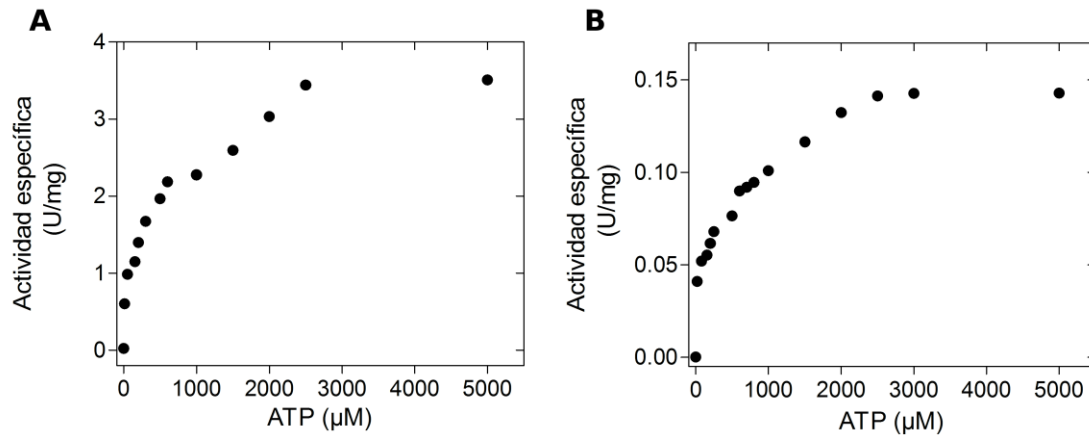


Figura 16. Curvas de saturación para ATP de *StHMPK* (A) y *TtHMPK* (B), utilizando 500 μM de HMP.

3.5 Caracterización biofísica de *StHMPK* y *TtHMPK*

Se determinó el estado de oligomerización de *StHMPK* y *TtHMPK* mediante cromatografía de exclusión molecular, empleando una columna Superdex-75g y eluyendo en condiciones isocráticas. Primero, se calibró la columna utilizando un estándar de masas moleculares conocidas y ajustando una ecuación de decaimiento exponencial (Figura 17). Posteriormente, se determinaron los tiempos de retención de ambas enzimas y se interpolaron en la curva calculada en la figura 17 para estimar masas moleculares. Los resultados se muestran en la tabla III.

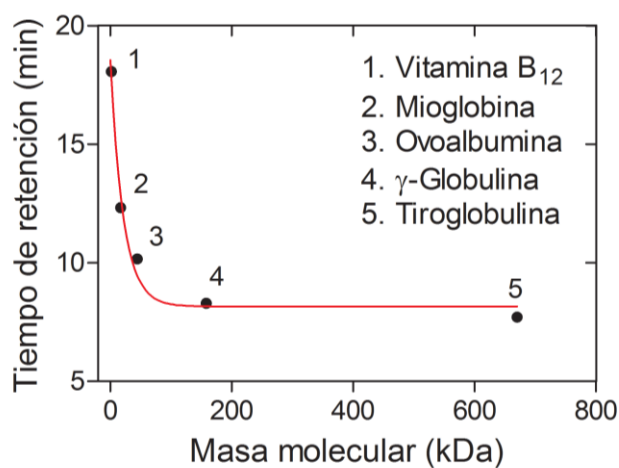


Figura 17. Calibración de la columna de exclusión molecular. 1) Vitamina B₁₂, 1.35 KDa. 2) Mioglobina (caballo), 17 KDa. 3) Ovoalbumina (Pollo), 44 KDa. 4) γ-Globulina (bovino), 158 KDa. 5) Tiroglobulina (bovino), 670 KDa.

Tabla III. Tiempos de retención y masas moleculares estimadas de *St*HMPK y *Tt*HMPK mediante filtración molecular.

Proteína	Tiempo de retención (min)	Masa molecular (kDa)
<i>St</i> HMPK	10.28	33.9
<i>Tt</i> HMPK	10.62	30.7

Dado que las masas moleculares teóricas para el monómero de *St*HMPK y *Tt*HMPK son de 30.9 y 29.1 KDa, respectivamente (Gasteiger y cols., 2005), se concluye que ambas enzimas son monoméricas bajo las condiciones evaluadas. Esto contrasta notablemente con lo reportado para la HMPK de *E. coli*, que fue descrita como

tetrámero y con lo que propone Cheng para *StHMPK* en base a la estructura cristalográfica, donde se indica que se trataría de un dímero.

Posteriormente, dado lo diferente de los nichos ecológicos, se estudió la estabilidad de ambas enzimas frente a desplegamiento por agentes físicos (temperatura) y químicos (cloruro de guanidinio). Se determinó la temperatura media de desplegamiento de *StHMPK* y *TtHMPK* mediante dicroísmo circular, evaluando la pérdida de señal a 222 nm en rampas de temperatura de 1 °C/min y 3 °C/min. Los resultados obtenidos se muestran en la figura 18 y se resumen en la tabla IV.

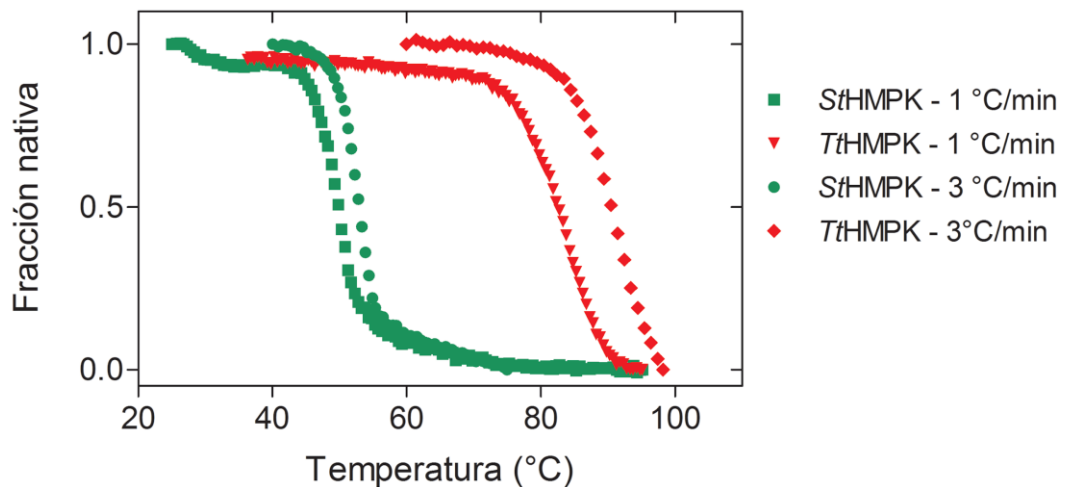


Figura 18. Estabilidad térmica de *StHMPK* (verde) y *TtHMPK* (rojo), en rampas de temperatura con un gradiente de 1 °C/min (círculos para *StHMPK* y rombos para *TtHMPK*) y de 3 °C/min (cuadrados y triángulos para *StHMPK* y *TtHMPK*, respectivamente).

Tabla IV. Valores de la temperatura media de desplegamiento (T_m), para *S*tHMPK y *T*tHMPK con sus respectivos errores estándar del ajuste.

Proteína	T_m a 1 °C/min (°C)	T_m a 3 °C/min (°C)
<i>S</i> tHMPK	50.0 ± 0.1	52.78 ± 0.072
<i>T</i> tHMPK	83.25 ± 0.12	91.40 ± 0.13

En ambos experimentos se registró una diferencia de más de 30 °C entre la T_m de *S*tHMPK y *T*tHMPK, confirmando la alta termoestabilidad de esta última. Además, los resultados indican que el valor de T_m para *T*tHMPK es mucho más sensible al gradiente de temperatura que *S*tHMPK, presentando una variación de ~8 °C, lo que sugiere que este sistema requiere mayor tiempo para alcanzar el equilibrio de desplegamiento.

La estabilidad frente a agentes químicos se evaluó monitoreando los cambios en la fluorescencia intrínseca de triptófanos en presencia de concentraciones crecientes de cloruro de guanidinio. Los espectros de fluorescencia se muestran en la figura 19A y 19B. Para *S*tHMPK (figura 19A) se observa un aumento en la intensidad de fluorescencia emitida a bajas concentraciones de GndCl, llegando a un máximo a 200 mM; posteriormente, la señal decae y la longitud de onda de máxima emisión aumenta. En el caso de *T*tHMPK, también se observa un aumento sostenido en la intensidad de fluorescencia emitida a medida que aumenta la concentración de GndHCl, alcanzando un máximo a 1.2 M, para luego decaer sin corrimientos significativos en la longitud de onda de la emisión máxima.

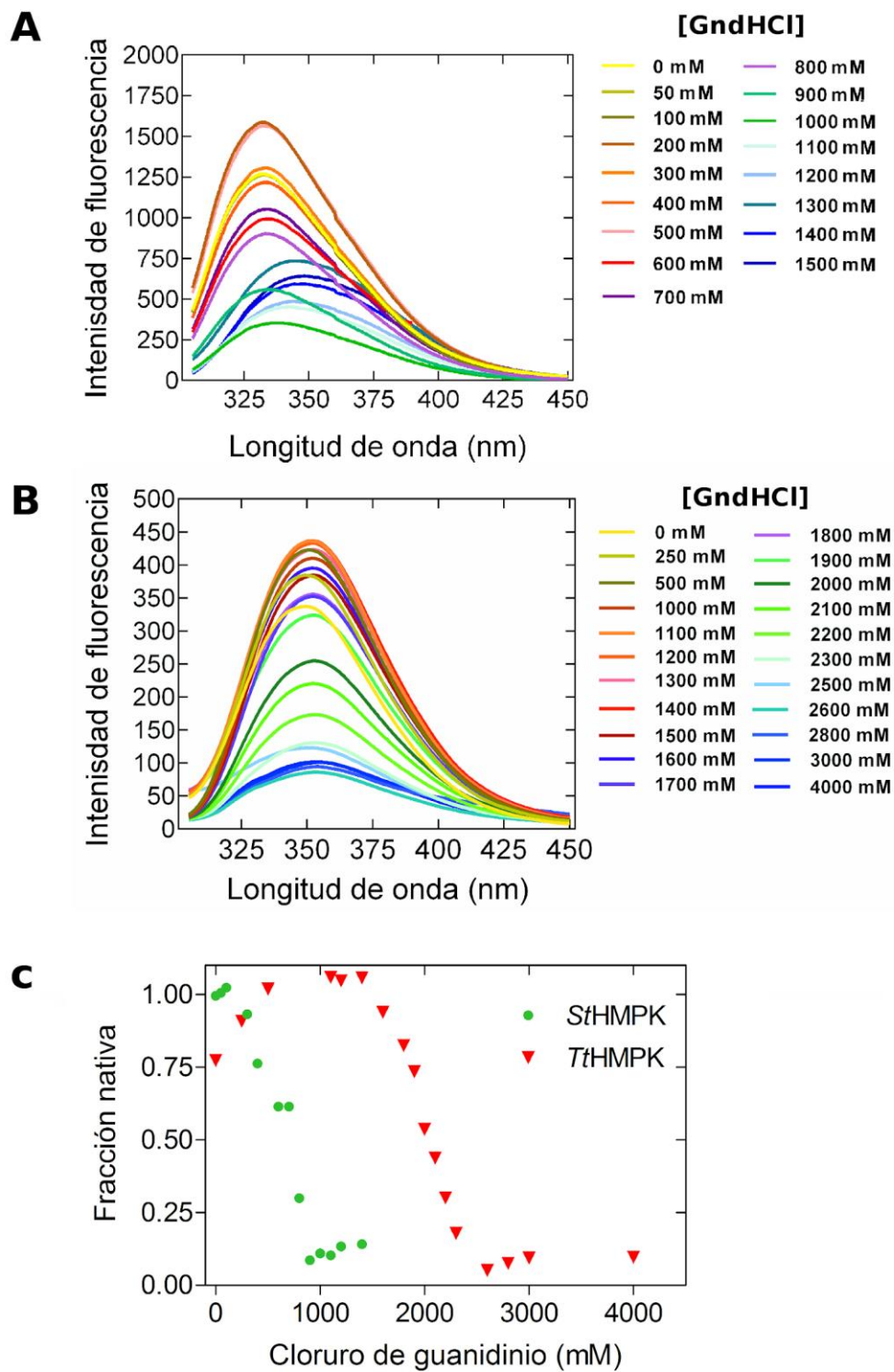


Figura 19. Despliegamiento de *StHMPK* y *TtHMPK* frente a GndCl monitoreada por fluorescencia intrínseca de triptófanos. A) Espectros de fluorescencia de *StHMPK*. B) Espectros de fluorescencia de *TtHMPK*, C) Curva de despliegamiento por GndCl de *StHMPK* (verde) y *TtHMPK* (rojo).

En base a los espectros obtenidos, se determinó la concentración media desnaturante de GndCl (C_m), tomando la intensidad de fluorescencia emitida a 320 nm para *StHMPK* y la intensidad de fluorescencia máxima para *TtHMPK*. Los datos fueron normalizados y ajustados a un modelo de dos estados. En el caso de *TtHMPK*, se ignoraron los 3 primeros puntos para el ajuste. Los resultados obtenidos se muestran en la figura 19C y las C_m s calculadas se indican en la tabla V.

Tabla V. Valores de C_m para *StHMPK* y *TtHMPK*, con sus respectivos errores estándar de ajuste.

Proteína	C_m (mM)
<i>StHMPK</i>	660 ± 43.0
<i>TtHMPK</i>	2000 ± 14.7

Al igual que en los experimentos de desplegamiento por temperatura, los resultados obtenidos indican que la proteína proveniente de *T. thermophilus* es significativamente más estable, presentando un valor de C_m de al menos 3 veces mayor que el de su homólogo mesófilo.

3.6 Análisis estructural de StHMPK y TtHMPK.

Como era de esperar, los resultados mostrados en la sección anterior revelan notables diferencias en la estabilidad de la HMPK proveniente de un organismo mesófilo, en relación a su homólogo termófilo. Dado que ambas enzimas cuentan con

estructuras determinadas por cristalografía de rayos X, éstas fueron utilizadas como base para intentar pesquisar los atributos estructurales responsables de las diferencias en estabilidad observadas.

Las proteínas son polímeros altamente dinámicos, por lo que la información extraíble de estructuras estáticas es limitada (Karplus & McCammon, 2002). En solución, las proteínas existen como un ensamble de conformaciones distintas, donde pueden ocurrir distintos rearrreglos estructurales restringidos por las interacciones covalentes y no covalentes presentes en la proteína y que transcurren en escalas temporales que van desde los picosegundos a los milisegundos. Para estudiar las propiedades dinámicas de *St*HMPK y *Tt*HMPK, se realizaron simulaciones de dinámica molecular para las proteínas en estado apo empleando el campo de fuerza AMBER *ff99SB*. Cada sistema fue simulado por 50 ns en el ensamble isotérmico - isobárico (1 atm y 300 K) considerando los 10 primeros nanosegundos como fase de equilibrado y los 40 nanosegundos posteriores como fase productiva.

De las trayectorias resultantes se pueden obtener dos parámetros importantes; el RMSD que da cuenta del movimiento de la cadena polipeptídica, expresado como la distancia cuadrática media de los $C\alpha$ entre las distintas conformaciones y el RMSF que expresa la fluctuación cuadrática media de los $C\alpha$ (Figura 19).

El RMSD de ambas proteínas muestra que convergen a estados estables transcurridos ~10 ns de simulación, con la diferencia que las magnitudes de las variaciones de RMSD son alrededor de 1 Å mayores para *St*HMPK (Figura 20A). Cuando se evalúa el RMSF, se observa que si bien el perfil de fluctuación es

prácticamente idéntico entre ambas proteínas, la intensidad de dichas fluctuaciones varía drásticamente, siendo mayores las de *St*HMPK (Figura 20B). Estos resultados apuntan hacia la existencia de diferencias en la flexibilidad local de las proteínas, siendo *Tt*HMPK mucho más rígida que su homólogo mesófilo. Al mapear las diferencias de fluctuaciones en las estructuras de las proteínas (figura 21), se observa que en ambos casos, las fluctuaciones se concentran en los *loops* cercanos al sitio activo, y que éstas son mucho mayores en *St*HMPK.

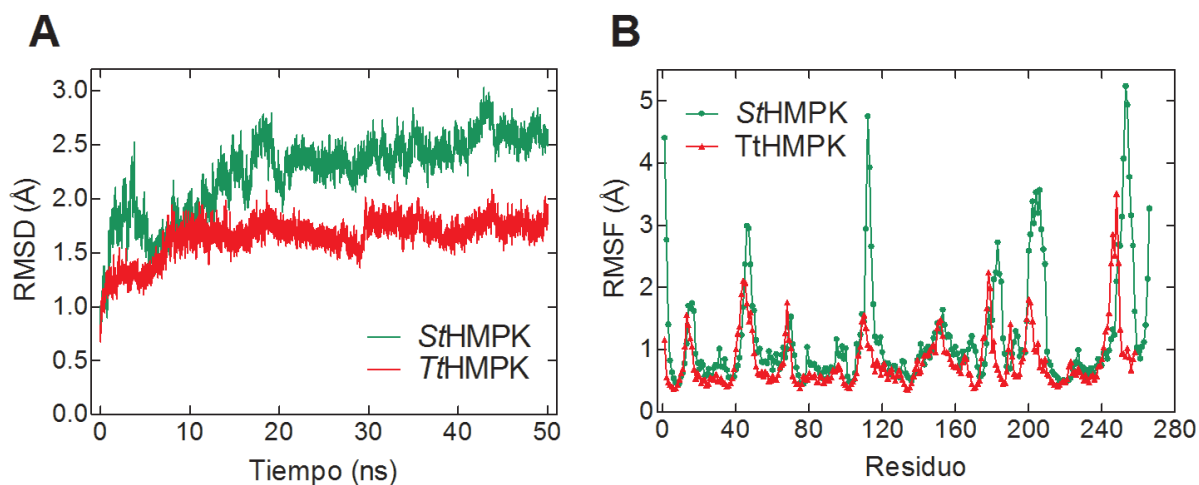


Figura 20. RMSD (A) y RMSF (B) para *St*HMPK (verde) y *Tt*HMPK (rojo) resultantes de 50 ns de simulación

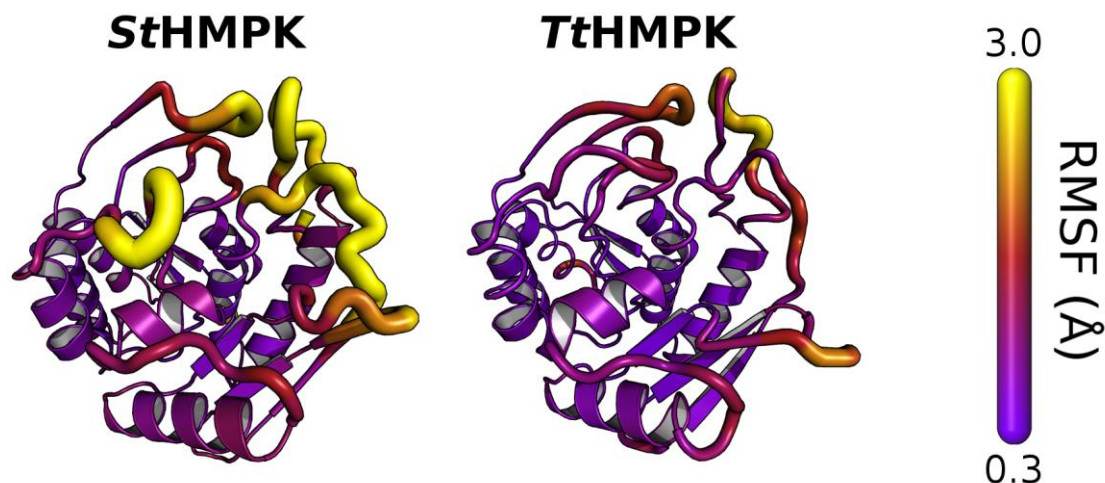


Figura 21. Diferencias en las fluctuaciones estructurales de *StHMPK* y *TtHMPK*, mapeadas en las estructuras de mínima energía. El grosor de los *loops* y el color ilustran la magnitud de las fluctuaciones estructurales.

Para tener una mayor robustez de los datos acerca de los cambios en flexibilidad observados en las proteínas, se calcularon matrices de correlación cruzada de residuos. Las matrices de correlación cruzada permiten detectar los movimientos concertados que ocurren al interior de la estructura de una proteína, en base al sentido de los vectores de desplazamiento de los $C\alpha$ (Hünenberger y cols., 1995). Valores positivos en la matriz indican movimientos correlacionados, lo que implica que los $C\alpha$ se mueven en el mismo sentido y sugieren la presencia de interacciones atractivas, mientras que valores negativos indican que los $C\alpha$ se mueven en sentidos opuestos lo que sugiere la existencia de interacciones repulsivas, y valores cercanos a 0 indican que no existe correlación de movimientos. Los resultados de las matrices de correlación para *StHMPK* y *TtHMPK* muestran diferencias significativas en la presencia de movimientos

concertados entre ambas proteínas (Figura 22). *Tt*HMPK posee una mayor cantidad de movimientos acoplados, indicado por la mayor cantidad de zonas cian en la matriz, además, existe una reducción notoria en la cantidad de movimientos anti - correlacionados en esta proteína, ilustrados en rosa. Si se observa la diagonal de la matriz, se evidencia que en *Tt*HMPK hay una mayor área con valores positivos de correlación, lo que implica que hay un mayor acople en los movimientos de los residuos adyacentes en el polipéptido, lo que reafirma la idea de que la proteína termoestable es mucho más rígida que su homólogo proveniente de un organismo mesófilo.

Los resultados de cálculos de RMSF y matrices de correlación cruzada indican fuertemente la existencia de diferencias significativas en la flexibilidad conformacional entre ambas enzimas. Esto implica que deben existir diferencias en la red de interacciones no covalentes presentes en la estructura de ambas proteínas que expliquen las diferencias en la flexibilidad y estabilidad de proteínas. Múltiples estudios en relación a la termoestabilidad de proteínas de organismos extremófilos han reportado que existen diferencias sustanciales en las interacciones no covalentes presentes en una proteína termófila en relación a homólogos mesoestables, donde destacan un aumento en la cantidad de interacciones electrostáticas, aumento en el volumen del núcleo hidrofóbico y en el número de puentes de hidrógenos (Pucci & Rooman, 2017).

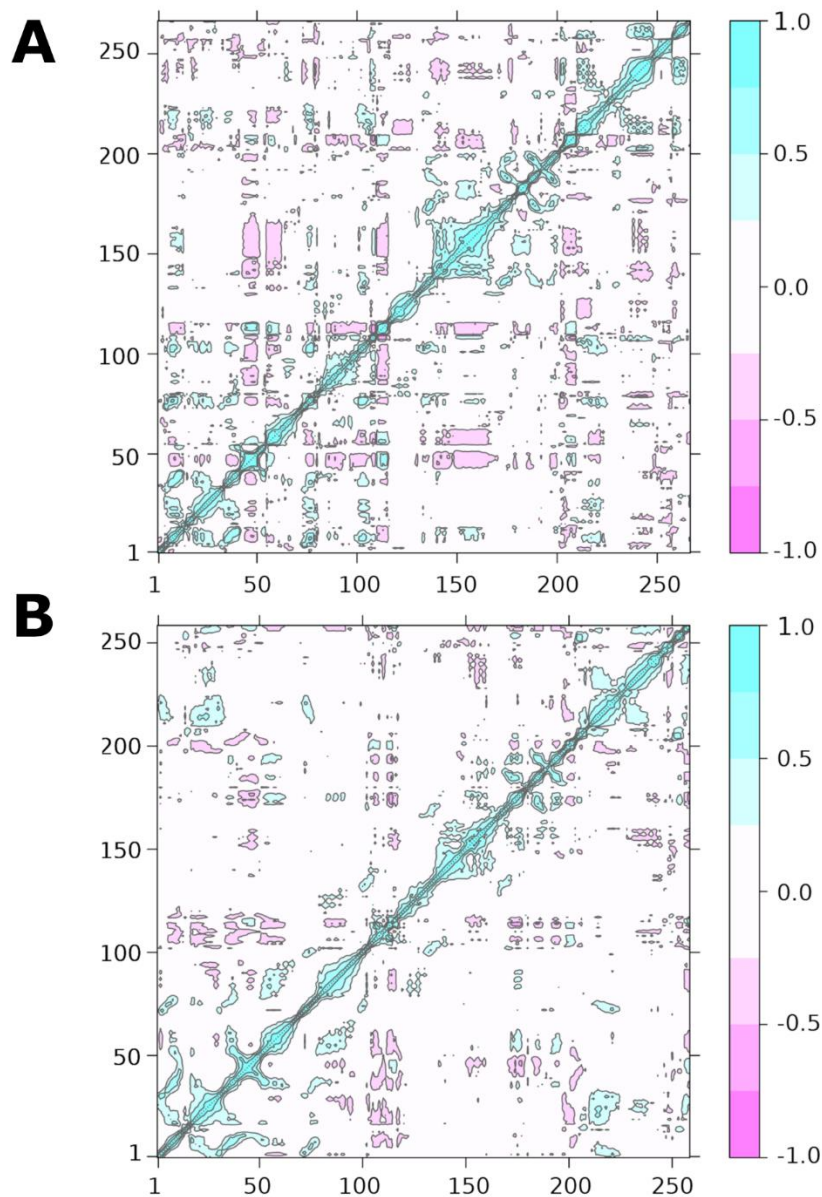


Figura 22. Matrices de correlación cruzada de residuos de A) *StHMPK* y B) *TtHMPK*. Las correlaciones positivas se muestran en cian y las correlaciones negativas se muestran en rosado.

Por lo tanto, se analizaron las interacciones intramoleculares presentes en cada proteína, con el objetivo de determinar variaciones estructurales que pudiesen ser

responsables de las diferencias de flexibilidad observadas. Para esto, primero se determinaron las estructuras representativas de los ensambles conformacionales explorados durante los 40 ns en equilibrio. Esto se realizó mediante análisis de *clusters*, definiendo un RMSD $C\alpha > 1 \text{ \AA}$ como una variación estructural significativa y un tiempo de vida mínimo de 125 ps por ensamble. Para la simulación de *StHMPK* se obtuvieron 4 estructuras que representan 4 ensambles conformacionales distintos, de los cuales el principal ocupa 45% del tiempo de simulación, otro el 14.3% y los últimos 2 menos del 20%. Por otra parte, en la simulación de *TtHMPK* existe un único ensamble conformacional, en el que la proteína se mantiene por el 99.6% del tiempo de la simulación, lo que reafirma que esta proteína es mucho más rígida que su homólogo mesófilo.

Uno de los aspectos estructurales que varía significativamente entre proteínas mesófilas y termófilas es el núcleo hidrofóbico, el cual suele ser más voluminoso en estas últimas (Carter y cols., 2001). Utilizando la estructura representativa del ensamble conformacional más poblado de *StHMPK*, y la estructura representativa de *TtHMPK*, se analizó la distribución de residuos en el núcleo de éstas. Se definió como residuos pertenecientes al núcleo todos aquellos con una superficie molecular expuesta $< 5 \text{ \AA}$. Los resultados obtenidos muestran que no hay diferencias evidentes en el volumen del núcleo hidrofóbico (figura 23A y 23B) como en otros casos reportados en literatura. Sin embargo, cuando se analiza la composición aminoacídica del núcleo, se aprecia que *StHMPK* posee una mayor cantidad de residuos polares que *TtHMPK*. Este tipo de aminoácidos constituye el 18.3 % del total de residuos en el núcleo de *StHMPK* y sólo un 10.5 % en *TtHMPK*. En esta proteína, se observa

concomitantemente un aumento significativo en el contenido de alanina. lo que sugiere un aumento en la hidrofobicidad del núcleo en la proteína termófila.

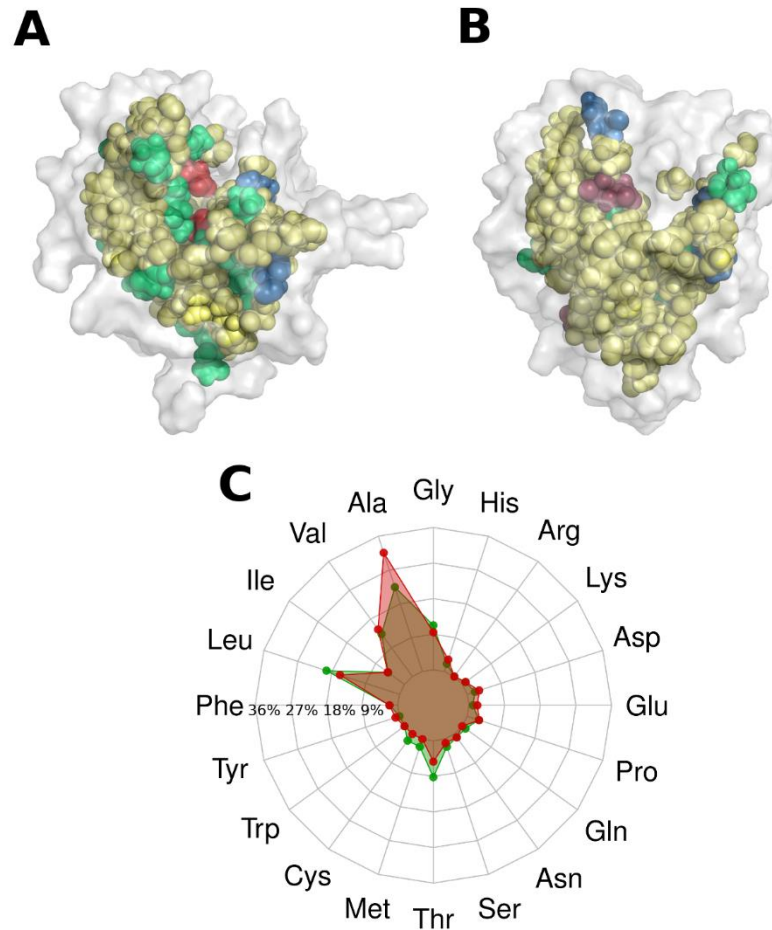


Figura 23. Diferencias en el núcleo hidrofóbico de *StHMPK* y *ThHMK*. A) Núcleo de *StHMPK*. Se muestra la superficie molecular en gris y los residuos pertenecientes al núcleo se representan como esferas de van der Waals. Los residuos hidrofóbicos (Ala, Gly, Ile, Leu, Tyr, Phe, Trp y Met) se muestran en amarillo, los residuos polares (Cys, Thr, Ser, Asn, Gln y Pro) se muestran en verde, los residuos con carga parcial positiva (His, Arg y Lys) se muestran en azul y los residuos con carga parcial negativa (Asp y Glu) se muestran en rojo. B) *Ídem* que en A, pero para *ThHMPK*. C) Composición aminoacídica del núcleo de las proteínas. En verde se muestra la composición de *StHMPK* y en rojo la de *ThHMPK*.

Posteriormente, se analizaron las interacciones electrostáticas presentes en ambas proteínas. Primero, se analizaron las interacciones existentes en las estructuras representativas de los ensambles conformacionales de mayor duración empleando el servidor *Protein Interaction Calculator* (PIC) (Tina y cols., 2007), estableciendo como distancia de corte 3.5 Å (figura 24A y 24B). En base a este análisis se encontraron 12 pares interactuantes en la estructura de *StHMPK* y 18 en la de *TtHMPK*. Un hecho interesante es que las nuevas interacciones electrostáticas presentes en la proteína termófila aparecen en la cercanía de interacciones conservadas, reforzando la red iónica en la estructura. Al analizar los pares iónicos a lo largo de toda la simulación, (definidos como interacciones entre cadenas laterales de residuos de carga opuesta a menos de 3.5 Å de distancia y persistentes por más del 5 % del tiempo estable de la simulación) hay un aumento notorio en la cantidad de interacciones transitorias en *TtHMPK*, pero no hay diferencias en el tiempo de vida promedio de éstos (figura 24C), por lo que la flexibilidad de *TtHMPK* no sólo está restringida por los pares mostrados en la figura 23B, sino que también por interacciones de corta duración.

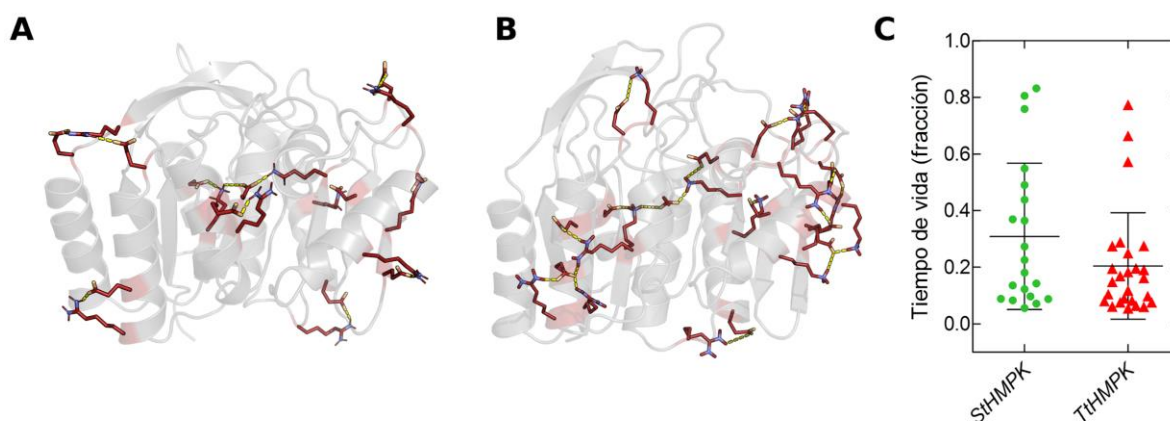


Figura 24. Diferencias en las interacciones electrostáticas entre la proteína mesófila y termófila. A) Estructura de *StHMPK*, los residuos interactuantes se destacan en rojo. B) *Ídem* que en A, pero para *TtHMPK*. C) Pares interactuantes en *StHMPK* (verde, n= 20) y *TtHMPK* (rojo, n= 26) y tiempos de vida (como fracción del tiempo de simulación) de cada uno.

Finalmente, se analizó la red de puentes de hidrógeno a lo largo de toda la dinámica. Se consideraron sólo los puentes de hidrógeno formados donde el ángulo descrito por los átomos dador – hidrógeno – aceptor era $> 120^\circ$, la distancia entre el hidrógeno y el aceptor era $< 3.5 \text{ \AA}$ y que persistieron por al menos el 5 % de los 40 ns estables de simulación. Además, se separaron según si los pares dador – aceptor son parte de la cadena principal del polipéptido o de las cadenas laterales de los aminoácidos. Los resultados obtenidos se muestran en la tabla VI.

Tabla VI. Interacciones por puentes de hidrógeno en *StHMPK* y *TtHMPK*.

Proteína	P. de H. Cadena principal				P. de H. Cadena lateral			
	Tiempo de vida (ns)				Tiempo de vida (ns)			
	N°	\bar{X}	Mediana	SD	N°	\bar{x}	Mediana	SD
<i>StHMPK</i>	280	22.7	26.2	14.4	145	10.2	7.1	8.7
<i>TtHMPK</i>	255	24.4	27.5	14.1	152	8.7	6.4	7.4

Se indica la cantidad total de pares dador – aceptor interactuantes (N°), el promedio de duración de las interacciones (\bar{x}), la mediana de la duración de las interacciones y la desviación estándar (SD).

En ambas proteínas hay una tendencia similar a concentrar una mayor cantidad de interacciones entre átomos de la cadena principal que entre átomos de cadenas laterales, y que estas últimas parecen ser más lábiles, ya que poseen menores tiempos de vida. *StHMPK* posee una cantidad levemente superior de puentes de hidrógeno totales que *TtHMPK*, lo que dista de otros reportes de enzimas termófilas (Vogt & Argos, 1997), aunque esto puede deberse a que la primera es levemente más grande que su homóloga (266 residuos contra 258). Sin embargo, no existen diferencias drásticas entre la cantidad o la persistencia de las interacciones por puente de hidrógeno entre la proteína termófila y la mesófila que pudiesen ayudar a explicar las diferencias de termoestabilidad.

4. Discusión

4.1 Doble fosforilación consecutiva y especificidad por sustrato en HMPKs

Los resultados expuestos en los primeros puntos de este seminario de título demuestran que las HMPKs de *S. typhimurium* y *T. thermophilus* son capaces de fosforilar HMP y HMP-P a expensas de ATP, tal como su homólogo de *E. coli*. Por lo que pueden ser consideradas como enzimas bifuncionales, dada la diferencia de los grupos químicos que fosforila. La baja identidad de secuencia existente entre estas enzimas (< 50 %) y la disparidad de los nichos en que habitan sus respectivos organismos sugieren que ha transcurrido un largo tiempo de divergencia evolutiva entre ambos, tal como muestra el árbol filogenético construido, pero pese a esto, los resultados apuntan a que los mecanismos estructurales que sustentan la doble fosforilación, y que permiten catalizar la transferencia de fosfatos sobre dos grupos funcionales distintos han sido conservados durante la evolución de los distintos linajes bacterianos de esta enzima.

La bifuncionalidad catalítica no es un fenómeno aislado dentro de la superfamilia riboquinasa. Tradicionalmente, se ha definido la bifuncionalidad de enzimas como la capacidad de realizar dos actividades metabólicamente relevantes, y por ende, la pérdida de una de ellas afecta negativamente el *fitness* del organismo (Allewell, 2012). Como antecedente dentro de la superfamilia, se tiene el caso de las quinasas de azúcares ADP-dependientes que participan en la vía modificada de Embden-Meyerhof presentes en el orden *Methanococcales*, las cuales poseen actividad glucoquinasa y fosfofructoquinasa

(Castro-Fernández y cols., 2014). Por otra parte, en las quinasas de cofactores enzimáticos, la piridoxal quinasa de *B. subtilis* y *S. aureus* son capaces de fosforilar piridoxal y HMP, siendo la primera actividad relevante esta última reacción relevante para el salvataje de tiamina en este organismo. Si bien la HMPK lleva a cabo dos funciones catalíticas independientes que son indispensables para la supervivencia del organismo y por lo tanto debiese considerarse una enzima bifuncional, su capacidad de catalizar dos fosforilaciones consecutivas difiere de los ejemplos mencionados anteriormente en dos aspectos medulares: En primer lugar, la doble fosforilación de las HMPKs parece ser un carácter estrictamente conservado a lo largo de la evolución de todas las taxas bacterianas, a diferencia de la bifuncionalidad metabólica de las quinasas ADP-dependientes, que sólo está presente en los órdenes *Methanococcales* y *Methanosarcinales*, y la bifuncionalidad observada en las piridoxal quinasas, que se acota a clados específicos que emergieron recientemente a partir de HMPK, como muestra la reconstrucción filogenética. Esto sugiere que debe existir una ventaja adaptativa clara en concentrar ambos pasos del metabolismo de la biosíntesis de tiamina en una sola enzima. Los valores de k_{cat} derivados a partir de las curvas de saturación de HMP son relativamente bajos en relación al resto de las enzimas de la superfamilia riboquinasa, un hecho que ya ha sido reportado en otras enzimas de la familia de quinasas de cofactores enzimáticos y que se especula que estaría asociado a la alta reactividad de los metabolitos pertenecientes a estas vías, por lo que bajas tasas de conversión catalítica evitarían la acumulación excesiva de éstos (Toney., 2005). De acuerdo a lo anterior, es posible que la bifuncionalidad de las HMPKs permita mantener

un flujo metabólico eficiente mediante un aumento local en la concentración de sustratos disponibles, sin necesidad de aumentar las constantes catalíticas.

El segundo aspecto que diferencia a las HMPKs de las otras enzimas bifuncionales de la superfamilia riboquinasa, se relaciona con el grupo funcional que actúa como aceptor de fosfato. En el caso de la actividad glucoquinasa/fosfofructoquinasa observada en las quinasa ADP-dependientes los sustratos aceptores de fosfato son isómeros, pero que difieren en el grado de fosforilación previa. En ambos sustratos la fosforilación catalizada por la enzima se produce en un hidroxilo que ocupa una posición estereoquímica equivalente, por lo tanto, no se requieren mayores rearrreglos de la geometría del sitio activo para acomodar fructosa-6-fosfato o glucosa, y de hecho una sustitución aminoacídica puntual es responsable de la viabilidad de la unión de ambos sustratos y de transformar una enzima específica en bifuncional (Castro-Fernández y cols., 2017). De manera similar, la HMPK/PLK homóloga acomoda sustratos relativamente similares y fosforila hidroxilos en posiciones equivalentes.

El caso de las HMPKs es más complejo que los mencionados anteriormente, dadas las implicancias fisicoquímicas y estructurales que subyacen a su bifuncionalidad. Si bien entre ambos sustratos la única diferencia química observada es la presencia de un fosfato en lugar de un hidroxilo, es precisamente en esta región distinta donde ocurre el proceso catalítico. Esto implica que grupos con características estereoquímicas y distribuciones de carga sumamente distintas, son capaces de unirse a la misma región adyacentes al fosfato γ del ATP y a la cisteína catalítica (Cys213 en *S*tHMPK), cosa de

mantener las distancias de transferencia necesaria para la catálisis. Dado que el fosfato es un grupo de volumen considerable, el reacomodo de los residuos catalíticos en torno a los sustratos no es trivial y los detalles concretos de cómo ocurre esto aún son desconocidos. En el trabajo realizado por Cheng y cols., se propone un mecanismo catalítico en base a la posición de sulfatos presentes en la estructura cristalográfica de *Sr*HMPK. Este implica la participación de un bolsillo aniónico en la fosforilación del HMP-P formado por los residuos H209, G210, T211, G212 y C213 y el reacomodo conformacional del ATP, según si se fosforila HMP o HMP-P, lo que mantendría la geometría lineal entre el grupo dador y el receptor de fosfato necesaria para la reacción SN_2 , pero, no existe evidencia experimental ni teórica referente a esta proposición. Sin embargo, los resultados del alineamiento de secuencias múltiples parecen indicar que la treonina propuesta por Cheng y cols. podría tener un rol en la doble fosforilación, ya que es estrictamente conservada en HMPKs, mientras que en HMPK/PLKs (que sólo son capaces de fosforilar una vez) este residuo está sustituido por una alanina, lo que constituye un cambio significativo en la naturaleza química de la cadena lateral.

La idea de que el ciclo catalítico de las HMPKs requiere de distintos estados conformacionales del complejo enzima-ATP implica la posibilidad que existan distintas afinidades por ATP, lo cual parece ser apoyado por la forma inusual obtenida en la curva de saturación de ATP. Esta hipótesis deberá ser probada mediante caracterizaciones cinéticas más detalladas, por ejemplo, realizando curvas de saturación por HMP-P, lo cual no es posible en nuestro laboratorio, ya que requeriría la síntesis orgánica de este compuesto. Otra alternativa sería realizar curvas de saturación a distintas

concentraciones de HMP para estudiar cómo cambia la forma de la curva, ya que si uno de los pasos es limitante con respecto al otro, se espera que predomine uno de los componentes de la doble hipérbola observada cuando el HMP está en concentración saturante.

Otro aspecto interesante referente a los resultados de la primera parte de este seminario se relaciona con la alta especificidad por HMP que muestran ambas enzimas estudiadas. Tanto *St*HMPK como *Tt*HMPK resultaron ser sumamente específicas por HMP, lo cual resulta inesperado si se consideran los resultados de los análisis filogenéticos, que muestran que las HMPKs están estrechamente relacionadas con el grupo de enzimas bifuncionales HMPK/PLK. Estudios realizados en el laboratorio, mediante la metodología de reconstrucción de secuencias ancestrales muestran que la capacidad de fosforilar PL observada en las enzimas HMPK/PLK de organismos como *Bacillus* y *Staphylococcaceae* correspondería a una novedad adaptativa, ya que el ancestro hipotético entre ambos grupos enzimáticos era específico por HMP (Castro-Fernández y cols., 2014). Al comparar los sitios activos de ambas enzimas resulta evidente el alto grado de similitud entre ellos, siendo las diferencias más significativas: 1) la presencia de una glutamina en el *hairpin*- β en las HMPKs que es sustituido por una metionina en las HMPK/PLKs, 2) la presencia de una treonina en la cercanía a la cisteína catalítica en HMPKs, la cual podría formar interacciones por puente de hidrógeno con el sustrato, mientras que en HMPK/PLKs hay una alanina, 3) La pérdida de la cisteína en la posición 110 necesaria para la formación del hemitioacetal en HMPKs. La interacción polar generada por la glutamina 44 con el amino del HMP

parece ser clave para explicar la especificidad de estas enzimas, sobretodo si se considera que las proteínas ancestrales específicas por HMP analizadas en el estudio de Castro-Fernández y cols. presentan este residuo, y además, el alineamiento de secuencias múltiples muestra que está estrictamente conservada en todas las HMPKs. Sin embargo, se ha descrito exhaustivamente que los cambios de especificidad por sustrato pueden deberse no sólo a mutaciones en el sitio activo, sino que también existe un fuerte componente epistático, por lo que mutaciones que afecten la dinámica del sitio activo o los cambios conformacionales requeridos para la catálisis pueden modular la especificidad por un sustrato o ligando (Starr & Thornton, 2016; Wilson y cols., 2015). Por lo tanto, se requerirá realizar estudios mutacionales para validar si los aminoácidos propuestos son suficientes para explicar los cambios de especificidad por sustrato o si hay participación de interacciones epistáticas.

4.2 Biofísica y biología estructural de las HMPKs

Uno de los resultados inesperados de este trabajo fue el estado de oligomerización observado para ambas enzimas. La HMPK de *E. coli* está reportada como tetramero, y análisis de interfase de contacto mediante el servidor web PISA de la estructura cristalográfica de *St*HMPK sugieren que es dimérica (Krissinel & Kenrick, 2007), mientras que los experimentos de cromatografía de exclusión molecular apuntan a que *St*HMPK y *Tt*HMPK son monómeros. Experimentos recientemente llevados a cabo en laboratorio muestran que la cola de polihistidina empleada en el proceso de purificación no tiene efectos en los resultados observados (datos no incluidos en este seminario), lo que descarta artefactos que pudiesen ser introducidos por estos residuos.

Cabe la posibilidad de que esta enzima sea capaz de formar estados de oligomerización superior, pero que sean poco estables y se produzca la disociación durante la cromatografía. Esto deberá corroborarse mediante otras metodologías que permitan calcular radio hidrodinámico, como SAXS o DLS.

Los resultados obtenidos en la evaluación de la estabilidad de ambas enzimas fueron consistentes con lo esperado en base al nicho ecológico en que habitan los organismos, existiendo más de 30 °C de diferencia entre la T_m de la proteína mesófila y termófila, y más de 1 M de GndCl entre la C_m de ambas. Uno de los aspectos más interesantes de este resultado es el impacto que tiene la velocidad de gradiente en la rampa de temperatura en *StHMPK* y *TtHMPK*, dado que en esta última, la T_m determinada en la rampa a 3 °C/min es significativamente mayor que la T_m determinada a 1 °C/min, mientras que para la primera, la diferencia es menos acentuada. El desplazamiento de la T_m a valores más altos producto de aumentos en la velocidad del gradiente de la rampa se produce por el retraso de la muestra en alcanzar el equilibrio entre el estado nativo y el estado plegado, lo que hace que la señal dicroica sea mayor que la que debiese ser realmente para esa temperatura según la ecuación de Gibbs ($\Delta G = RT \ln K_{eq}$, donde ΔG es la diferencia de energía libre entre el estado plegado y desplegado, R es la constante de los gases, T es la temperatura y K_{eq} es la constante de equilibrio entre los estados). Esto implica que *TtHMPK* requiere un mayor tiempo para alcanzar el equilibrio entre el estado nativo y el estado desplegado, lo cual es relevante cuando se busca explicar las diferencias de estabilidad con su homólogo mesófilo. Tiempos lentos de equilibrado son indicativos de cinéticas de desplegamiento lentas, las

que a su vez, sugieren la existencia de una mayor barrera energética entre el estado nativo y el estado desplegado. Últimamente, la barrera energética del desplegamiento de proteínas se ha posicionado como uno de los factores claves para explicar cómo la naturaleza, a través del proceso de selección natural, ha generado proteínas capaces de tolerar temperaturas extremas (Lim y cols., 2017).

Finalmente, para explicar las diferencias de estabilidad observadas a nivel atómico, se realizaron simulaciones de dinámica molecular empleando como base las estructuras cristalográficas de *St*HMPK y *Tt*HMPK. Los análisis estructurales a partir de las trayectorias de dinámica molecular de estas enzimas revelaron que existe una diferencia significativa en la movilidad de los *loops* adyacentes al sitio activo, siendo mucho más rígidos en *Tt*HMPK y que presenta una mayor cantidad de movimientos concertados en su estructura. Además, se observó que esta proteína puebla un único ensamble conformacional por casi todo el tiempo de simulación, mientras que su homólogo mesófilo transita entre 4 estados conformacionales con diferencias estructurales considerables. La existencia de diferencias de flexibilidad entre ambas estructuras es consistente con estudios previos de termoestabilidad proteica, ya que una mayor rigidez compensa la alta energía cinética de los medios de alta temperatura, por lo que es posible mantener el andamiaje catalítico en la conformación necesaria para realizar la reacción (Åqvist y cols, 2017). Esto además explica la baja actividad observada en proteínas termoestables a temperaturas por debajo de las del crecimiento óptimo de sus organismos, dado que una estructura muy rígida requiere de una mayor

temperatura para que se produzcan los movimientos conformacionales requeridos para el proceso catalítico (Nguyen y cols., 2017)

El estudio de las interacciones apunta a que la alta termoestabilidad observada en la proteína de *T. thermophilus* estaría dada por un componente entrópico y un componente entálpico. El primero, se relaciona con las diferencias observadas en el núcleo hidrofóbico de las proteínas. En *TtHMPK*, se observó una disminución en el contenido de aminoácidos polares y un aumento en la cantidad de alanina, lo que se traduce en un aumento en la hidrofobicidad del núcleo, comparado con *StHMPK*. Este aumento en la hidrofobicidad del núcleo de *TtHMPK* implica que durante el proceso de desplegamiento debe exponerse una superficie apolar mayor, lo que tiene un costo entrópico más alto que lo que ocurriría con *StHMPK*, donde el núcleo rico en residuos polares puede formar interacciones por puente de hidrógeno con el solvente, reduciendo el costo energético de la transición de plegado a desplegado. Por otra parte, también existe un aporte entálpico considerable a la estabilidad superior de *TtHMPK*, dado por un incremento en la cantidad de interacciones electrostáticas, que son evidenciables al comparar estructuras representativas de los ensambles conformacionales más persistentes en el tiempo de cada proteína. Además, el aumento de interacciones transitorias en *TtHMPK* puede generar restricciones significativas a los grados de libertad de los aminoácidos, disminuyendo la flexibilidad del polipéptido y por ende, aumentar su termoestabilidad. Interesantemente, no se observan diferencias significativas en la red de puentes de hidrógeno, pese a que se ha reportado que muchas proteínas termoestables tienen una mayor cantidad de este tipo de interacciones (Pucci &

Rooman, 2017). Cabe destacar que las adaptaciones a nivel de red de interacciones no covalentes observadas en proteínas termófilas no parecen adherir a patrones universales, si no que parece ser el producto de un proceso evolutivo particular (Petsko, 2004), por lo que el estudio de proteínas termoestables no caracterizadas puede aportar al descubrimiento y descripción de nuevos mecanismos para incrementar la termoestabilidad proteica.

5. Conclusiones

- La HMPK de *Salmonella typhimurium* y *Thermus thermophilus* son enzimas bifuncionales, capaces de fosforilar HMP y HMP-P.
- Las HMPKs están muy relacionadas con las HMPK/PLK de bacterias Gram positivas, pero sus propiedades catalíticas han divergido significativamente, siendo altamente específicas por HMP y catalizando dos fosforilaciones consecutivas. Estructuralmente, estas diferencias podrían asociarse a la presencia de un residuo de treonina ubicado cerca de la cisteína catalítica y a la presencia de una glutamina en el *hairpin*- β , que forma interacciones polares con el sustrato HMP.
- Ambas enzimas presentan un comportamiento Michaeliano frente a HMP y poseen una K_M similar para este sustrato (en el orden de micromolar), pero difieren en un orden de magnitud en su k_{cat} y en su V_{max} , siendo más rápida la enzima del organismo mesófilo a la temperatura estudiada.

- Cuando se realizan curvas de saturación para ATP, se observa una desviación del modelo Michaelis-Menten. Se requerirán análisis más detallados para estimar los parámetros cinéticos.
- Ambas enzimas son monómeros en las condiciones evaluadas, a diferencia de lo reportado para la HMPK de *E. coli* y a lo predicho en base a las estructuras cristalográficas.
- *Tt*HMPK es significativamente más estable frente a temperatura y agentes químicos que *St*HMPK. Estas diferencias de estabilidad observadas parecen estar asociadas a diferencias en la flexibilidad conformacional de estas enzimas, siendo mucho más rígida la enzima proveniente de *Thermus thermophilus*.
- Las diferencias de flexibilidad pueden explicarse por modificaciones en la red de interacciones no covalentes. *Tt*HMPK tiene una mayor cantidad de interacciones electrostáticas, especialmente de vida corta, que podrían estar restringiendo las fluctuaciones conformacionales. Además, *Tt*HMPK presenta una reducción en el contenido de residuos polares en su núcleo, lo que incrementaría el efecto hidrofóbico que favorece el plegamiento.

6. Referencias

- Abarca-Lagunas, M. J., Rivas-Pardo, J. A., Ramírez-Sarmiento, C. A., & Guixé, V.** (2015). Dissecting the functional roles of the conserved NXXE and HXE motifs of the ADP-dependent glucokinase from *Thermococcus litoralis*. *FEBS letters*, 589(21), 3271-3276.
- Allewell, N. M.** (2012) Thematic minireview series on enzyme evolution in the post-genomic era. *Journal of Biological Chemistry*. 287, 1–2
- Allner, O., Nilsson, L., & Villa, A.** (2012). Magnesium ion–water coordination and exchange in biomolecular simulations. *Journal of Chemical Theory and Computation*, 8(4), 1493-1502.
- Åqvist, J., Kazemi, M., Isaksen, G. V., & Brandsdal, B. O.** (2017). Entropy and enzyme catalysis. *Accounts of chemical research*, 50(2), 199-207.
- Begley, T. P., Downs, D. M., Ealick, S. E., McLafferty, F. W., Van Loon, A. P., Taylor, S., Campobasso, N., Chiu, H., Kinsland, C., Reddick, J., & Xi, J.** (1999). Thiamin biosynthesis in Prokaryotes. *Archives of microbiology*, 171(5), 293-300.
- Carter Jr, C. W., LeFebvre, B. C., Cammer, S. A., Tropsha, A., & Edgell, M. H.** (2001). Four-body potentials reveal protein-specific correlations to stability changes caused by hydrophobic core mutations. *Journal of molecular biology*, 311(4), 625-638.
- Case D.A, T.E. Cheatham, III, T. Darden, H. Gohlke, R. Luo, K.M. Merz, Jr., A. Onufriev, C. Simmerling, B. Wang and R. Woods** (2005). The Amber biomolecular simulation programs. *Journal of Computational Chemistry*. 26, 1668-1688
- Castro-Fernandez, V., Bravo-Moraga, F., Ramirez-Sarmiento, C. A., & Guixé, V.** (2014). Emergence of pyridoxal phosphorylation through a promiscuous ancestor during the evolution of hydroxymethyl pyrimidine kinases. *FEBS Letters*, 588(17), 3068–3073.
- Castro- Fernandez, V., Bravo- Moraga, F., Herrera- Morande, A., & Guixé, V.** (2014). Bifunctional ADP- dependent phosphofructokinase/glucokinase activity in the order Methanococcales–biochemical characterization of the mesophilic enzyme from *Methanococcus maripaludis*. *The FEBS journal*, 281(8), 2017-2029.

Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2011). ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics*, 27(8), 1164-1165.

Dyguda, E., Szefczyk, B., & Sokalski, W. A. (2004). The mechanism of phosphoryl transfer reaction and the role of active site residues on the basis of ribokinase-like kinases. *International Journal of Molecular Sciences*, 5(4), 141-153.

Fiser, A., Do, R. K. G & Sali, A. (2000). Modeling of loops in protein structures. *Protein science*, 9(9), 1753-1773.

Gasteiger, E., Hoogland, C., Gattiker, A., Wilkins, M. R., Appel, R. D., & Bairoch, A. (2005). Protein identification and analysis tools on the ExPASy server. In *The proteomics protocols handbook*. Human press. pp. 571-607

Grant, B. J., Rodrigues, A. P., ElSawy, K. M., McCammon, J. A., & Caves, L. S. (2006). Bio3d: an R package for the comparative analysis of protein structures. *Bioinformatics*, 22(21), 2695-2696.

Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic biology*, 59(3), 307-321.

Haas R. H. (1988) Thiamin and the brain. *Annual Review of Nutrition*. 8: 483–515

Humphrey, W., Dalke, A., & Schulten, K. (1996). VMD: visual molecular dynamics. *Journal of molecular graphics*, 14(1), 33-38.

Hünenberger, P. H., Mark, A. E., & Van Gunsteren, W. F. (1995). Fluctuation and cross-correlation analysis of protein motions observed in nanosecond molecular dynamics simulations. *Journal of molecular biology*, 252(4), 492-503.

Karplus, M., & McCammon, J. A. (2002). Molecular dynamics simulations of biomolecules. *Nature Structural and Molecular Biology*, 9(9), 646.

Krissinel E. & Henrick K. (2007). Inference of macromolecular assemblies from crystalline state. *Journal of Molecular Biology*. 372, 774-797.

Li, H., Robertson, A. D., & Jensen, J. H. (2005). Very fast empirical prediction and rationalization of protein pKa values. *Proteins: Structure, Function, and Bioinformatics*, 61(4), 704-721.

- Lim, S. A., Bolin, E. R., Hart, K. M., Harms, M. J., & Marqusee, S.** (2018). An Evolutionary Trend towards Kinetic Stability in the Folding Trajectory of Ribonucleases H. *Biophysical Journal*, 114(3), 577a.
- Meagher, K. L., Redman, L. T., & Carlson, H. A.** (2003). Development of polyphosphate parameters for use with the AMBER force field. *Journal of computational chemistry*, 24(9), 1016-1025.
- Miranda-Ríos, J., Navarro, M., & Soberón, M.** (2001). A conserved RNA structure (thi box) is involved in regulation of thiamin biosynthetic gene expression in bacteria. *Proceedings of the National Academy of Sciences*, 98(17), 9736-9741.
- Mizote, T., Tsuda, M., Smith, D. D. S., Nakayama, H., & Nakazawa, T.** (1999). Cloning and characterization of the thiD/J gene of *Escherichia coli* encoding a thiamin-synthesizing bifunctional enzyme, hydroxymethylpyrimidine kinase/phosphomethylpyrimidine kinase. *Microbiology*, 145(2), 495-501.
- Newman, J. A., Das, S. K., Sedelnikova, S. E., & Rice, D. W.** (2006). The crystal structure of an ADP complex of *Bacillus subtilis* pyridoxal kinase provides evidence for the parallel emergence of enzyme activity during evolution. *Journal of molecular biology*, 363(2), 520-530.
- Nguyen, V., Wilson, C., Hoemberger, M., Stiller, J. B., Agafonov, R. V., Kutter, S., & Kern, D.** (2017). Evolutionary drivers of thermoadaptation in enzyme catalysis. *Science*, 355(6322), 289-294.
- Nodwell, M. B., Koch, M. F., Alte, F., Schneider, S., & Sieber, S. A.** (2014). A subfamily of bacterial ribokinases utilizes a hemithioacetal for pyridoxal phosphate salvage. *Journal of the American Chemical Society*, 136(13), 4992-4999.
- Park, J. H., Burns, K., Kinsland, C., & Begley, T. P.** (2004). Characterization of Two Kinases Involved in Thiamine Pyrophosphate and Pyridoxal Phosphate Biosynthesis in *Bacillus subtilis*: 4-Amino-5-Hydroxymethyl-2-Methylpyrimidine Kinase and Pyridoxal Kinase. *Journal of Bacteriology*, 186(5), 1571-1573.
- 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.
- Petsko, G. A.** (2001). Structural basis of thermostability in hyperthermophilic proteins, or “There's more than one way to skin a cat”. In *Methods in enzymology*. Academic Press. (Vol. 334, pp. 469-478).

Pucci, F., & Rooman, M. (2017). Physical and molecular bases of protein thermal stability and cold adaptation. *Current opinion in structural biology*, 42, 117-128.

Racine, J. (2006). gnuplot 4.0: a portable interactive plotting utility. *Journal of Applied Econometrics*, 21(1), 133-141.

Rodionov, D. A., Vitreschak, A. G., Mironov, A. A., & Gelfand, M. S. (2002). Comparative genomics of thiamin biosynthesis in procaryotes New genes and regulatory mechanisms. *Journal of Biological chemistry*, 277(50), 48949-48959.

Roe, D. R., & Cheatham III, T. E. (2013). PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. *Journal of chemical theory and computation*, 9(7), 3084-3095.

Roe, D. R., Bergonzo, C., & Cheatham III, T. E. (2014). Evaluation of enhanced sampling provided by accelerated molecular dynamics with Hamiltonian replica exchange methods. *The Journal of Physical Chemistry B*, 118(13), 3543-3552.

Safo, M. K., Musayev, F. N., Hunt, S., di Salvo, M. L., Scarsdale, N., & Schirch, V. (2004). Crystal structure of the PdxY Protein from Escherichia coli. *Journal of bacteriology*, 186(23), 8074-8082.

Schowen R (1998).Thiamin-dependent enzymes.In *Comprehensive Biological Catalysis*, vol 2. Edited by Sinnott M. San DiegoAcademic Press; 1998:217-266.

Schrödinger Release (2018).: Maestro, Schrödinger, LLC, New York, NY, 2018.

Shen, M. Y., & Sali, A. (2006). Statistical potential for assessment and prediction of protein structures. *Protein science*, 15(11), 2507-2524.

Sigrell, J. A., Cameron, A. D., Jones, T. A., & Mowbray, S. L. (1998). Structure of Escherichia coli ribokinase in complex with ribose and dinucleotide determined to 1.8 Å resolution: insights into a new family of kinase structures. *Structure*, 6(2), 183-193.

Starr, T. N., & Thornton, J. W. (2016). Epistasis in protein evolution. *Protein Science*, 25(7), 1204-1218.

Tina, K. G., Bhadra, R., & Srinivasan, N. (2007). PIC: protein interactions calculator. *Nucleic acids research*, 35(suppl_2), W473-W476.

- Toney, M. D.** (2005). Reaction specificity in pyridoxal phosphate enzymes. *Archives of biochemistry and biophysics*, 433(1), 279-287.
- Vogt, G., & Argos, P.** (1997). Protein thermal stability: hydrogen bonds or internal packing?. *Folding and Design*, 2, S40-S46.
- Vriend, G.** (1990). WHAT IF: a molecular modeling and drug design program. *Journal of molecular graphics*, 8(1), 52-56.
- Wang, J., Wang, W., Kollman, P. A., & Case, D. A.** (2001). Antechamber: an accessory software package for molecular mechanical calculations. *Journal of American Chemical Society*, 222, U403.
- Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., & Case, D. A.** (2004). Development and testing of a general amber force field. *Journal of computational chemistry*, 25(9), 1157-1174.
- Wilson, C., Agafonov, R. V., Hoemberger, M., Kutter, S., Zorba, A., Halpin, J., & Kern, D.** (2015). Using ancient protein kinases to unravel a modern cancer drug's mechanism. *Science*, 347(6224), 882-886.
- Yang, Y., Zhao, G., & Winkler, M. E.** (1996). Identification of the pdxK gene that encodes pyridoxine (vitamin B6) kinase in *Escherichia coli* K-12. *FEMS microbiology letters*, 141(1), 89-95.
- Yang, Y., Tsui, H. C. T., Man, T. K., & Winkler, M. E.** (1998). Identification and function of the pdxY gene, which encodes a novel pyridoxal kinase involved in the salvage pathway of pyridoxal 5'-phosphate biosynthesis in *Escherichia coli* K-12. *Journal of bacteriology*, 180(7), 1814-1821.