

UCH-FC  
DOC-BMCN  
R173  
C.1



**UNIVERSIDAD DE CHILE**

**FACULTAD DE CIENCIAS**

**IMPORTANCIA DE LA TOPOLOGIA DE PLEGAMIENTO EN  
LA ESTABILIDAD CONFORMACIONAL Y EN LAS  
PROPIEDADES DE LA INTERFAZ DE ASOCIACION EN  
MIEMBROS DE LA SUPERFAMILIA RIBOQUINASA**

Tesis entregada a la Universidad De Chile en cumplimiento parcial  
de los requisitos para optar al Grado Académico de  
**Doctor en Ciencias con Mención en Biología Molecular, Celular y  
Neurociencias**

Por



**CESAR ANTONIO RAMIREZ SARMIENTO**

Director de Tesis

Dra. Victoria Cristina Guixé Leguía

Enero, 2014

**UNIVERSIDAD DE CHILE**

**FACULTAD DE CIENCIAS**

**INFORME DE APROBACION**

**TESIS DE DOCTORADO**

Se informa a la Escuela de Posgrado de la Facultad de Ciencias que la Tesis de Doctorado presentada por el candidato

CESAR ANTONIO RAMIREZ SARMIENTO

Ha sido aprobada por la Comisión de Evaluación de la Tesis como requisito para optar al grado de Doctor en Ciencias con mención en Biología Molecular, Celular y Neurociencias, en el examen de Defensa Privada de Tesis rendido el día 24 de Enero de 2014.

**Director de Tesis**

Dra. Victoria Guixé



*[Handwritten signature]*

**Comisión Evaluadora**

Dr. Osvaldo Alvarez (Presidente)

Dra. María Cecilia Rojas

Dr. Danilo González

Dr. José Martínez

*[Handwritten signatures of the evaluation committee members]*

*We should enjoy and make the most of life,  
not because we are in constant fear of what might happen to us  
in a mythical afterlife, but because we have only one opportunity to live.*

– Greg Graffin y Steve Olson, *Anarchy Evolution*.

*Dedicado a mi familia.*

## AGRADECIMIENTOS

En primer lugar, quiero agradecer a mi tutora, Dra. Victoria Guixé, por la orientación científica y profesional brindada durante el desarrollo de mi tesis doctoral, y por permitirme plantear mis ideas en torno a los cimientos de su proyecto. Extiendo mis agradecimientos a la Dra. Elizabeth A. Komives, por su confianza al recibirme en su laboratorio, por ayudarme a transformar mi proyecto en resultados y por catalizar mi razonamiento científico crítico. También agradezco a los miembros de la comisión de mi tesis doctoral: Dres. Osvaldo Álvarez, María Cecilia Rojas, Danilo González y José Martínez, por la buena acogida de mi proyecto y el valioso aporte de sus intervenciones.

Agradezco a los coautores de los trabajos publicados en este tiempo, entre ellos: Mauricio Báez, Christian Wilson, Jorge Babul, Ricardo Cabrera, Alejandro Blanco, Pablo Villalobos, Humberto Pereira, Richard Garratt, Jeffrey Noel, Joanna Sułkowska, y José Onuchic. Al Dr. Jorge Babul, por permitir inmiscuirme en su proyecto y apoyar mis decisiones; a Mauricio, por impulsar mis ideas y por contribuir a mi desarrollo como científico y a Jeff, por instruirme en el estudio teórico del plegamiento de proteínas.

Estos años habrían pasado en vano de no ser por Felipe, con quien comí las mejores hamburguesas del lugar más “fome” de Alemania; Brian, que compartió los secretos del *chili sauce* y del estilo OKC; Jeff, por presentarme las mejores películas (Rifftrax!!); Pato, por hacerme sentir como en casa; Diego, por representar lo mejor de Bs. As.; y César, por cambiar mis pipetas por guitarras y el laboratorio por escenarios. También recibí el estímulo de mis amigos y compañeros del Laboratorio de Bioquímica y Biología Molecular, en especial de Rodrigo, Pablo, Diego y Ricardo, del Komives Lab y del NSB en UCSD. A ellos les debo las charlas, risas, comidas, asados y el fútbol.

Le agradezco enormemente a mi familia (que hace ~4 años creció bastante) por apoyarme en todo momento, en especial a mi madre, mi padre y Manuel, por inculcarme que la perseverancia da frutos, y a Nicole, mi esposa, por su apoyo y compañía absolutos, por cocinar increíble (*ñami!*) y por transformar cada día en algo mejor.

## PUBLICACIONES

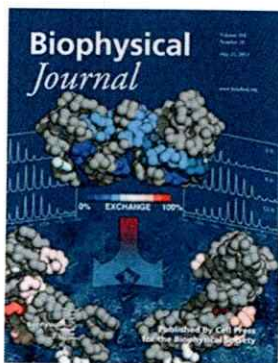
### Artículos publicados durante el período de tesis.

Navarro, F., **Ramírez-Sarmiento, C. A.**, Guixé, V. (2013) Catalytic and regulatory roles of species involved in metal–nucleotide equilibriums in human pyridoxal kinase. *Biometals* **26**, 805–812.

Baez, M., Cabrera, R., Pereira, H. M., Blanco, A., Villalobos, P., **Ramírez-Sarmiento, C. A.**, Caniuguir, A., Guixé, V., Garratt, R. C., Babul, J. (2013) A Ribokinase Family Conserved Monovalent Cation Binding Site Enhances the MgATP-induced Inhibition in *E. coli* Phosphofructokinase-2. *Biophys J* **105**, 185–193.

**Ramírez-Sarmiento, C. A.**, Baez, M., Wilson, C. A. M., Babul, J., Komives, E. A., Guixé, V. (2013) Observation of Solvent Penetration during Cold Denaturation of *E. coli* Phosphofructokinase-2. *Biophys J* **104**, 2254–2263.†

† *Escogido para la portada de la revista Biophysical Journal.*



Sułkowska, J. I., Noel, J. K., **Ramírez-Sarmiento, C. A.**, Rawdon, E. J., Millett, K. C., Onuchic, J. N. (2013) Knotting pathways in proteins. *Biochem. Soc. Trans.* **41**, 523–527.

Baez, M., Wilson, C. A. M., **Ramírez-Sarmiento, C. A.**, Guixé, V., Babul, J. (2012) Expanded Monomeric Intermediate upon Cold and Heat Unfolding of Phosphofructokinase-2 from *Escherichia coli*. *Biophysical Journal* **103**, 2187–2194.

## PRESENTACIONES EN CONGRESOS

### Nacionales

Ramírez C, Cabrera R, Guixé V. Análisis de determinantes estructurales de piridoxal quinasa humana mediante estudios cinéticos y bioinformáticos. *XXXII Reunión Anual de la Sociedad de Bioquímica y Biología Molecular de Chile, Oct 2009.*

Ramírez C, Merino F, Cabrera R, Guixé V. Evolución de la superfamilia riboquinasa. Determinación de residuos implicados en la especificidad de sustratos y caracterización de nuevos motivos asociados a la coordinación de metales. *LII Reunión anual de la Sociedad de Biología de Chile, Nov 2009.*

Werner N, Ramírez C, Guixé V. Role of conserved sequence motifs associated to metal coordination in catalysis and regulation of human pyridoxal kinase. *XXXIII Reunión Anual de la Sociedad de Bioquímica y Biología Molecular de Chile, Sept 2010.*

Ramírez C, Guixé V. Kinetic and evolutive aspects of pyridoxal kinases: Implications for the evolution of the ribokinase superfamily. *XXXIII Reunión Anual de la Sociedad de Bioquímica y Biología Molecular de Chile, Sept 2010.*

Ramírez C, Baez M, Wilson CAM, Babul J, Komives E, Guixé V. Following cold denaturation of phosphofructokinase-2 by amide hydrogen/deuterium (H/D) exchange mass spectrometry. *XXXIV Reunión Anual de la Sociedad de Bioquímica y Biología Molecular de Chile, Sept 2011.*

Ramírez C, Noel JK, Baez M. Kinetic folding simulations of the knotted ribbon-helix-helix (RHH) protein VirC2 using a simplified structured-based model. *XXXIV Reunión Anual de la Sociedad de Bioquímica y Biología Molecular de Chile, Sept 2011.*

Ramírez-Sarmiento CA, Villalobos P, Baez M, Guixé V, Babul J. Understanding non-trivial protein topologies using simplified molecular dynamics. *XXXV Reunión Anual de la Sociedad de Bioquímica y Biología Molecular de Chile, Oct 2012.*

## Internacionales

**Ramírez C, Guixé V.** Bioinformatic and kinetic studies in human pyridoxal kinase reveal a new motif associated to zinc coordination and enzyme activity regulation. *24th Annual Symposium of The Protein Society, San Diego, California, EEUU, Ago 2010.*

**Ramírez C, Guixé V.** Bioinformatic and kinetic studies in human pyridoxal kinase reveal new structural determinants associated to zinc coordination and activity regulation. *3rd Latin American Protein Society Meeting, Salta, Argentina, Oct 2010.*

**Ramírez C, Rivas-Pardo JA, Guixé V.** Similarities between the mechanical unfolding pathways of ribokinase superfamily enzymes resolved by molecular dynamics simulations: The case of a hyperthermophile glucokinase and a mesophile adenosine kinase. *55th Annual Biophysical Society Meeting, Baltimore, Maryland, EEUU, Mar 2011.*

**Ramírez C, Noel JK, Baez M.** Topological constraints and challenges imposed by knots in proteins. *2nd International Society for Computational Biology (ISCB) Latin American Meeting, Santiago, Chile, Mar 2012. Premio a la Mejor Presentación Oral.*

**Ramírez-Sarmiento CA, Baez M, Babul J, Komives EA, Guixé V.** Direct Observation Of Solvent Penetration During Cold Denaturation Of Pfk-2 By Amide Hydrogen Deuterium ( $H^2H$ ) Exchange Mass Spectrometry. *26th Annual Symposium of The Protein Society, San Diego, California, EEUU, Ago 2012.*

**Ramírez-Sarmiento CA, Baez M, Villalobos P, Komives EA, Guixé V, Babul J.** Structural changes upon dissociation of a highly cooperative dimeric protein. *27th Annual Symposium of The Protein Society, Boston, Massachusetts, EEUU, Jul 2013.*

**Ramírez-Sarmiento CA.** Observation of solvent penetration during solvent, temperature and mutational perturbations of *E. coli* Pfk-2 by HXMS. *International seminar Germany-Chile: from Plant Biology to Computational Chemistry and Molecular Bioinformatics, Talca, Chile, Sept 2013. Conferencista invitado.*

Abarca M, **Ramírez-Sarmiento CA**, Castro-Fernández V, Merino F, Guixé V. Biochemical characterization of metal-nucleotide binding in an ADP-dependent phosphofructokinase from *Pyrococcus horikoshii*. *XII Panamerican Association for Biochemistry and Molecular Biology Meeting, Puerto Varas, Chile, Nov 2013.*

**Ramírez-Sarmiento CA**, Baez M, Villalobos P, Babul J, Komives EA, Guixé V. Structural changes in *E. coli* phosphofructokinase-2 upon subunit dissociation induced by chemical, temperature and mutational perturbations. *XII Panamerican Association for Biochemistry and Molecular Biology Meeting, Puerto Varas, Chile, Nov 2013.*

**Ramírez-Sarmiento CA**, Merino F, Babul J. Addition of non-native electrostatic interactions in coarse-grained and all-atom structure-based models for protein folding and binding. *XII Panamerican Association for Biochemistry and Molecular Biology Meeting, Puerto Varas, Chile, Nov 2013.*

Villalobos P, **Ramírez-Sarmiento CA**, Babul J. Increased local flexibility of the forkhead domain of FOXP2 transcription factor is required for domain swapping. *XII Panamerican Association for Biochemistry and Molecular Biology Meeting, Puerto Varas, Chile, Nov 2013.*

**Ramírez-Sarmiento CA**. Expanded monomeric intermediate upon thermal and chemical denaturation of phosphofructokinase-2 from *Escherichia coli*. *2nd Zing Enzymes, Coenzymes and Metabolic Pathways Conference, Cancún, México, Nov 2013. Conferencista invitado.*



## FINANCIAMIENTO

- Proyecto FONDECYT 1070111, 1090336 y 1110137.
- Beca para estudios de Doctorado en Chile, CONICYT 2009-2013.
- Beca de congresos internacionales, CONICYT, 2010.
- Beca de pasantía en la Universidad de California, San Diego, Estados Unidos, MECESUP/STIPAs UCH 0717, 2011.
- Beca de pasantía en el extranjero, MECESUP UCH 0713, 2012.
- Beca de invitación de especialistas extranjeros, MECESUP UCH 0713, 2012.

## INDICE

<b>BIOGRAFIA</b>	<b>iii</b>
<b>AGRADECIMIENTOS</b>	<b>iv</b>
<b>PUBLICACIONES</b>	<b>v</b>
Artículos publicados durante el período de tesis.	v
<b>PRESENTACIONES EN CONGRESOS</b>	<b>vi</b>
Nacionales	vi
Internacionales	vii
<b>FINANCIAMIENTO</b>	<b>ix</b>
<b>INDICE</b>	<b>x</b>
<b>INDICE DE FIGURAS</b>	<b>xiv</b>
<b>INDICE DE TABLAS</b>	<b>xvi</b>
<b>LISTA DE ABREVIATURAS</b>	<b>xvii</b>
<b>RESUMEN</b>	<b>xviii</b>
<b>ABSTRACT</b>	<b>xx</b>
<b>INTRODUCCION</b>	<b>1</b>
Aspectos teóricos del plegamiento de proteínas.	1
Teoría del paisaje energético de proteínas y el principio de la mínima frustración	2
Rol de la topología nativa en la definición de la velocidad y mecanismo de plegamiento.	4
Rol de la topología nativa en el mecanismo de asociación de proteínas diméricas.	6
Importancia del establecimiento del rol de la topología del estado nativo en la función y (des)plegamiento de enzimas.	8
La familia riboquinasa como modelo de estudio del rol de la topología del estado nativo en el plegamiento y asociación homodimérica.	11
<b>HIPOTESIS DEL TRABAJO</b>	<b>19</b>

<b>OBJETIVO GENERAL</b>	<b>20</b>
<b>OBJETIVOS ESPECIFICOS</b>	<b>20</b>
<b>SINOPSIS</b>	<b>21</b>
<b><u>CAPITULO UNO: Desnaturación en frío de Pfk-2 de <i>Escherichia coli</i></u></b>	<b><u>22</u></b>
<b>1.1. INTRODUCCION</b>	<b>23</b>
<b>1.2. PRINCIPALES RESULTADOS Y CONCLUSIONES</b>	<b>29</b>
<b>1.3. AGRADECIMIENTOS</b>	<b>31</b>
<b><u>CAPITULO DOS: Caracterización biofísica del monómero aislado inducido por disociación del homodímero de Pfk-2 y Kdgk</u></b>	<b><u>51</u></b>
<b>2.1. INTRODUCCION</b>	<b>52</b>
<b>2.2. MATERIALES Y METODOS</b>	<b>59</b>
2.2.1. Purificación y almacenamiento de Pfk-2 de <i>Escherichia coli</i> .	59
2.2.2. Clonamiento, purificación y almacenamiento de Kdgk de <i>Escherichia coli</i> .	59
2.2.3. Cristalización de la enzima Kdgk de <i>Escherichia coli</i> en presencia de MgADP	61
2.2.4. Escaneo in silico de mutaciones de residuos de la interfaz por alanina en Pfk-2 y Kdgk de <i>Escherichia coli</i> .	62
2.2.5. Mutagénesis sitio-dirigida de Pfk-2 y Kdgk de <i>Escherichia coli</i> .	62
2.2.6. Sedimentación por velocidad de las enzimas Pfk-2 y Kdgk.	62
2.2.7. Intercambio isotópico para amidas de la cadena principal del intermediario monomérico inducido por GndHCl de Pfk-2 y de la mutante L93A.	63
2.2.8. Digestión proteolítica de Pfk-2 y secuenciación de los fragmentos peptídicos.	64
2.2.9. Espectrometría de masa.	66
2.2.10. Seguimiento del desplegamiento inducido por GndHCl de la mutante L93A de Pfk-2 y de Kdgk silvestre y las mutantes Y101A y W103A mediante difracción circular.	67

2.2.11. Difracción de rayos X de ángulo pequeño de la enzima Kdgk de <i>Escherichia coli</i> y de la mutante W103A.	67
<b>2.3. RESULTADOS</b>	<b>68</b>
2.3.1. Purificación y remoción del epítipo de histidina de la enzima Kdgk mediante proteólisis.	68
2.3.2. Resultados preliminares de la estructura resuelta por remplazo molecular de la enzima Kdgk de <i>Escherichia coli</i> .	70
2.3.3. Determinación de los residuos claves para la estabilidad del homodímero de Pfk-2 y Kdgk.	74
2.3.4. La mutante L93A de Pfk-2 y la mutante W103A de Kdgk son monoméricas.	80
2.3.5. La mutante L93A de Pfk-2 representa los primeros rearrreglos estructurales producto de la disociación.	85
2.3.6. La mutante L93A de Pfk-2 exhibe un desplegamiento no cooperativo en presencia de GndHCl.	90
2.3.7. Observación de un intermediario monomérico en el desplegamiento químico de la enzima Kdgk.	93
2.3.8. Confirmación del estado parcialmente desplegado de la mutante W103A de Kdgk mediante difracción de rayos X de ángulo pequeño.	95
<b>2.4. DISCUSION</b>	<b>101</b>
<b>2.5. AGRADECIMIENTOS</b>	<b>106</b>
<b><u>CAPITULO TRES: Estudios <i>in silico</i> del plegamiento de topologías complejas y efecto de la penetración del solvente.</u></b>	<b><u>108</u></b>
<b>3.1. INTRODUCCION</b>	<b>109</b>
<b>3.2. MATERIALES Y METODOS.</b>	<b>111</b>
3.2.1. Modelos de estudio	111
3.2.2. Modelo simplificado basado en estructura nativa.	112
3.2.2. Simulaciones moleculares utilizando modelos simplificados.	112

3.2.3. Potencial polarizante externo para muestreo acelerado de transiciones de plegamiento en modelos simplificados.	114
3.2.4. Potencial de solvatación de contactos nativos.	115
3.2.5. Análisis de las simulaciones.	115
<b>3.3. RESULTADOS.</b>	<b>116</b>
3.3.1. Aceleración del muestreo de configuraciones mediante la aplicación del potencial armónico sobre los contactos nativos.	116
3.3.2. El método de muestreo acelerado permite describir el paisaje energético de plegamiento de la quinasa humana de piridoxal.	118
3.3.3. El potencial de solvatación permite describir la hidratación del dominio B1 de la proteína G de <i>Streptomyces</i> .	123
<b>3.4. DISCUSION.</b>	<b>125</b>
<b>3.5. AGRADECIMIENTOS</b>	<b>126</b>
<b>DISCUSION GENERAL</b>	<b>128</b>
<b>CONCLUSIONES</b>	<b>135</b>
<b>REFERENCIAS</b>	<b>138</b>

## INDICE DE FIGURAS

Figura 1. Heterogeneidad estructural y funcional de la superfamilia riboquinasa.	13
Figura 2. Modificaciones topológicas a nivel del dominio menor y efectos en la asociación proteína-proteína de las enzimas de la familia riboquinasa.	16
Figura 3. Visualización de la estructura del dominio bimolecular de Pfk-2.	27
Figura 4. Esquema del experimento de intercambio isotópico hidrógeno/deuterio (HXMS) de las amidas de la cadena principal de la enzima Pfk-2 medido por espectrometría de masas.	65
Figura 5. Purificación y remoción del epítopo de histidinas en el extremo amino-terminal de la enzima recombinante Kdgk de <i>Escherichia coli</i> .	69
Figura 6. Cristales obtenidos para la enzima Kdgk de <i>Escherichia coli</i> en presencia del ligando MgADP.	71
Figura 7. Estructura resuelta de la enzima Kdgk de <i>Escherichia coli</i> en presencia del ligando MgADP.	73
Figura 8. Cambios de estabilidad de la asociación de subunidades estimados mediante mutaciones sitio-específicas <i>in silico</i> .	76
Figura 9. Conservación de la posición de los residuos claves para la estabilidad del homodímero en la estructura del dominio menor	77
Figura 10. Cambio en el grado de frustración local como consecuencia de la oligomerización de Pfk-2 y Kdgk.	79
Figura 11. Disociación de las enzimas Pfk-2 y Kdgk de <i>Escherichia coli</i> mediante mutagénesis sitio-dirigida de la interfaz de asociación proteína-proteína.	81
Figura 12. Efectos de la adición de agentes caotrópicos sobre la distribución de especies oligoméricas en la enzima Kdgk silvestre y la mutante Y101A.	83
Figura 13. Caracterización estructural del intermediario monomérico de Pfk-2 inducido por GndHCl mediante HXMS.	88
Figura 14. Caracterización estructural de la mutante L93A de Pfk-2 mediante HXMS.	89
Figura 15. Mapeo de las regiones afectadas por la disociación inducida por mutagénesis sitio-dirigida del residuo L93 en la estructura de la enzima Pfk-2.	91

Figura 16. Pérdida no cooperativa de estructura secundaria de la mutante monomérica de Pfk-2 en presencia de GndHCl.	92
Figura 17. Presencia de un intermediario monomérico en el desplegamiento de la enzima Kdgk.	94
Figura 18. Difracción de rayos X de ángulo pequeño para las especies dimericas y monoméricas de Kdgk.	97
Figura 19. Propiedades hidrodinámicas de la enzima Kdgk y la mutante W103A en solución.	99
Figura 20. Determinación <i>ab initio</i> de la forma del homodímero y monómero aislado de la enzima Kdgk de <i>Escherichia coli</i> .	100
Figura 21. Diferencias en la accesibilidad al solvente del dímero nativo de Pfk-2 y del monómero aislado obtenido mediante diversas perturbaciones globales y locales.	104
Figura 22. Modelos de estudio para la inclusión de potenciales adicionales en modelos nativo-céntricos.	113
Figura 23. Transiciones entre el estado plegado y desplegado de distintas proteínas modelo con mecanismo de plegamiento de dos estados mediante potenciales harmónicos sobre Q.	117
Figura 24. Propiedades termodinámicas de las proteínas CI2, SH3 y ubiquitina para diferentes fuerzas del potencial harmónico sobre Q.	119
Figura 25. Probabilidad de formación de los contactos nativos de las proteínas CI2, SH3 y ubiquitina en el estado de transición de plegamiento.	120
Figura 26. Análisis termodinámico del proceso de plegamiento de la quinasa humana de piridoxal mediante el uso de modelos nativo-céntricos y potenciales harmónicos sobre Q.	122
Figura 27. Intermediarios penetrados por solvente en el desplegamiento termodinámico del dominio B1 de la proteína G de <i>Streptomyces</i> .	124

## INDICE DE TABLAS

Tabla 1. Parámetros obtenidos del procesamiento de los patrones de difracción colectados y posterior refinamiento de la estructura de Kdgk de <i>Escherichia coli</i> en presencia de MgADP.	72
Tabla 2. Propiedades hidrodinámicas de las especies observadas mediante ultracentrifugación analítica para las enzimas silvestres y mutantes de Kdgk y Pfk-2 *	84
Tabla 3. Fragmentos peptídicos de la enzima Pfk-2 y la mutante L93A analizados mediante espectrometría de masas.	87



## LISTA DE ABREVIATURAS

**ATP**, adenosina trifosfato

**ADP**, adenosina difosfato

**BME**, 2-mercaptoetanol

**D<sub>2</sub>O**, agua deuterada

**DTT**, ditioneitol

**GndHCl**, hidrocloreuro de guanidinio.

**H<sub>2</sub>O**, agua

**HXMS**, intercambio isotópico de hidrógeno/deuterio de las amidas de la cadena principal, seguido por espectrometría de masas.

**Mg**, magnesio

**Pfk-2**, fosfofructoquinasa-2

**Kdgk**, 2-ceto-3-desoxigluconato quinasa

**Q**, fracción de contactos nativos

**TCEP**, tris(2-carboxietil)fosfina

**TFA**, ácido trifluoroacético

## RESUMEN

El proceso de plegamiento y asociación de proteínas monoméricas y homodiméricas pequeñas está dominado por la topología del estado nativo y las interacciones presentes en este ensamble conformacional. En proteínas compuestas por más de 100 residuos, la frustración energética que se origina por la formación de contactos no-nativos se ha evadido mediante la presencia de múltiples dominios que se pliegan de manera paralela e independiente. Sin embargo, en varios casos estos dominios están conectados por medio de la topología reentrante de la cadena polipeptídica, lo cual ha sido descrito como un rasgo determinante del acoplamiento y de la cooperatividad del proceso de desplegamiento entre los dominios involucrados. Los miembros homodiméricos de la familia riboquinasa presentan esta discontinuidad topológica entre un dominio de menor tamaño que permite la asociación entre subunidades y el dominio tipo Rossmann que es responsable de la unión de los sustratos.

La primera parte de este trabajo describe el acoplamiento entre los procesos de disociación y desplegamiento parcial de las enzimas homodiméricas Pfk-2 y Kdgk de *Escherichia coli*, pertenecientes a la familia riboquinasa, producto de la aplicación de distintas perturbaciones globales (cambios de temperatura, adición de agentes caotrópicos) y locales (mutaciones sitio-dirigidas de la interfaz proteína-proteína). Todas estas perturbaciones causan la disociación de las subunidades y el desplegamiento parcial hacia una especie monomérica altamente hidratada, pero las propiedades hidrodinámicas y estructurales precisas de esta especie dependen de la perturbación aplicada. Nuestros resultados sugieren que los primeros rearrreglos estructurales ocurren en la región amino-terminal del dominio tipo Rossmann, el cual está directamente conectado al dominio menor a través de la topología reentrante de la cadena polipeptídica, lo cual sugiere que, como una consecuencia de la topología del estado nativo, la estabilidad del monómero depende de la formación del homodímero. Además,

la posición e identidad de los residuos que contribuyen mayormente a la estabilidad del homodímero son concebidas en función de la topología de la interfaz proteína-proteína.

La segunda parte de este trabajo describe la implementación de estrategias de muestreo acelerado para el estudio de plegamiento termodinámico *in silico* utilizando modelos simplificados basados en la estructura nativa de proteínas, y la adición de potenciales de solvatación implícita que permiten la formación de interacciones directas y mediadas por solvente entre residuos en contacto en el estado nativo. Determinamos que las estrategias de muestreo acelerado facilitan el estudio de modelos proteicos cuyas topologías complejas constituyen un desafío computacional, mientras que la incorporación de configuraciones separadas por solvente permite capturar la estructura de estados intermedios en los cuales el centro hidrofóbico es penetrado por el agua circundante.

## ABSTRACT

The folding and binding process for small monomeric and homodimeric proteins is dominated by the native state topology and the observed interactions in this conformational ensemble. In proteins composed by more than 100 residues, energetic frustration that originates from the formation of non-native contacts is avoided by the presence of multiple domains that fold both independently and in parallel. However, in several cases these domains are connected through the reentrant topology of the polypeptide chain, which has been described as a key feature to determine coupling and cooperativity of the unfolding process between the domains involved. The homodimeric members from the ribokinase family possess this topological discontinuity between a small domain responsible of subunit association and the Rossmann-like domain responsible for substrate binding.

The first part of this work describes the coupling between dissociation and partial unfolding of the homodimeric enzymes Pfk-2 and Kdgk of *Escherichia coli* that belong to the ribokinase family, due to the application of several global (temperature shifts, addition of chaotropic agents) and local perturbations (site-directed mutagenesis of the protein-protein interface). All these perturbations cause subunit dissociation and partial unfolding leading to a highly hydrated monomeric species but the hydrodynamic and structural properties of this species are defined by the applied perturbation. Our results suggest the first structural rearrangements occur in the amino-terminal region of the Rossmann-like domain, which is directly connected to the small domain through the reentrant topology of the polypeptide chain, suggesting that monomer stability depends on dimer formation as a consequence of the native state topology. Moreover, the

position and identity of the residues contributing most of the stability of the homodimer are envisioned as a function of the topology of the protein-protein interface.

The second part of this work describes the implementation of enhanced sampling strategies for *in silico* thermodynamic folding studies using simplified models based on the protein native state, and the addition of implicit solvation potentials that allow the formation of direct and solvent-mediated interactions between residues in contact in the native state. We determined that the enhanced sampling strategy eases the study of protein models whose complex topology are computationally challenging, while the incorporation of solvent-separated configurations allows capturing the structure of intermediate states where the hydrophobic core is penetrated by the surrounding water.

## INTRODUCCION

### Aspectos teóricos del plegamiento de proteínas.

Junto con la elucidación de la primera estructura a alta resolución de la proteína mioglobina mediante cristalografía de rayos X en el año 1959 [1], surgió una pregunta fundamental en el área de la biofísica y la bioquímica: ¿cuáles son los mecanismos que participan de la obtención o pérdida de esta estructura compacta, única y coherente con la función que desempeña? Este proceso, conocido como (des)plegamiento de proteínas, es un problema teórico y experimental altamente desafiante, dado que busca entender la “especificidad” de la organización espacial tridimensional alcanzada por una proteína, a pesar del gran espacio conformacional inicial al cual puede acceder la cadena polipeptídica en el estado desplegado. Adicionalmente, esta configuración única debe alcanzarse en una escala de tiempo biológicamente accesible [2].

El trabajo pionero acerca del desplegamiento reversible de la ribonucleasa pancreática bovina llevado a cabo por Anfinsen [3], sugirió que la información requerida para alcanzar el estado nativo se encuentra codificada en la secuencia aminoacídica de las proteínas y en su interacción con el ambiente. Dada la reversibilidad del proceso y la singularidad de la estructura nativa, la cual es termodinámicamente estable y altamente ordenada, Anfinsen sugirió que esta conformación corresponde al mínimo energético global [3]. Adicionalmente, Levinthal propuso que las proteínas alcanzan el estado nativo a partir del estado desplegado mediante rutas de plegamiento específicas, siguiendo una secuencia de estados intermediarios consecutivos [4]. Al no tener que acceder exhaustivamente a todas las configuraciones admitidas por la cadena polipeptídica, el tiempo que tarda una proteína en plegarse se reduce a la escala de segundos o menos [2, 4].

Sin embargo, a lo largo del tiempo se ha demostrado experimentalmente que, para gran parte de las proteínas pequeñas de un solo dominio, el proceso de plegamiento

*in vitro* ocurre sin la acumulación de intermediarios e incluso en escalas de tiempo de milisegundos [5]. De esta afirmación se desprenden dos aspectos fundamentales: i) el plegamiento de proteínas consiste en un mecanismo tan simple como una reacción “todo o nada” entre los estados plegado y desplegado [6]; ii) este proceso ocurre de manera rápida [5]. En consecuencia, se requirió de una nueva teoría acerca del (des)plegamiento de proteínas para dar cabida a estas consideraciones.

#### Teoría del paisaje energético de proteínas y el principio de la mínima frustración

El marco teórico más atractivo para satisfacer estas consideraciones corresponde al del paisaje energético del plegamiento proteico [7]. Este modelo sugiere que, en condiciones que favorecen el plegamiento, las proteínas siguen un gradiente desde cualquiera de los infinitos estados altamente entrópicos y carentes de interacciones intramoleculares, hacia un estado de baja entropía, altamente organizado y que presenta todas las interacciones intramoleculares nativas (observadas en la estructura resuelta mediante cristalografía de rayos X o resonancia magnética nuclear) [8]. Durante este proceso, la proteína va acumulando interacciones favorables que disminuyen la energía, promoviendo la compactación y reduciendo el espacio de fases de conformaciones posibles [9]. Por lo tanto, pueden existir muchas rutas de plegamiento para una proteína particular (es decir, no se requiere de una vía plegamiento determinada ni de la acumulación de intermediarios estables a través de esta vía), pero aquellas que permiten adquirir una energía de estabilización más rápidamente y que posean un bajo costo entrópico serán las predominantes [10].

Dentro de este marco, una propiedad obligatoria del mecanismo de plegamiento de las proteínas actuales es el de evitar las conformaciones no-nativas durante el proceso de plegamiento. Por consiguiente, se ha propuesto que las secuencias proteicas actuales disminuyen la exhaustividad de la búsqueda conformacional mediante el favorecimiento de la formación de contactos presentes en el estado nativo, minimizando la formación de

contactos no-nativos que conducen a trampas cinéticas, estados intermediarios o conformaciones no-nativas [11].

Este concepto, conocido también como el principio de mínima frustración [12], permite simplificar la búsqueda en el espacio de configuraciones posibles durante el proceso de plegamiento y alcanzar el estado nativo de manera eficiente y rápida [5]. De acuerdo a este principio, las secuencias han evolucionado de manera de acumular interacciones estabilizantes, las cuales son mayoritariamente interacciones entre residuos que establecen contactos nativos, que permiten que las proteínas alcancen su estructura nativa independientemente de la conformación inicial de la cadena polipeptídica y utilizando numerosas rutas paralelas con conformaciones isoenergéticas [13]. Como consecuencia, la acumulación de estos contactos que favorecen el plegamiento ocurre al mismo tiempo que se evita que las interacciones accesibles sean potencialmente “frustrantes”. Dado que las interacciones no-nativas (“frustrantes”) son minimizadas durante este proceso, la evolución ha permitido esculpir el paisaje energético en la forma de un embudo (“funnel-like shaped landscapes”), cuya “rugosidad” (presencia de intermediarios transitorios y trampas conformacionales no-nativas) depende principalmente de los conflictos remanentes que surgen entre las interacciones estabilizantes durante el proceso de plegamiento [14].

En este escenario, para una proteína cuya probabilidad de formar interacciones conflictivas es mínima (“frustración energética” mínima), la topología del estado nativo es el factor determinante del mecanismo de plegamiento de la cadena polipeptídica hacia una conformación estable [15]. Esto se debe a que la topología nativa, en ese caso definida como el empacamiento y conectividad de los elementos de estructura secundaria en un dominio proteico [16], impone restricciones geométricas al espacio de movimientos que una cadena polipeptídica puede ejecutar de manera de establecer los contactos nativos de manera correcta.



### Rol de la topología nativa en la definición de la velocidad y mecanismo de plegamiento.

La evidencia directa del rol de la topología en el plegamiento proteico viene de la mano del establecimiento de una métrica adecuada para su descripción numérica. Esta métrica, conocida como orden de contacto, consiste en la separación de los contactos nativos en la estructura primaria de la proteína y representa un índice cuantitativo de la complejidad topológica de una proteína [17], sin consideración de otros detalles estructurales como el tamaño y el contenido de estructura secundaria [18, 19], y aun así se correlaciona de manera significativa con las velocidades de plegamiento de proteínas de un solo dominio y mecanismo de dos estados [20]. La inclusión del largo de la cadena polipeptídica y de la forma de la proteína en el estado nativo, permite ampliar esta correlación hacia la predicción de la velocidad de plegamiento de proteínas con múltiples dominios y/o que presentan un mecanismo de plegamiento de tres estados [18, 21, 22]. En consecuencia, las proteínas de topologías simples, estabilizadas principalmente por contactos nativos locales (cuya separación en la secuencia primaria es pequeña), tienden a alcanzar el estado nativo más rápido que proteínas estabilizadas por contactos no locales (más complejas), las cuales poseen mayores posibilidades de alcanzar trampas geométricas o conformaciones que deben desplegarse parcial o totalmente para permitir el alcance del estado nativo.

Esto sugiere que las características principales del plegamiento de proteínas pueden predecirse con solo conocer la energía de estabilización de los elementos de la estructura nativa y el costo entrópico del empaque de estos elementos en el espacio tridimensional, propiedades que son extraíbles a partir del mapa de contactos nativos de la proteína [9, 10, 23, 24]

El rol de la topología ha sido demostrado experimentalmente mediante el estudio de las consecuencias estructurales (tanto en el estado nativo como en el estado de transición de plegamiento) de la realización de mutaciones puntuales en distintos lugares de una proteína dada a través de la medición y el análisis comparativo del

plegamiento cinético y al equilibrio entre la proteína silvestre y las distintas variantes mutagénicas [25]. Este mapeo estructural del estado de transición del plegamiento de proteínas, conocido como análisis de valores  $\Phi$  [26], ha sido aplicado exitosamente a proteínas como barnasa [27], CI2 [28], src SH3 [29, 30], GCN4 [31], CheY [32], el dominio de activación de procarboxipeptidasa 2 (ADA2h) [33], acil fosfatasa [34], apoflavodoxina [35], entre otras. Dado que muchas de las proteínas estudiadas son estructuralmente similares entre sí [32, 34, 35], pese al bajo porcentaje de identidad de secuencia, estos antecedentes han permitido demostrar que proteínas con topologías nativas similares poseen velocidades de plegamiento y estados de transición de plegamiento similares [36]. Por otro lado, permutaciones circulares en proteínas como lisozima T4 [37], CI2 [38], SH3 de espectrina [39] y S6 [40] han demostrado que proteínas con arquitecturas y secuencias equivalentes, pero con cambios en la organización topológica de la cadena polipeptídica conducen a modificaciones del estado de transición de plegamiento, de la cooperatividad del mecanismo de plegamiento y de las cinéticas de (des)plegamiento, dependiendo de los cambios inducidos por los rearrreglos de la cadena polipeptídica en el orden de contacto [41].

La teoría del paisaje energético canalizado (“funnel-like”) y el principio de mínima frustración son los cimientos fundamentales para entender el rol de la topología en el plegamiento de proteínas. En base a éstos, se han desarrollado modelos computacionales simplificados (modelos tipo Gō, [42]) en donde solo se considera la información pertinente a la estructura nativa (mediante la inclusión implícita del estado nativo de la proteína por medio del mapa de contactos nativos obtenido a partir de la estructura cristalográfica) en la función de energía potencial, mientras que la frustración energética es dramáticamente reducida; con ello se prescinde de interacciones conflictivas para alcanzar el estado nativo [43]. Como consecuencia, estos modelos biomoleculares permiten atenuar la dificultad impuesta por la escala de tiempo en la cual ocurre el proceso de plegamiento [44] y determinar el grado de control topológico del proceso de plegamiento en ausencia de interacciones que desestabilicen el estado

nativo [45]; en otras palabras, estos modelos exhiben un paisaje energético perfectamente canalizado. Esta estrategia ha permitido corroborar que la topología es el factor determinante de proteínas como CI2, SH3, Rnasa H, barnasa, CheY, dihidrofolato reductasa e interleukina-1 $\beta$  [43, 45], reproduciendo las características del mecanismo de plegamiento observadas experimentalmente.

#### Rol de la topología nativa en el mecanismo de asociación de proteínas diméricas.

La topología también es un factor determinante del plegamiento de proteínas diméricas, como ha sido comprobado mediante estudios de plegamiento cinético y al equilibrio y/o análisis de valores  $\Phi$  para diversas proteínas, como el represor Arc [46], represor cro del fago lambda [47] y represor de triptófano [48]. Los efectos de la asociación de subunidades se manifiestan en la determinación del mecanismo de (des)plegamiento, de manera que éste puede ser tan simple como la transición de dos estados observada para proteínas monoméricas de un solo dominio, pasando de un dímero nativo ( $N_2$ ) a un monómero desplegado ( $2U$ ), a ser un mecanismo de tres estados con la presencia de un intermediario monomérico ( $2I$ ) o dimérico ( $I_2$ ) [49]. Además, la cantidad de contactos nativos formados en el estado de transición, así como también la importancia de las cadenas laterales de los residuos participantes de los contactos inter o intramoleculares, en los pasos entre el(los) estado(s) de transición y los estados basales observados se manifiesta en los cambios inducidos por mutaciones de estos residuos en las constantes cinéticas de desplegamiento y/o de replegamiento [46].

Una aproximación simple para determinar la importancia de la topología en el plegamiento y asociación de proteínas diméricas se obtiene al correlacionar el mecanismo de asociación determinado experimentalmente en función de la hidrofobicidad de su interfaz de asociación y de la razón entre los contactos nativos establecidos intermolecularmente y aquellos pertinentes al monómero aislado [50]. De este diagrama de fase se desprende que, en general, un homodímero que presenta un mecanismo de asociación y plegamiento de dos estados se caracteriza por una

proporción elevada de contactos intermoleculares por sobre contactos intramoleculares y una interfaz más hidrofóbica en comparación a homodímeros con mecanismos de tres estados [50].

Aplicando los modelos computacionales simplificados basados en potenciales energéticos centrados en el estado nativo [43] para el estudio del mecanismo de asociación y plegamiento de homodímeros, se logró confirmar que la topología del estado nativo juega un rol fundamental en la determinación del proceso de reconocimiento y asociación de subunidades para once homodímeros diferentes, siendo capaz de reproducir las observaciones experimentales macroscópicas en cuanto a si los procesos de asociación y plegamiento se encuentran acoplados en un estado o si existen intermediarios monoméricos o diméricos involucrados [50]. Este análisis se ha extendido a la determinación del acoplamiento entre plegamiento y asociación de heterodímeros, homotrímeros y homotetrámeros [51], así como también para la determinación de un mecanismo especial de homodimerización generado mediante intercambio de segmentos (“domain swapping”, [52]). En este último caso, el mecanismo de asociación consiste en el intercambio de segmentos de estructura o dominios completos con la parte simétricamente idéntica de la subunidad adyacente, de manera que los contactos nativos intramoleculares son ahora restituidos de manera intermolecular [53], por lo que las proteínas que presentan este mecanismo de asociación poseen al menos dos estados nativos definidos correspondientes al mínimo energético: monomérico y dimérico [54]. Mediante el uso de estos modelos simplificados con interacciones simétricas intermoleculares, se demostró que es posible reproducir el estado homodimérico con intercambio de segmentos resuelto experimentalmente con solo conocer el mapa de contactos representativo del estado nativo del monómero aislado [55], así como también dilucidar el rol de la concentración de proteína en la superación de la gran barrera energética que separa ambos estados nativos [56].

Estos antecedentes reafirman que el proceso de asociación de proteínas (y el acoplamiento con el proceso de plegamiento) son guiados por un embudo de energía libre, determinado por interacciones favorables que se establecen en alguna vecindad del paisaje energético del estado nativo [57, 58], y que el grado de rugosidad del piso del paisaje energético (donde se define el mínimo energético) sea un indicador tanto de los conformeros disponibles para desplazar el equilibrio hacia el establecimiento de la asociación homo(hetero)-molecular como de la especificidad de la interacción [59]. En particular, estos conformeros se caracterizan por la presencia de múltiples regiones con eventos de desplegamiento local independientes, de manera que las superficies de interacción homo(hetero)-molecular se componen de áreas heterogéneas con baja y alta estabilidad [60]. La presencia de una desestructuración dinámica parcial y local permite que el proceso de asociación sea más efectivo y rápido, dado que facilita la interacción inespecífica de largo alcance previo al fortalecimiento de las interacciones intermoleculares y la definición de la especificidad de la unión [61, 62]. En este proceso de interacción inicial, el acomodo correcto de la cadena principal de las proteínas es crucial para la complementariedad entre las interfaces de asociación [63], y esta información es suficiente para predecir gran parte de las estructuras de los complejos homo(hetero)-moleculares resueltos hasta ahora [64]. De esta manera, los cambios en la base del paisaje energético para un evento de unión dado, retrata i) los cambios en las poblaciones de cada conformero [65] y ii) el paisaje energético es dinámico [66].

#### Importancia del establecimiento del rol de la topología del estado nativo en la función y (des)plegamiento de enzimas.

La mayoría de las determinaciones acerca de la importancia de la topología del estado nativo en el plegamiento de proteínas se han obtenido en función de numerosos trabajos con proteínas de tamaños pequeños, casos en los que generalmente no se observan conformaciones no-nativas parcialmente ordenadas de manera experimental [67]. Al contrario, el plegamiento de proteínas de mayor tamaño (>100 residuos

aminoácidos en largo de cadena polipeptídica) generalmente involucra el paso por uno o más intermediarios que son necesarios en la vía de plegamiento para alcanzar el estado nativo de manera eficiente [68], o bien el proceso de plegamiento cinético se encuentra dominado por el escape desde conformaciones no-nativas de baja energía que enlentecen el proceso [36]. Se ha sugerido que estos intermediarios se producen ya que las cadenas polipeptídicas de gran extensión tienden a colapsar en sí mismas en soluciones acuosas, lo que conduce a la formación de estados compactos que pueden contener elementos estructurales reminiscentes de los elementos encontrados en el estado nativo [69]. La propensión de estos intermediarios se incrementa en cuanto la topología es más compleja, es decir, cuando la estabilización está dada por interacciones de largo alcance [70] (en otras palabras, un orden de contacto elevado [18]).

Un gran número de proteínas con cadenas polipeptídicas de más de 100 aminoácidos se componen de varios dominios, cada uno de los cuales alcanza el estado nativo a través de rutas paralelas e independientes [71], acompañado de una etapa final en la cual se establecen los últimos rearrreglos en y entre dominios para establecer el estado nativo funcional [69]. En efecto, en procariontes, cerca del 65% de las proteínas están formadas por combinación de dominios, mientras que en el caso de eucariontes la abundancia de proteínas multidominio es mayor al 80% [72]. Esto permite simplificar el problema al nivel del plegamiento de proteínas pequeñas de un solo dominio [73].

En otros casos, a pesar de la presencia de varios dominios, éstos son discontinuos, ya que se encuentran interrumpidos por la topología de la cadena polipeptídica, de tal manera que existen más de dos puntos de cruces de ésta entre los dominios [74]. Estos eventos topológicos son habituales en proteínas con múltiples dominios, manifestándose como el empacamiento de elementos estructurales de un dominio localizados de manera no contigua en la secuencia primaria [75]. Uno de los mejores ejemplos de este tipo de discontinuidades es el de lisozima T4, en donde una hélice en el extremo amino-terminal forma parte del dominio carboxilo-terminal [76]. Permutaciones circulares para permitir la continuidad en secuencia primaria del dominio

carboxilo-terminal demostraron que la conectividad de los dominios a través de la topología de la cadena polipeptídica es un determinante fundamental de la cooperatividad del plegamiento y que la permutación lleva al desacoplamiento energético de los dominios, por lo que es posible observar el (des)plegamiento independiente de cada uno de ellos [37]. En base a esto, se ha sugerido que la discontinuidad topológica ha surgido como una estrategia evolutiva que promueve la cooperatividad, para así evitar que las proteínas aumenten su grado de frustración y alcancen trampas cinéticas [77].

A la luz de estos antecedentes, es un ejercicio fundamental el determinar el control que ejerce la topología del estado nativo en la función y plegamiento de las enzimas, ya que éstas representan una población significativa de las proteínas totales que se expresan en la célula. Solo en el genoma de *Escherichia coli* es posible detectar 676 genes [78] que codifican para 607 enzimas que participan de las 744 reacciones del metabolismo de moléculas pequeñas [79]. Esto corresponde al 14% de todas las proteínas de *E. coli*, y al 25% si se descartan las proteínas de función desconocida [78]. Estas enzimas son responsables de utilizar diversos metabolitos, cuya concentración intracelular total observada es aproximadamente 300 mM (100 millones de metabolitos por célula), con concentraciones de metabolito por sobre la constante cinética de Michaelis-Menten para gran parte de las enzimas [80]. De estas enzimas, 311 son monoméricas, mientras que el resto son oligoméricas, correspondiendo principalmente a homodímeros y homotetrámeros [79]. Además, cerca del 50% de estas enzimas (271) son comunes entre *Escherichia coli* y *Saccharomyces cerevisiae* [81].

La teoría del paisaje energético [7] y el principio de la mínima frustración [12] ya han sido aplicados para la determinación teórica del rol de la topología en la dinámica estructural y los cambios conformacionales asociados a la función catalítica de enzimas como la quinasa de adenilato [82, 83], el dominio catalítico de la proteína quinasa A [84] y la proteína de fosforilación de la familia Src de quinasas de tirosina Csk [85]. Además, se ha utilizado para estudiar teórica y/o experimentalmente el

plegamiento de lisozima T4 [37, 77, 86] y acil fosfatasa [34]. Estos datos sugieren que la presencia de un mayor grado de frustración energética es necesario para la catálisis enzimática [87] y que, a pesar de la mayor rugosidad del paisaje energético de plegamiento de proteínas con >100 residuos [68], las características elementales de su mecanismo pueden ser comprendidas en función de la topología del estado nativo.

La familia riboquinasa como modelo de estudio del rol de la topología del estado nativo en el plegamiento y asociación homodimérica.

Un modelo de estudio atractivo para establecer el rol de la topología en el plegamiento y función de enzimas lo constituyen los homodímeros pertenecientes a la superfamilia riboquinasa [88], la cual está conformada por enzimas que catalizan la transferencia del fosfato terminal desde un sustrato dador hacia el grupo hidroximetil de diversas coenzimas, azúcares y nucleósidos [89, 90]. La reacción catalizada por estas enzimas requiere de un catión divalente (por lo general, magnesio), el cual forma un complejo metal-nucleótido que corresponde al verdadero sustrato dador de la enzima [91]. Si bien el sustrato dador corresponde generalmente a MgATP [91], algunas enzimas de eucariontes y arqueas utilizan MgADP [92].

La arquitectura común para las enzimas de esta superfamilia corresponde a un plegamiento  $\alpha\beta$  de tipo Rossmann [93], consistente en una sábana  $\beta$  central compuesta de al menos ocho hebras  $\beta$  en disposición principalmente paralela, rodeada de hélices  $\alpha$  por ambas caras [88]. Este dominio (denominado dominio mayor) puede subdividirse en un dominio Rossmann [93], compuesto por seis hebras  $\beta$  y cuatro hélices  $\alpha$  próximas al extremo amino-terminal, complementado por un conjunto de hebras  $\beta$  y hélices adicionales localizadas en la región carboxilo terminal que definen un dominio adicional conocido como  $\beta$ -meandro [94]. Sobre este andamiaje, se suele observar una serie de modificaciones topológicas, como cambios en la conectividad de las hebras de la sábana central, adición de elementos de estructura secundaria [95] o la aparición de un dominio de menor tamaño cuya arquitectura basal corresponde a una sábana  $\beta$  formada por dos

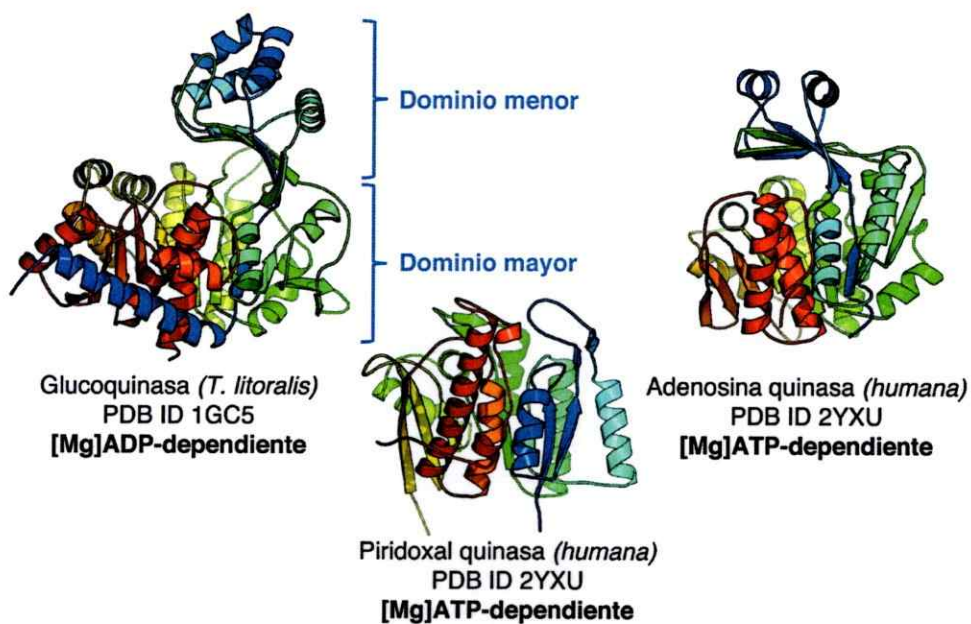


pares de hebras  $\beta$  que emergen desde el dominio Rossmann [94]; es decir, este dominio está inserto en el dominio Rossmann como una discontinuidad topológica [74], de manera que estos dominios están conectados por cuatro segmentos de la cadena polipeptídica.

En base a las diferencias de especificidad por los sustratos aceptores y dadores del grupo fosforilo y a las modificaciones topológicas sobre la arquitectura tridimensional común, la superfamilia riboquinasa puede ser subdividida en tres agrupaciones o familias: las quinasas de azúcares dependiente de ATP que presentan un dominio menor, las quinasas de azúcares dependientes de ADP que presentan un dominio menor y las quinasas de coenzimas dependientes de ATP que no poseen un dominio menor (Figura 1). Las diferencias topológicas observadas entre los miembros de estas familias tienen una estrecha relación con i) el mecanismo catalítico, ii) la composición de residuos del sitio activo, iii) la interfaz de asociación entre subunidades y iv) el estado oligomérico.

Mientras el dominio mayor contiene a la mayor parte de los residuos necesarios para unir los sustratos y facilitar la catálisis enzimática [88], el dominio de menor tamaño actúa como una tapa que se cierra luego de la unión de sustratos [96], posiblemente para protegerlos de la hidrólisis [88], aunque también puede aportar con residuos dirigidos hacia el sitio activo [97]. En las estructuras resueltas mediante cristalografía de rayos X para algunas de las enzimas monoméricas y homodiméricas, es posible observar que la unión de los sustratos dadores y aceptores del grupo fosforilo depende de residuos adicionales provenientes del dominio menor, como en el caso de fosfofructoquinasa-2 de *Escherichia coli*, tagatosa-6-fosfato quinasa de *Staphylococcus aureus* y glucoquinasa de *Thermococcus litoralis* [94, 97, 98, 99].

Los residuos que forman parte del sitio activo de estas enzimas pueden provenir del dominio mayor o menor de la misma subunidad o de subunidades adyacentes [88, 94, 100], lo cual depende del estado de agregación de estas enzimas, que puede corresponder a monómeros (adenosina quinasa [101, 102], glucoquinasa [92],



**Figura 1. Heterogeneidad estructural y funcional de la superfamilia riboquinasa.**

Las enzimas pertenecientes a la superfamilia riboquinasa puede ser agrupadas en tres familias según la composición de dominios del monómero aislado (presencia o ausencia de un dominio de menor tamaño) y el tipo de sustrato dador del grupo fosforilo (ATP o ADP).



fosfofructoquinasa [103]), dímeros (riboquinasa [88], fosfofructoquinasa-2 [104], 4-amino-5-hidroximetil pirimidina quinasa [105], piridoxal quinasa [106, 107]), trímeros (4-metil-5- $\beta$ -hidroxietil tiazol quinasa [100, 108]) y tetrámeros (quinasa putativa YxKO [109]). A su vez, los distintos tipos de estados de agregación dependen de la localización de la superficie de interacción proteína-proteína, la cual se extiende sobre el dominio menor [88, 94], el dominio mayor [105, 106, 109] o los sitios activos [100, 108].

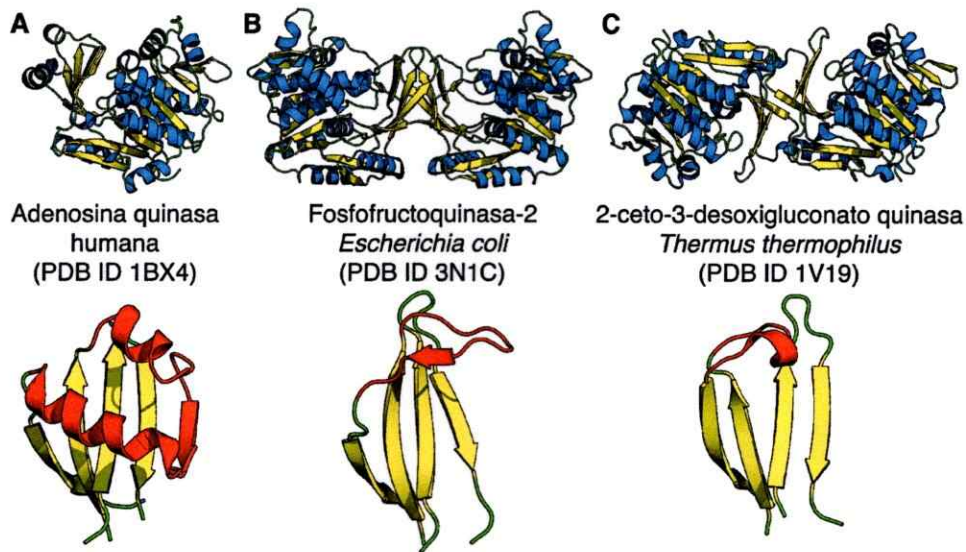
Dentro de la familia de quinasas de azúcares ATP-dependientes que presentan un dominio menor (en adelante, familia riboquinasa), es posible encontrar miembros monoméricos [101, 102] y diméricos [94, 97, 98]. En ambos casos, los residuos fundamentales para la composición del sitio activo provienen tanto del dominio menor como del dominio mayor de la enzima [101, 104]; sin embargo, en el caso de algunas de las enzimas diméricas estos residuos también pueden provenir de la subunidad adyacente, dado que la superficie de interacción proteína-proteína se localiza sobre el dominio menor [88, 94].

La evidencia estructural sugiere que las modificaciones topológicas a nivel del dominio menor son responsables de los cambios en el estado de agregación y del origen de los residuos que conforman el sitio activo. Las enzimas monoméricas presentan estructuras secundarias (mayoritariamente hélices  $\alpha$ ) adicionales en el dominio menor que ocluyen la asociación proteína-proteína al prevenir las interacciones entre las hebras  $\beta$  de dos subunidades [94, 97]. Este es el caso de adenosina quinasa humana (hAK, Figura 2A) y de *Toxoplasma gondii*, en donde el empacamiento de una hebra  $\beta$  y dos hélices  $\alpha$  contra la cara exterior de la sábana  $\beta$  del dominio menor permite su estabilización [101, 102]. Una estrategia similar se observa para la familia de quinasas de azúcares ADP-dependientes, como la glucoquinasa de *Thermococcus litoralis*, en donde son cuatro las hélices  $\alpha$  empacadas contra la sábana  $\beta$  del dominio menor [92].

En el caso de las enzimas diméricas, si bien la sábana  $\beta$  del dominio menor es estabilizada por la homodimerización, existen pequeñas diferencias estructurales que cambian la orientación de los monómeros en el contexto dimérico y, por consiguiente, la

forma general del dímero [110]. En el caso de enzimas como riboquinasa o fosfofructoquinasa-2 (Pfk-2, Figura 2B) de *Escherichia coli*, la segunda hebra  $\beta$  del dominio menor presenta una alteración de la red repetitiva de puentes de hidrógeno (“ $\beta$ -bulge”, [111]), lo que introduce una vuelta en  $90^\circ$  que permite que la región amino-terminal de esta hebra establezca interacciones con la última hebra  $\beta$  carboxilo-terminal del dominio menor de la subunidad adyacente, constituyendo así un dominio bimolecular compacto similar a una estructura de barril achatado (“ $\beta$ -clasp”) [88, 94]. Para otras quinasas homodiméricas como 2-ceto-3-desoxigluconato quinasa (Kdgc) de *Thermus thermophilus* (Figura 2C) y *Thermotoga maritima*, la vuelta entre la primera y segunda hebra  $\beta$  del dominio menor contiene una hélice  $3_{10}$  que impide la formación del dominio bimolecular observado para el homodímero de riboquinasa y fosfofructoquinasa-2 de *Escherichia coli*, por lo que en este caso las subunidades se asocian a través de la interacción directa entre la segunda hebra  $\beta$  del dominio menor de cada subunidad, la cual es facilitada por el posicionamiento antiparalelo de estas hebras, estableciendo así una conformación similar a la de una sábana  $\beta$  antiparalela pronunciadamente curvada [110, 112]. Esto también se observa en las estructuras de Kdgc de *Sulfobulus solfataricus* [113] y aminoimidazol ribosido quinasa de *Salmonella enterica* [95].

Estos antecedentes estructurales indican que las diferencias topológicas a nivel del dominio menor determinan el requerimiento de homodimerización en esta familia de quinasas ATP-dependientes. Considerando además la topología reentrante de la cadena polipeptídica entre el dominio menor y el dominio mayor, estos argumentos sugieren que la estabilidad de la subunidad aislada (y no solo del dominio menor) depende fuertemente de la formación del homodímero y, en consecuencia, que los procesos de disociación y desplegamiento pueden ser cooperativos. Experimentos de desplegamiento químico al equilibrio utilizando Pfk-2 de *Escherichia coli* como modelo de estudio, han demostrado que esta enzima posee un mecanismo de plegamiento de tres estados ( $N \Leftrightarrow 2I \Leftrightarrow 2U$ ), en el cual el paso disociativo  $N \Leftrightarrow 2I$  conduce a la acumulación de un



**Figura 2. Modificaciones topológicas a nivel del dominio menor y efectos en la asociación proteína-proteína de las enzimas de la familia riboquinasa.**

La familia riboquinasa está compuesta por enzimas monoméricas (adenosina quinasa humana) y dimericas (fosfofructoquinasa-2 de *Escherichia coli* y 2-ceto-3-desoxigluconato quinasa de *Thermus thermophilus*), cuya estructura resuelta se muestra en representación de cintas en la parte superior. El estado oligomérico depende de la adición de elementos de estructura secundaria en el dominio menor, representado en la parte inferior, sobre el cual se extiende la interfaz de asociación homodimérica. Estas inserciones estructurales (indicadas en rojo) modifican la forma general del homodímero y la orientación de los monómeros en el contexto oligomérico.

A. La estructura resuelta de adenosina quinasa humana presenta dos hélices  $\alpha$  y una hebra  $\beta$  adicional que ocluyen la sábana  $\beta$  del dominio menor y estabilizan al monómero aislado.

B. La estructura resuelta de fosfofructoquinasa-2 de *Escherichia coli* presenta un quiebre en la segunda hebra  $\beta$  del dominio menor, lo que le permite extenderse y establecer interacciones de puentes de hidrógeno con el dominio menor de la subunidad adyacente, formando un dominio bimolecular.

C. La estructura resuelta de 2-ceto-desoxigluconato quinasa de *Thermus thermophilus* presenta una hélice  $3_{10}$  que impide la formación del dominio bimolecular observado en el caso de fosfofructoquinasa-2 de *Escherichia coli*.

intermediario monomérico marginalmente estable y sin actividad enzimática, el cual posee un radio hidrodinámico mayor al esperado para un monómero compacto, una mayor accesibilidad al solvente y una retención de solamente 30% de la estructura secundaria presente en el dímero [114], lo que sugiere una alta cooperatividad entre los procesos de disociación y desplegamiento. Para otras enzimas de la superfamilia riboquinasa que carecen de dominio menor y, por ende, de la compleja discontinuidad topológica observada para Pfk-2, es posible obtener subunidades aisladas a concentraciones de orden micromolar sin conducir a la formación de un monómero parcial o totalmente desplegado, como en el caso de piridoxal quinasa de oveja [115]. Esto sugiere fuertemente que la cooperatividad entre los procesos de desplegamiento y disociación de subunidades es un fenómeno dependiente de la complejidad topológica y del tipo de interfaz de asociación.

Adicionalmente, los experimentos de desplegamiento químico reversible de piridoxal quinasa de oveja, utilizando como sonda análogos fluorescentes de ATP y piridoxal, demostraron que el subdominio proteico que contiene el sitio de unión a piridoxal (cercano al extremo amino-terminal) se despliega a concentraciones menores de cloruro de guanidino que el subdominio que contiene el sitio de unión de ATP (0.1 M versus 0.75 M) [115]. Esto sugiere que no todas las regiones del dominio mayor, común para todos miembros de la superfamilia, presentan la misma estabilidad frente a agentes caotrópicos.

Considerando los resultados observados para Pfk-2 de *Escherichia coli* [114] y para piridoxal quinasa de oveja [115], y la topología de los miembros de la familia riboquinasa [88, 94], es importante determinar si, en el caso de proteínas homodiméricas, la disociación de las subunidades lleva a un desplegamiento local o global de la proteína, producto de posibles diferencias de estabilidad entre los dominios Rossmann y dominio menor, los cuales están conectados a través de la topología discontinua de la cadena polipeptídica, y el dominio  $\beta$ -meandro en el extremo carboxilo-terminal. También es importante determinar si los cambios topológicos localizados en el

dominio menor de las enzimas homodiméricas de la familia riboquinasa inciden en el acoplamiento entre los procesos de disociación y desplegamiento.

En este trabajo doctoral, se realizó un estudio de la disociación en subunidades de las enzimas Pfk-2 y Kdgk de *Escherichia coli* inducido mediante distintos tipos de perturbaciones globales (cambios de temperatura, adición de agentes caotrópicos) y locales (mutagénesis sitio-dirigida). Tanto este proceso, como el posible desplegamiento parcial o total de estas enzimas, fueron medidos a través de diversas estrategias espectroscópicas y espectrométricas, como ultracentrifugación analítica, espectrometría de masas, dicroísmo circular y difracción de rayos X de ángulo pequeño. Esto, con el fin de determinar i) las regiones locales de la proteína que se despliegan producto de la inducción de la disociación de subunidades; ii) la importancia de las modificaciones topológicas a nivel del dominio menor en la estabilidad conformacional, la interacción proteína-proteína y el acoplamiento entre los procesos de desplegamiento y disociación.

## HIPOTESIS DEL TRABAJO

Las modificaciones topológicas y la discontinuidad polipeptídica del dominio menor, sobre el cual se extiende la interfaz de asociación proteína-proteína de las enzimas homodiméricas Pfk-2 y Kdgk de *Escherichia coli*, determinan tanto el mecanismo de (des)plegamiento y su acoplamiento con el proceso de disociación, como la estabilidad conformacional local de los distintos dominios contenidos en la arquitectura del monómero aislado.



## OBJETIVO GENERAL

Determinar la estabilidad conformacional local y el acoplamiento de los procesos de disociación y desplegamiento de las enzimas Pfk-2 y Kdgk de *Escherichia coli*, a través de la inducción de la disociación de las subunidades mediante diversas perturbaciones globales (cambios de temperatura, adición de agentes caotrópicos) y locales (mutagénesis sitio-dirigida de residuos localizados en la interfaz monómero-monómero) y el seguimiento de estos cambios estructurales por medio de diversas estrategias experimentales clásicas y nóveles.

## OBJETIVOS ESPECIFICOS

1. Describir los cambios estructurales globales y locales, y la concurrencia de los procesos de disociación y desplegamiento de subunidades, durante la desnaturación inducida por exposición a baja temperatura de la enzima homodimérica Pfk-2 de *Escherichia coli*.
2. Caracterizar el ensamble conformacional del monómero aislado mediante perturbaciones globales (cambios de temperatura, adición de agentes caotrópicos) y locales (mutaciones sitio-dirigidas en la interfaz de asociación) de la enzima homodimérica Pfk-2 de *Escherichia coli*.
3. Determinar la estructura cristalográfica de la enzima Kdgk de *Escherichia coli* y describir los cambios en las propiedades estructurales e hidrodinámicas producto de la disociación de las subunidades mediante perturbaciones globales (adición de agentes caotrópicos) y locales (mutaciones sitio-dirigidas en la interfaz de asociación).
4. Optimizar el estudio *in silico* del proceso de desplegamiento y/o de penetración del solvente de proteínas con topologías complejas mediante el uso de modelos simplificados basados en potenciales nativo-céntricos y estrategias de muestreo acelerado.

## SINOPSIS

En el capítulo 1, se describe cinética y termodinámicamente el mecanismo de desnaturación a baja temperatura de la enzima Pfk-2 de *Escherichia coli*, mediante la combinación de diversas estrategias experimentales, como ultracentrifugación analítica, fluorescencia e intercambio isotópico hidrógeno/deuterio seguido por espectrometría de masas (HXMS).

En el capítulo 2, se describen las propiedades hidrodinámicas y estructurales de los homodímeros de Kdgk y Pfk-2 de *Escherichia coli* y de las subunidades aisladas de estas enzimas, obtenidas mediante la adición de agentes caotrópicos y mediante mutaciones sitio-dirigidas de la interfaz de asociación, las cuales son predichas *in silico* a partir de las estructuras cristalográficas resueltas. Estas observaciones experimentales se realizaron mediante ultracentrifugación analítica, dicroísmo circular, HXMS, cristalografía de rayos X y difracción de rayos X de ángulo pequeño.

En el capítulo 3, se describe la implementación de un potencial armónico de muestreo acelerado y un potencial de solvatación de contactos nativos, para el estudio *in silico* del proceso de plegamiento y de penetración del solvente en el centro hidrofóbico de proteínas, utilizando modelos simplificados basados en la estructura nativa de proteínas y simulaciones moleculares.

CAPITULO UNO:  
Desnaturación en frío de Pfk-2 de  
*Escherichia coli*

## 1.1. INTRODUCCION

La enzima fosfofructoquinasa-2 (Pfk-2) corresponde a una de las dos isoenzimas presentes en *Escherichia coli* [116], y cataliza la reacción de fosforilación de fructosa-6-fosfato (Fru-6-P) para formar fructosa-1,6-bifosfato. Esta enzima constituye, junto con la riboquinasa de *Escherichia coli* [117], uno de los miembros mejor estudiados de las quinasa de azúcares de la familia riboquinasa [88].

La enzima Pfk-2 corresponde a un homodímero de aproximadamente 66 KDa (33 KDa por cada subunidad). Mediante estudios de cinética enzimática se demostró que, a diferencia de la isoenzima homotetramérica fosfofructoquinasa-1, Pfk-2 exhibe una curva de saturación hiperbólica para los sustratos dador (MgATP) y aceptor (Fru-6-P) del grupo fosforilo [118]. No obstante, Pfk-2 es inhibida por MgATP cuando la concentración de Fru-6-P es baja [119]. Estudios de fluorescencia intrínseca demostraron que la inhibición es ejercida mediante la unión de este complejo metal-nucleótido a un sitio alostérico presente en la enzima [120]. De esta manera, el complejo MgATP tiene una función doble: actuar como sustrato dador de fosforilo y como regulador alostérico, inhibiendo la reacción catalizada por Pfk-2 [91], lo que también desplaza el equilibrio hacia la formación de un tetrámero [121]. La actividad quinasa sobre Fru-6-P también puede ser inhibida por  $ATP^{4-}$  [91] y fosfato [122], mientras que un segundo catión divalente libre  $Mg^{+2}$  opera como un activador esencial de Pfk-2, siendo incluso capaz de atenuar la inhibición inducida por la unión alostérica de MgATP [122]. Debido a la variedad de efectores fisiológicos que controlan la actividad de esta enzima, se sugirió que Pfk-2 juega un papel muy importante en la regulación de la glicólisis, pese a que su expresión *in vivo* es baja (representa el 5-10% de la actividad quinasa de Fru-6-P total [123]).

Es importante notar que las características estructurales particulares del dímero de Pfk-2 definen en gran parte las propiedades funcionales de esta enzima. En primer lugar, la estabilidad de la estructura cuaternaria es fundamental; la disociación de las

subunidades provoca la pérdida de la actividad enzimática [124]. Esto se debe a las características de la interfaz de asociación entre subunidades, la cual puede ser descrita como un barril  $\beta$  aplanado y cerrado, compuesto por 8 hebras provenientes del dominio menor de ambos monómeros. Este arreglo estructural (Figura 3) es posible debido a la presencia de los residuos Tre31, Ala32 y Pro33 en la hebra  $\beta_3$  de cada dominio menor, los cuales introducen una interrupción en la continuidad de la hebra y una vuelta con un giro de  $90^\circ$  que permite que la parte distal de la hebra  $\beta_3$  interactúe con la hebra  $\beta_7$  del dominio menor adyacente [94]. De esta manera, la interfaz de asociación constituye un dominio tipo barril  $\beta$  cuyo centro hidrofóbico es bimolecular, en donde los residuos Tir23, Lis27 y Arg29 de un monómero se proyectan hacia la subunidad adyacente, estableciendo interacciones importantes con MgATP en el sitio activo y alostérico de Pfk-2, participando así en la actividad enzimática y en la inhibición por sustrato [94, 104].

Se ha sugerido que esta intrincada arquitectura de asociación entre subunidades también incide en el mecanismo de plegamiento de Pfk-2. El desplegamiento reversible inducido a través de la adición de cloruro de guanidinio (GndHCl) seguido por cinética enzimática, fluorescencia intrínseca del residuo Trp88 (cuya ubicación, por debajo del dominio bimolecular de la interfaz de asociación, permite dar cuenta de procesos de cambio en el estado de agregación [125]), dicroísmo circular, cromatografía de exclusión por tamaño y fluorescencia extrínseca de la sonda hidrofóbicasulfonato de 1-anilino-8-naftaleno (ANS) demostraron que i) el mecanismo de plegamiento sigue un modelo de 3 estados,  $N \Leftrightarrow 2I \Leftrightarrow 2D$ ; ii) el intermediario monomérico presenta una mayor accesibilidad al solvente y su estabilidad es marginal en comparación a la del dímero nativo; iii) la transición  $N \Leftrightarrow 2I$  ocurre de manera concertada con la pérdida de actividad enzimática y del 70% de la estructura secundaria observada para el dímero nativo de Pfk-2, lo que sugiere que el proceso de disociación y desplegamiento son cooperativos [114]. El mismo fenómeno de acumulación de un intermediario monomérico hidratado se observa para el desplegamiento del tetrámero de Pfk-2, estado

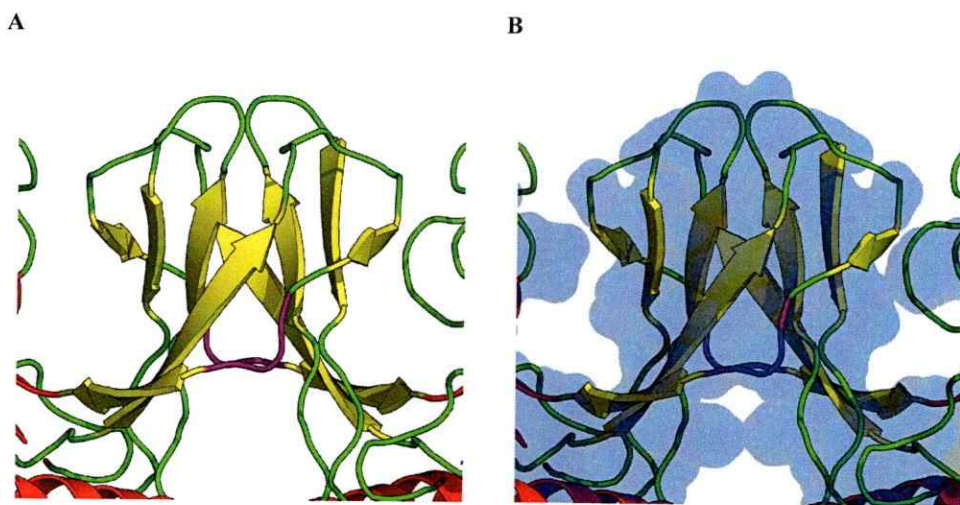
oligomérico inducido por la unión de MgATP al sitio alostérico de la enzima [125]. Estudios de las cinéticas de plegamiento de Pfk-2 por fluorescencia y dicroísmo circular refuerzan esta idea, al demostrar que la asociación es el paso limitante para la formación de gran parte de los contactos nativos de esta enzima [126].

Si bien la cooperatividad entre la disociación y el desplegamiento resulta esperable desde el punto de vista estructural [127] dadas las características de dominio compacto de la interfaz de asociación y la topología reentrante del dominio menor respecto de la arquitectura tipo Rossmann del dominio mayor [95], no es evidente si la pérdida de estructura secundaria y el aumento de la accesibilidad al solvente ocurren local o globalmente a lo largo de la proteína. Al respecto, el análisis mediante dicroísmo circular y fluorescencia intrínseca de la cinética de desplegamiento de Pfk-2 en 2 M de GndHCl describe dos fases, en donde la fase rápida correspondería a la formación de un intermediario dimérico generado por el desplegamiento del módulo  $\beta$ -meandro de la región C-terminal de Pfk-2, de manera independiente del resto de la proteína, seguido de la disociación y desplegamiento del resto de la proteína, alcanzando así el intermediario monomérico observado en condiciones de equilibrio [126]. Sin embargo, es importante señalar que estas herramientas espectroscópicas dan cuenta solo de cambios globales, o en el mejor escenario, sugieren de manera parcial los cambios estructurales que pueden ocurrir localmente en la estructura de una proteína particular. En este sentido, es importante: i) encontrar un fenómeno adecuado para el estudio de estos cambios estructurales en el tiempo; ii) desarrollar un método apropiado para detectar los cambios locales a lo largo de la estructura de Pfk-2 durante este proceso cinético.

Respecto al primer punto, recientemente hemos descrito el desplegamiento a bajas temperaturas de Pfk-2 [128]. El fenómeno de desplegamiento a temperaturas justo por sobre el punto de congelamiento del agua es esperable desde el punto de vista termodinámico. Al expandir la entalpía y entropía en función de la temperatura de transición (donde  $G_N - G_D = 0$ ) a través de la capacidad calórica a presión constante ( $\Delta C_p$ ), la ecuación de Gibbs-Helmholtz describe una curva convexa dependiente de la

temperatura y, como consecuencia, la existencia de dos transiciones de desplegamiento térmico en proteínas. Estas transiciones son conocidas como desplegamiento por calentamiento y desplegamiento por enfriamiento (una excelente descripción de los fundamentos teóricos y el procedimiento algebraico para describir termodinámicamente este fenómeno se encuentran en refs. 129 – 132).

Si bien el desplegamiento a alta temperatura ha sido ampliamente estudiado, el desplegamiento en frío ocurre generalmente a temperaturas bajo el punto de congelamiento del agua y sólo es observable en ciertos modelos proteicos [133, 134] cuyos  $\Delta C_p$  son grandes y positivos y sus  $\Delta H$  son pequeños [135]. Esto ha llevado a que, en la mayoría de los casos estudiados hasta la fecha, se haya recurrido a una perturbación adicional (presencia de agentes caotrópicos, aumento de presión, cambios de pH, mutaciones) para que este proceso pueda ser estudiado *in vitro* sobre la temperatura de congelamiento [136, 137]. Además, existe una fuerte discusión acerca de si este proceso de desplegamiento es cooperativo (como en el caso de ubiquitina, [138]) o no (como en el caso de Yfh1 y CTL9, [133, 139]), y si es similar o no al fenómeno de desplegamiento a alta temperatura [140]. Esto se debe a que en experimentos de escaneo diferencial de calorimetría se ha observado que el desplegamiento a alta temperatura es un proceso endotérmico, mientras que el proceso de desplegamiento en frío generalmente no muestra cambio en entalpía ( $\Delta H \sim 0$ ). Sin embargo, estudios recientes han demostrado que esta caracterización energética anómala no se debe a la ausencia de cooperatividad, sino que es un efecto aparente como consecuencia de la gran barrera energética existente entre los estados plegado y desplegado a bajas temperaturas, lo que impone una estabilidad cinética frente al proceso de desplegamiento en frío [141]. De este modo, la velocidad de este fenómeno a baja temperatura es, por lo general, muy lenta, alcanzando el estado final de este proceso en el transcurso de días e incluso semanas. Este proceso es notablemente acelerado en presencia de una perturbación adicional [141]. Por consiguiente, estos resultados demuestran la importancia del análisis de la cinética de desplegamiento y replegamiento de proteínas en el estudio *in*



**Figura 3. Visualización de la estructura del dominio bimolecular de Pfk-2.**

Este dominio de aspecto tipo barril  $\beta$  aplanado, se forma producto de la interacción del dominio menor proveniente de cada subunidad de la enzima Pfk-2. **A.** Representación en cintas de la estructura secundaria de este dominio, correspondiente a hebras beta (amarillo). Los residuos T31, A32 y P33 que permiten la interacción de la hebra  $\beta 3$  con la  $\beta 7$  del dominio adyacente se indican en color rosa. **B.** Representación en cintas equivalente a Figura 3A, incluyendo la superficie correspondiente a residuos hidrofóbicos del dominio bimolecular de Pfk-2 en azul.



*vitro* de la denaturación en frío, en particular en ausencia de condiciones desestabilizantes, para una adecuada caracterización estructural y para la determinación de la energética y cooperatividad de este proceso.

Para estudiar este fenómeno en Pfk-2, primero se optó por la determinación del contenido de estructura secundaria y terciaria y de los cambios en el estado de agregación de la enzima Pfk-2 a distintas temperaturas y en condiciones de equilibrio termodinámico mediante metodologías espectroscópicas clásicas, como dicroísmo circular, fluorescencia intrínseca del residuo Trp88, fluorescencia extrínseca de la sonda hidrofóbica ANS, dispersión dinámica de la luz (DLS) y experimentos de velocidad de sedimentación por ultracentrifugación analítica (AUC). Exceptuando las metodologías de DLS y AUC, estas técnicas espectroscópicas han sido generalmente utilizadas para el estudio del desplegamiento de proteínas a baja temperatura [142], junto con calorimetría [137], resonancia magnética nuclear [133] y dinámica molecular [143]. Sin embargo, con la notable excepción de RMN, estas técnicas solo dan cuenta de propiedades globales y no permiten obtener información sobre cambios estructurales locales. Por otro lado, el uso de la RMN está limitado por el tamaño del modelo de estudio y por el marcaje de la proteína con isótopos de H, N y C, por lo que en casos como el de Pfk-2 el estudio a través de esta metodología es complejo.

Por ende, para la determinación de los cambios estructurales durante la cinética de desplegamiento en frío de la enzima Pfk-2, hemos optado por la combinación de metodologías clásicas (fluorescencia intrínseca y AUC) junto con el desarrollo de una estrategia de cuantificación del intercambio de hidrógeno por deuterio entre las amidas de la cadena principal de la proteína y un solvente deuterado (generalmente agua deuterada) mediante espectrometría de masas. Esta estrategia experimental (en adelante HXMS) es una gran herramienta para el estudio de interacciones proteína-proteína y para la identificación de interfaces de asociación [144], cambios en la dinámica estructural como consecuencia de la interacción con otras proteínas [145] e identificación de regiones intrínsecamente desordenadas [146], permitiendo obtener

información estructural a partir del análisis y cuantificación del intercambio isotópico. En ésta, la reacción de intercambio isotópico es detenida a distintos tiempos, las muestras son proteolizadas y los péptidos resultantes son analizados mediante espectrometría de masas, permitiendo obtener, a partir del análisis y cuantificación del número de deuterones incorporados, información estructural de los cambios ocurridos sobre regiones locales de la proteína en estudio. Dado que el intercambio ocurre entre la proteína y el solvente circundante, el grado de intercambio es una medida de la accesibilidad al solvente de las amidas de la cadena principal [147].

Estas estrategias experimentales permitieron la caracterización estructural global del proceso de desplegamiento en frío de la enzima Pfk-2 en ausencia de perturbaciones adicionales. Demostramos que a 3°C se alcanza un intermediario monomérico que reúne propiedades similares a las del monómero parcialmente desplegado inducido por GndHCl. Estos resultados, junto a la caracterización de la cinética de cambio estructural local, nos permitió demostrar que el proceso ocurre a lo largo de toda la estructura proteica de manera concurrente con el proceso de disociación de subunidades, alcanzando un estado intermediario penetrado por el solvente, cuya incorporación de deuterio es ~2.5 mayor que en el caso de la enzima en condiciones nativas a 25°C [148].

## 1.2. PRINCIPALES RESULTADOS Y CONCLUSIONES

Mediante mediciones de dicroísmo circular y fluorescencia intrínseca del residuo Trp88 de la enzima Pfk-2 a distintas temperaturas, se observó la presencia de transiciones de desplegamiento a baja y alta temperatura en ausencia de perturbaciones adicionales [128]. En el caso de la transición de desplegamiento en frío, ésta se completa luego de la exposición por alrededor de 24 h a una temperatura cercana a los 3°C, justo por sobre el punto de congelamiento del agua. Como es esperado para una enzima dimérica como Pfk-2, el aumento de la concentración de proteína provoca una disminución de la temperatura de transición en frío y un aumento de la temperatura de transición por calentamiento, lo que es consistente con un modelo de disociación

dímero-monómero. Esta disociación de subunidades fue confirmada mediante experimentos de sedimentación por ultracentrifugación analítica y dispersión dinámica de luz a distintas temperaturas. Mediante el uso de la sonda fluorescente extrínseca ANS, se demostró que la especie inducida por frío no representa el estado completamente desplegado, sino que un intermediario monomérico que presenta regiones hidrofóbicas expuestas, sugiriendo su similitud con el intermediario monomérico observado en el desplegamiento inducido por GndHCl. Esta idea fue confirmada al observar la transición desde la especie monomérica a 3°C hasta el estado desplegado mediante la adición de GndHCl. Los experimentos realizados y los parámetros termodinámicos calculados demostraron que: i) la transición dímero-monómero inducida por frío es altamente cooperativa y energéticamente equivalente a la misma transición inducida por GndHCl; ii) el desplegamiento químico del monómero inducido por frío y por GndHCl presentan un cambio de energía libre similar. Dada la aparente lentitud del proceso de desplegamiento en frío de Pfk-2 y de otros modelos proteicos [141], éste resulta adecuado para realizar estudios de cambios estructurales locales en función del tiempo.

Para poder determinar los cambios estructurales globales y locales ocurridos durante la transición  $N \Leftrightarrow 2I$  inducida por la denaturación en frío de Pfk-2, primero se cuantificó la población del intermediario monomérico en condiciones de equilibrio a distintas temperaturas, mediante ultracentrifugación analítica. Luego, se determinó la tasa de denaturación de Pfk-2 a 3°C mediante fluorescencia intrínseca. Finalmente, se desarrolló una estrategia de HXMS para seguir los cambios estructurales locales de Pfk-2 en el tiempo durante el desplegamiento por incubación de la enzima en frío [148]. Se observó que la disociación total entre las subunidades de Pfk-2 ocurre a 3°C, por encima del punto de congelamiento del agua, y que este proceso es lento, ocurriendo en una escala de tiempo de horas, sin la acumulación de monómeros compactos. El intercambio de hidrógeno por deuterio aumenta a lo largo de toda la estructura proteica a medida que ocurre el proceso de denaturación, hasta alcanzar un estado intermediario monomérico final cuyo intercambio es menor al esperado para una proteína sin estabilidad

conformacional (es decir, un estado completamente desplegado). La tasa de intercambio isotópico calculada para las amidas de la cadena principal de Pfk-2 es similar para todos los péptidos analizados, incluyendo un péptido localizado en la interfaz de asociación, y está en buen acuerdo con la velocidad de denaturación determinada mediante fluorescencia intrínseca. Estos resultados no solo confirman las características estructurales del intermediario inducido por el fenómeno de denaturación en frío, sino que constituyen una demostración experimental de las teorías que sugieren que el mecanismo molecular de la denaturación en frío corresponde a la penetración del solvente en el centro de plegamiento de las proteínas [149].

Los resultados y conclusiones del análisis cinético y termodinámico de la denaturación en frío de la enzima Pfk-2 se presentan de manera extendida a continuación, mediante la presentación íntegra de los dos artículos publicados referentes a este trabajo [128, 148].

### 1.3. AGRADECIMIENTOS

El capítulo 1 aparece en *Biophysical Journal*, (2012), Báez, M., Wilson, C. A. M., Ramírez-Sarmiento, C. A., Guixé, V., Babul, J.; y en *Biophysical Journal*, (2013), Ramírez-Sarmiento, C. A., Báez, M., Wilson, C. A. M., Babul, J., Komives, E. A., Guixé, V.; éste último fue seleccionado para la portada del fascículo correspondiente de *Biophysical Journal*. El autor de este trabajo doctoral es coautor e investigador secundario del primer artículo, y es autor e investigador primario del segundo artículo. Los autores agradecen al Dr. Richard Garratt, de la Universidad de Sao Paulo, por la lectura crítica del manuscrito *Biophysical Journal*, (2012), Báez, Wilson, Ramírez-Sarmiento, Guixé, Babul; al Dr. Jon Fitchett, de Lilly Biotechnology Center, San Diego, California, Estados Unidos, por brindarnos acceso al espectrómetro de masas Synapt G2; y a la Dra. Elizabeth A. Komives, por su participación en el desarrollo e implementación de los experimentos de HXMS, los cuales fueron realizados en la Universidad de California, San Diego, Estados Unidos. Los autores también agradecen

al Grupo de Cristalografía, Instituto de Física de Sao Carlos, Universidad de Sao Paulo, Brasil, por la facilitación del equipamiento de dispersión dinámica de luz, y al Centro de Biofísica, Departamento de Química y Bioquímica, Universidad de California, San Diego, por la facilitación del equipamiento de ultracentrifugación analítica. Este trabajo recibió financiamiento del Fondo Nacional de Desarrollo Científico y Tecnológico (FONDECYT) Nos. 1090336 y 1110137. CAR-S y CAMW fueron beneficiarios de una beca de estudios doctorales de la Comisión Nacional de Investigación Científica y Tecnológica (CONICYT). CAMW fue beneficiario de una beca de apoyo a tesis doctoral de CONICYT No. 24090160 y un premio de apoyo a la investigación de Fermelo Chile S.A. CAR-S fue beneficiario de una beca de apoyo a la investigación en el extranjero del organismo de Mejoramiento de la Calidad y Equidad de la Educación Superior (MECESUP) en conjunto con el programa de Ciencias, Tecnología e Innovación para las Americas (STIPAS, *Sciences, Technology and Innovation Program for the Americas*) No. UCH0717 y de una beca de apoyo a la investigación en el extranjero de MECESUP No. UCH0713.

## Expanded Monomeric Intermediate upon Cold and Heat Unfolding of Phosphofructokinase-2 from *Escherichia coli*

Mauricio Baez,<sup>†</sup> Christian A. M. Wilson,<sup>‡</sup> César A. Ramírez-Sarmiento,<sup>‡</sup> Victoria Guixé,<sup>‡</sup> and Jorge Babul<sup>†\*</sup>

<sup>†</sup>Departamento de Bioquímica y Biología Molecular, Facultad de Ciencias Químicas y Farmacéuticas and <sup>‡</sup>Departamento de Biología, Facultad de Ciencias, Universidad de Chile, Santiago, Chile

**ABSTRACT** Folding studies have been focused mainly on small, single-domain proteins or isolated single domains of larger proteins. However, most of the proteins present in biological systems are composed of multiple domains, and to date, the principles that underlie its folding remain elusive. The unfolding of Pfk-2 induced by GdnHCl has been described by highly cooperative three-state equilibrium ( $N_2 \leftrightarrow 2I \leftrightarrow 2U$ ). This is characterized by a strong coupling between the subunits' tertiary structure and the integrity of the dimer interface because "I" represents an unstructured and expanded monomeric intermediate. Here we report that cold and heat unfolding of Pfk-2 resembles the  $N_2 \leftrightarrow 2I$  step of chemically induced unfolding. Moreover, cold unfolding appears to be as cooperative as that induced chemically and even more so than its heat-unfolding counterpart. Because Pfk-2 is a large homodimer of 66 kDa with a complex topology consisting of well-defined domains, these results are somewhat unexpected considering that cold unfolding has been described as a special kind of perturbation that decouples the cooperative unfolding of several proteins.

### INTRODUCTION

The folding of small proteins has proven to be very informative for understanding the folding code of single-domain proteins. However, large and multidomain proteins with molecular masses  $>20$  kDa are much more representative in typical biological systems (1). Phosphofructokinase-2 (Pfk-2) from *Escherichia coli*, a member of the ribokinase family of enzymes (2), is a medium-sized homodimer, composed of 33-kDa subunits, which show a complex topology. Each subunit is composed of a large domain based on the architecture of a  $\alpha/\beta/\alpha$  sandwich and a four-stranded  $\beta$ -sheet that covers the active site as a lid. The two lids of the dimer are packed together to form an orthogonal  $\beta$ -barrel—an intertwined structure, stabilized by a single hydrophobic core that includes residues from both subunits, called a " $\beta$ -clasp" (3). Thus, the overall architecture of Pfk-2 and its close homologs can best be described as a homodimer composed of two  $\alpha/\beta/\alpha$  domains and one bimolecular domain (the  $\beta$ -clasp) that unites them (see Fig. 1). The complexity of the fold is further enhanced by the fact that the large domain is discontinuous, i.e., it is not formed by a contiguous stretch of polypeptide chain, but instead has a small four-stranded  $\beta$ -sheet inserted within it.

We have studied how the folding and subunit association of Pfk-2 occurs, looking for equilibrium or transient intermediates accumulated during guanidine hydrochloride (GdnHCl)-induced unfolding or folding experiments. No compact monomers were detected under equilibrium conditions (4) or during folding (5), supporting the fact that the

$\beta$ -clasp and an important portion of the  $\alpha/\beta/\alpha$  domain do not seem to behave independently with respect to their conformational stability and structural integrity. However, global unfolding transitions induced by chaotropic agents can hide intermediates or low-energy structures populated under native conditions, as has been shown by hydrogen-exchange experiments (6).

Given a large and positive  $\Delta C_p$  value, usually observed for the overall unfolding of a polypeptide chain (7,8), the thermal stability of proteins is dominated by concave free-energy dependence with temperature. Accordingly, the Gibbs-Helmholtz equation for a two-state system predicts two thermal unfolding transitions for a protein—cold unfolding and heat unfolding (9). Nevertheless, protein unfolding induced by cold has been documented for only a few proteins because nonnative species induced by cold are predicted to be populated below the freezing point of water.

Interestingly, some studies have shown that cold and heat unfolding might occur by different mechanisms because equilibrium intermediates are frequently found along the cold- but not the heat-induced transition pathway (10–15), although some proteins seem to unfold by a two-state mechanism at both temperature extremes (16–18). Freire et al. (19) have rationalized the different mechanism for both events as a natural consequence of the thermodynamics associated with exposing hydrophobic surfaces upon partial unfolding of a given structure. Within this framework, cold unfolding is proposed to be a special kind of perturbation able to probe the presence of low-energy substructures populated under native conditions but hidden upon heat or chemical perturbation.

In this work, the reversible thermal unfolding of Pfk-2 was characterized under equilibrium conditions. Strikingly, in addition to the typical heat-induced transition, Pfk-2 shows a clear cold-induced transition which is completed at

Submitted June 18, 2012, and accepted for publication September 28, 2012.

\*Correspondence: jbabul@uchile.cl

Christian A. M. Wilson's present address is QB3 Institute, University of California at Berkeley, Berkeley, CA.

Editor: Patricia Clark.

© 2012 by the Biophysical Society  
0006-3495/12/11/2187/8 \$2.00

<http://dx.doi.org/10.1016/j.bpj.2012.09.043>



FIGURE 1 (Ribbon representation) Homodimeric structure of Pfk-2 (PDB:3CQD). For a single subunit, the five  $\beta$ -strands are located in the  $\beta$ -clasp interface (red); large domain (gray). (Sticks) Side chain of a single tryptophan residue (W88) on the interface between subunits, used for intrinsic fluorescence measurements. Figure was prepared with the software PyMol (www.pymol.org).

$\sim 3^\circ\text{C}$ . Different from several other proteins, the direct observation of a cold-induced transition in Pfk-2 did not require such additional perturbations as chaotropic agents (11,20), high pressure (21), extreme pH (22), or point mutations (16), typically used to decrease the protein stability to allow us to observe cold unfolding transitions above the freezing point of water. Cold-induced unfolding of Pfk-2 seems to be a highly cooperative process significantly disrupting the dimer structure without generating additional intermediates, as was also found to be the case in chemical unfolding.

## MATERIALS AND METHODS

### Purification and storage of Pfk-2

Pfk-2 was purified and stored as described by Babul (23). At least four independent preparations of Pfk-2 were used in this work. The specific activity of Pfk-2 (80–100 U/mg) was used to check the quality of the protein preparations. Previous to the unfolding or refolding experiments, the storage buffer was changed to standard buffer (50 mM Tris pH 8.2, 5 mM  $\text{MgCl}_2$ , and 2 mM dithreitol) by using a HiTrap desalting column (Amersham Biosciences, Uppsala, Sweden). The enzyme was concentrated using a Centricon-60 concentrator (Amicon, Beverly, MA) and further concentration was achieved using a Microcon Ultracel YM-10 unit (Millipore, Billerica, MA). Protein concentration was determined by the Bradford assay (24) using bovine serum albumin as the standard and is expressed in terms of the monomer concentration. Pipette tips, cuvettes, and all material necessary for the manipulation of the enzyme were equilibrated at the working temperature before withdrawing the protein from the Eppendorf tube.

### Circular dichroism measurements

Circular dichroism (CD) spectroscopy was performed using a model No. J-810 instrument (JASCO, Essex, UK). The spectra were measured in the far ultraviolet region, from 260 to 190 nm. Cuvettes with optical paths of 0.1, 1, and 10 mm were employed depending on the protein concentration used. Each spectrum was obtained from the accumulation of at least three scans at the working temperature.

### Intrinsic and extrinsic fluorescence measurements

Measurements were made using a model No. PC-5031 spectrofluorimeter (Shimadzu, Kyoto, Japan). Protein samples at several different GdnHCl

concentrations or temperatures were excited at 295 nm and emission spectra were recorded from 300 to 480 nm. The concentration of the stock 8-anilino-1-naphthalene sulfonic acid (ANS) solution (Molecular Probes, Eugene, OR) was determined using a  $\epsilon = 7800 \text{ M}^{-1} \text{ cm}^{-1}$  at 372 nm in methanol. Pfk-2 samples unfolded at different GdnHCl concentrations were incubated in 100  $\mu\text{M}$  ANS for 3 h in the dark at 3 or  $20^\circ\text{C}$ . The mixture was excited at 380 nm and the emission spectra recorded from 400 to 580 nm.

### Sedimentation experiments

Sedimentation velocity experiments were conducted on an Optima XL-I analytical ultracentrifuge (Beckman-Coulter, Palo Alto, CA) equipped with an An-60 Ti rotor. Experiments were performed at 25 and  $4^\circ\text{C}$  using a rotor speed of 41,000 rpm and a protein concentration of  $\sim 40 \mu\text{M}$ . Pfk-2 was incubated at  $3^\circ\text{C}$  for 24 h in standard buffer (using tris(2-carboxyethyl)phosphine instead of dithreitol as the reducing agent) and the centrifuge, rotor, and centerpieces were cooled down to  $4^\circ\text{C}$  before centrifugation. The moving boundary was monitored by absorbance optics at 280 nm. Sedimentation velocity data were analyzed using the software program SEDFIT (25) and the continuous size-distribution model was chosen to generate the sedimentation coefficient distribution plots (25). It is worth noting that this model can be applied to interacting molecules if the reaction is slow on the timescale of sedimentation and if the species are stable during sedimentation.

### Dynamic light scattering measurements

Dynamic light scattering (DLS) measurements were performed on a DynaPro-MS800 instrument (Protein Solutions, Chicago, IL). Before starting the experiment, the solution was centrifuged for 10 min at 14,000 rpm in an Eppendorf microfuge to remove any particulate matter. Cold-induced transitions were obtained by increasing the temperature of the sample, previously unfolded at  $3^\circ\text{C}$ , using a protein concentration of 30  $\mu\text{M}$ . The hydrodynamic radius was calculated via the Stokes-Einstein equation from the diffusion coefficient, which was obtained from the measured autocorrelation function, using the DYNAMICS software supplied with the instrument (Protein Solutions).

### Unfolding induced by GdnHCl and low temperatures

Cold unfolded Pfk-2 was diluted at various GdnHCl concentrations and left for 24 h to reach equilibrium at  $3^\circ\text{C}$ . Each property, intrinsic and extrinsic fluorescence, circular dichroism (CD), and dynamic light scattering (DLS), was measured at  $3^\circ\text{C}$  in the presence of different GdnHCl concentrations. Then, the sample temperature was increased to  $20^\circ\text{C}$ , until reaching a stable signal for the measured properties (at least 6 h) to get equilibrium conditions.

### Model analysis of the transitions obtained by varying the GdnHCl concentration or temperature

The stability curves obtained in the presence of GdnHCl were analyzed according to a two-state  $1 \leftrightarrow U$  or three-state  $N_2 \leftrightarrow 2I \leftrightarrow 2U$  model (4).

The dependence of  $\Delta G$  on the temperature is given by the Gibbs-Helmholtz equation,

$$\Delta G(T) = \Delta H(T_g) \left(1 - \frac{T}{T_g}\right) + \Delta C_p \left[T - T_g - T \ln\left(\frac{T}{T_g}\right)\right],$$

where  $\Delta H(T_g)$  is the enthalpy of the reaction at  $T_g$  (the temperature for which  $\Delta G(T)$  is zero) and  $T$  is the measured temperature.  $\Delta C_p$  is the change

of specific heat of the reaction assumed to be constant in the temperature interval studied (9). For each temperature,  $\Delta G(T)$  was calculated using a bimolecular two-state model of unfolding  $N_2 \leftrightarrow 2I$ . Additional relationships between the observed dichroic signal at 220 nm and  $\Delta G(T)$  were derived as indicated in several publications (18,26).

## RESULTS

### Thermal unfolding of Pfk-2

Fig. 2 shows the temperature dependence of the CD signal of Pfk-2 at 222 nm. Clearly, the enzyme undergoes two thermal transitions at low (cold-unfolding) and elevated (heat-unfolding) temperatures within a short temperature interval (3–53°C). These transitions were reversible and were obtained under equilibrium conditions (see section S1 in the Supporting Material). As should be noted, cold and heat unfolding of Pfk-2 were captured completely because both transitions end with a substantial decrease of the CD signal at 222 nm defined by a clear nonnative plateau below 4°C. The observation of this behavior is uncommon in the absence of additional perturbations such as chaotropic agents (11,20), high pressure (21), extreme pH (22), or point mutations (16).

As shown in Fig. 3 A, the CD spectra taken between 3 and 23°C show a clear isosbestic point at ~207 nm, suggesting that cold unfolding of Pfk-2 is a two-state process. However, the CD spectra taken during the heat-induced transition did not show a clear isosbestic point (Fig. 3 B), which could be due to irreversible reactions. The refolding yield after heating the protein to 53°C was 80% (see section S2 in the Supporting Material), after which temperature, the thermal unfolding becomes completely irreversible (Fig. 2, shaded symbols; Fig. 3 B, shaded lines).

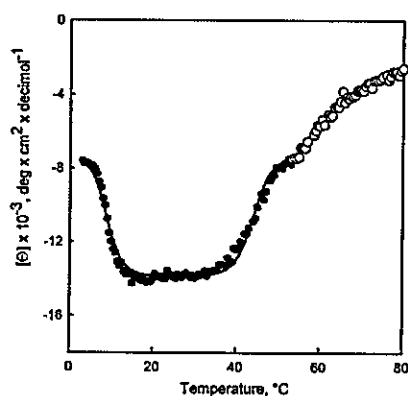


FIGURE 2 Cold and heat transitions of Pfk-2. The mean-residue molar ellipticity obtained at 220 nm is plotted as a function of the temperature. Above 53°C, heat unfolding becomes irreversible (shaded symbols). To reach equilibrium conditions, spectra were taken by increasing the temperature from 3°C at a velocity of 2°C/h. Representative results of three independent experiments are presented.

GdnHCl-induced dissociation of Pfk-2 is coupled with an extensive unfolding of the subunits (4,28) so the effect of the protein concentration on both cold and heat transitions was determined. As indicated in Fig. 4, increasing the protein concentration shifts the cold-induced transition toward lower temperatures and the heat-induced transition toward elevated temperatures (*inset*, Fig. 4). This behavior is expected for a dissociative equilibrium and has been reported for other dimeric proteins (18,29) supporting the fact that both transitions involve subunits dissociation.

Analytical ultracentrifugation measurements performed at 4°C confirm that Pfk-2 is dissociated at lower temperatures. As shown in Fig. 5, the sedimentation coefficient of Pfk-2 at this temperature is nearly half that observed at 25°C, although there is no complete dissociation. This can be seen from the presence of a small peak that matches the sedimentation coefficient for the dimer and accounts for 18% of the total signal (Table 1).

To gain further insight into the hydrodynamic changes that occur on cold treatment, the hydrodynamic radius of Pfk-2 was measured by DLS at several temperatures using a protein concentration of 30  $\mu$ M (Fig. 4). As can be seen, the cold-induced transition revealed by CD was accompanied by a substantial increase to the hydrodynamic radius ( $R_h$ ) reaching values of  $4.6 \pm 0.3$  nm at 3°C.

### Thermodynamic analysis

Taking into account the isosbestic point of the cold-induced unfolding together with the protein concentration dependence of the observed thermal transitions, the Gibbs-Helmholtz equation for a dissociative two-state system (see Materials and Methods) was used to fit both the heat- and cold-induced transitions. The continuous line in Fig. 2 represents the best fit for a dissociative two-state mechanism of unfolding operating at low and high temperatures. Nevertheless, there is a deviation at high temperatures, suggesting that a two-state model could be too simple.

The thermodynamic parameters are shown in Table 2 and were used to build the stability curve of Pfk-2 (Fig. 6). It is worth noting that the stability of Pfk-2 at 20°C predicted by the adjustment (~11 kcal/mol) is similar to previous estimates calculated from the dissociation and unfolding step  $N_2 \leftrightarrow 2I$  obtained in the presence of GdnHCl ( $12 \pm 0.6$  kcal/mol (4)). As expected from the large curvature of the stability curve (Fig. 6), thermal unfolding of Pfk-2 is dominated by a large  $\Delta C_p$  value (~7.8 kcal/mol $^\circ$ K), one of the largest reported in literature to date for a protein (30).

### GdnHCl-Induced unfolding of the cold denatured state of Pfk-2

Further insights about the structural properties of the cold unfolded state of Pfk-2 were obtained by comparing the



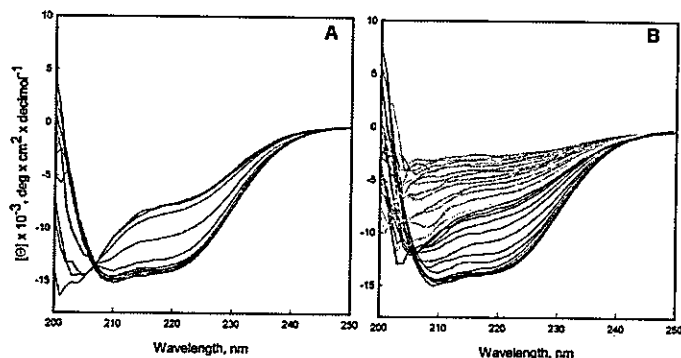


FIGURE 3 Cold and heat denaturation of Pfk-2 followed by circular dichroism (CD). CD spectra obtained from 3 to 23°C (A) and from 23 to 80°C (B). In panel B, the CD spectra obtained above 53°C are shown (shaded representation) where heat denaturation of Pfk-2 becomes irreversible.

GdnHCl-induced unfolding transitions starting from the native dimer at 20°C (Fig. 7, *solid symbols*) or from the cold unfolded state at 3°C (Fig. 7, *open symbols*). GdnHCl-induced transitions obtained from the heat-unfolded state were not pursued, due to the presence of aggregates that formed during the large incubation times required to perform these experiments.

Under equilibrium conditions, GdnHCl-induced unfolding of the native dimer has been described by a three-state

equilibrium process ( $N_2 \leftrightarrow 2I \leftrightarrow 2U$ ) characterized by a poorly structured and expanded monomeric intermediate, "I." The presence of a monomeric intermediate can be deduced from the two transitions separated by a wide plateau when the secondary structure change is monitored by CD at 222 nm (Fig. 7 A, *solid symbols*). The sharper CD transition observed at low GdnHCl concentrations represents the  $N_2 \leftrightarrow 2I$  step whereas the second one represents the  $I \leftrightarrow U$  step of unfolding (4). This intermediate was not observed during the GdnHCl-induced unfolding at 3°C, which shows a unique transition whose amplitude and shape resemble that of the  $I \leftrightarrow U$  step measured at 20°C. Indeed, the native baseline measured by CD at 3°C was coincident with the plateau that separates both transitions at 20°C (Fig. 7 A), suggesting that the cold-induced species may be similar or identical to the intermediate "I."

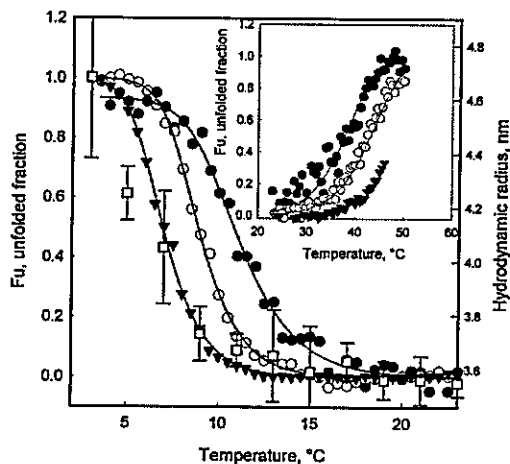


FIGURE 4 Effect of the protein concentration on heat- and cold-induced transitions and variation of Rh as measured by DLS. (Main graph) Fraction of unfolded protein within the cold temperature regime as estimated from the CD signal at 220 nm for several protein concentrations: 0.36  $\mu\text{M}$  ( $\bullet$ ), 3.6  $\mu\text{M}$  ( $\circ$ ), and 45  $\mu\text{M}$  ( $\blacktriangledown$ ). (Inset) Effect of protein concentration on the heat denaturation transition of Pfk-2 obtained at 0.36  $\mu\text{M}$  ( $\bullet$ ), 3.6  $\mu\text{M}$  ( $\circ$ ), and 30  $\mu\text{M}$  ( $\blacktriangledown$ ). The presence of visible aggregates prevented the evaluation of the entire heat transition using 30  $\mu\text{M}$  of protein ( $\blacktriangledown$ , inset). (Main graph,  $\square$ ) Hydrodynamic radius estimated by DLS as a function of the temperature. Rh values were taken by increasing the temperature in 2°C steps after equilibration for 1 h of a 30  $\mu\text{M}$  sample previously unfolded at 3°C.

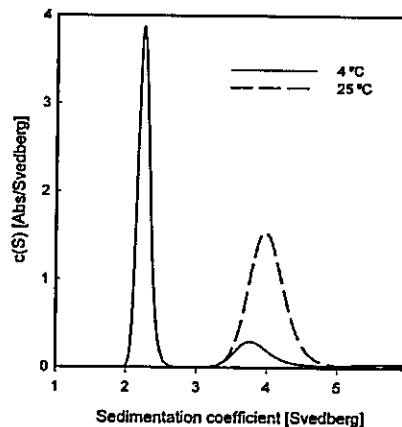


FIGURE 5 Dissociation of Pfk-2 induced by cold. Sedimentation velocity experiments were performed at 4°C (continuous line) and 25°C (dotted line) and protein samples were spun at 41,000 rpm. Pfk-2 was incubated on ice for 24 h before centrifugation. Sedimentation coefficient distributions were generated using a continuous size-distribution model.

**TABLE 1** Fractional population of the species observed by sedimentation velocity

	Monomer (%)	Dimer (%)	Other signals (%)
25°C	-	85 ± 2	15 ± 2
4°C	74 ± 2	18 ± 3	8 ± 2

Relative amounts of all species were calculated by integration of the area under each peak. The errors represent the standard deviation calculated from three independent experiments.

The similarity of the intermediate I to the cold-induced state is also supported by the GdnHCl-induced changes of the Trp-88 fluorescence and extrinsic fluorescence of ANS measured at 3°C. The monomeric intermediate of Pfk-2 must have a nonnative environment of its single Trp similar to that of the fully unfolded state (4) because the equilibrium transition observed by intrinsic fluorescence at 20°C fails to reproduce the I ↔ U step but describes quite well the sharp N<sub>2</sub> ↔ 2I transition (Fig. 7 B, solid symbols). In agreement with this behavior, the intrinsic fluorescence of Pfk-2 measured at 3°C did not show major changes as a function of the GdnHCl concentration (Fig. 7 B, open symbols) suggesting that the intermediate induced by cold also lacks the native environment of Trp-88.

Another feature of the monomeric intermediate is the presence of solvent-exposed hydrophobic regions (4). The extrinsic fluorescence of the hydrophobic probe ANS (31) shows an asymmetrical bell-shape curve that follows the fractional population variation of the monomeric intermediate deduced from the three-state equilibrium (N<sub>2</sub> ↔ 2I ↔ 2U) at 20°C (Fig. 7 C, inset, solid symbols). Conversely, the ANS intensity measured at 3°C shows a decrease as the GdnHCl concentration increases, suggesting that the cold-induced intermediate has solvent-exposed hydrophobic regions in the absence of GdnHCl (Fig. 7 C, open symbols). The fractional population change of Pfk-2, extracted upon fitting the single CD transition observed at 3°C to a two-state mechanism (Table 3), is in good agreement with this observation and matches quite well with the decrease of the ANS intensity induced by GdnHCl at 3°C (Fig. 7 C, open symbols).

**TABLE 2** Thermodynamic parameters for the thermal unfolding of Pfk-2

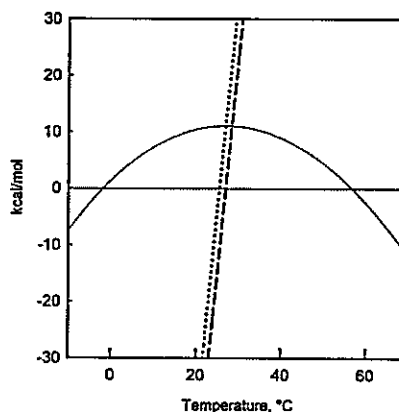
$\Delta H$ at $T_g = 57^\circ\text{C}^\dagger$	247 ± 10 kcal/mol
$\Delta H$ at $T_g = 0^\circ\text{C}^\ddagger$	-206 ± 10 kcal/mol
$\Delta C_p^\ddagger$	7.8 ± 0.5 kcal/mol
$\Delta ASA$	56,021 Å <sup>2</sup>
$\Delta C_p^\ddagger$	7.8–10.0 kcal/mol

\*Values where the  $\Delta G$  is zero at high temperatures.

†Values where the  $\Delta G$  is zero at low temperatures.

‡Measured from the experimental data.

§Expected values calculated from the literature parameterization of  $\Delta ASA$ . Standard errors were obtained from the fit procedure.



**FIGURE 6** Stability curve of Pfk-2. (Curved continuous line) Free-energy dependence for bimolecular equilibrium calculated from variation of the enthalpies (dotted line) and entropy (dashed line) with the temperature using the thermodynamic parameters of Table 2.

## DISCUSSION

### Cold- and heat-induced unfolding of Pfk-2 resembles the GdnHCl-induced unfolding/dissociation step

As indicated in Fig. 2, the amplitudes of both transitions (cold and heat-induced) are similar in terms of the molar ellipticity, suggesting that the cold- and heat-denatured states of Pfk-2 share some properties in terms of their secondary structure content. Moreover, both transitions are sensitive to the protein concentration, indicating that unfolding occurs concomitantly with dimer dissociation. Sedimentation measurements confirm this fact by demonstrating the presence of a monomer at low temperature with an Rh value of  $4.6 \pm 0.3$  nm as measured by DLS. Taking into account that a fully unfolded chain of 33 kDa is expected to have an Rh value of  $\sim 5.3$  nm (32), we conclude that the cold-induced transition produces largely unfolded subunits. Rh values of several proteins obtained by NMR measurements have shown that cold-induced unfolding can be accompanied by an expansion of the polypeptide chain reaching values close to the fully unfolded polypeptide in GdnHCl (33).

However, cold unfolding of Pfk-2 does not lead to full unfolding. Treatment of the cold-unfolded state of Pfk-2 with GdnHCl generates an additional sigmoidal transition when monitored by CD (Fig. 7 A) whose amplitude was similar to that which describes the I ↔ U equilibrium observed at 20°C. The CD transition induced by GdnHCl at 3°C was also accompanied by a large decrease of the ANS fluorescence without changes to the intrinsic fluorescence of Pfk-2. These characteristics match those described for the expanded intermediate "I" characterized under equilibrium conditions in the presence of GdnHCl at 20°C

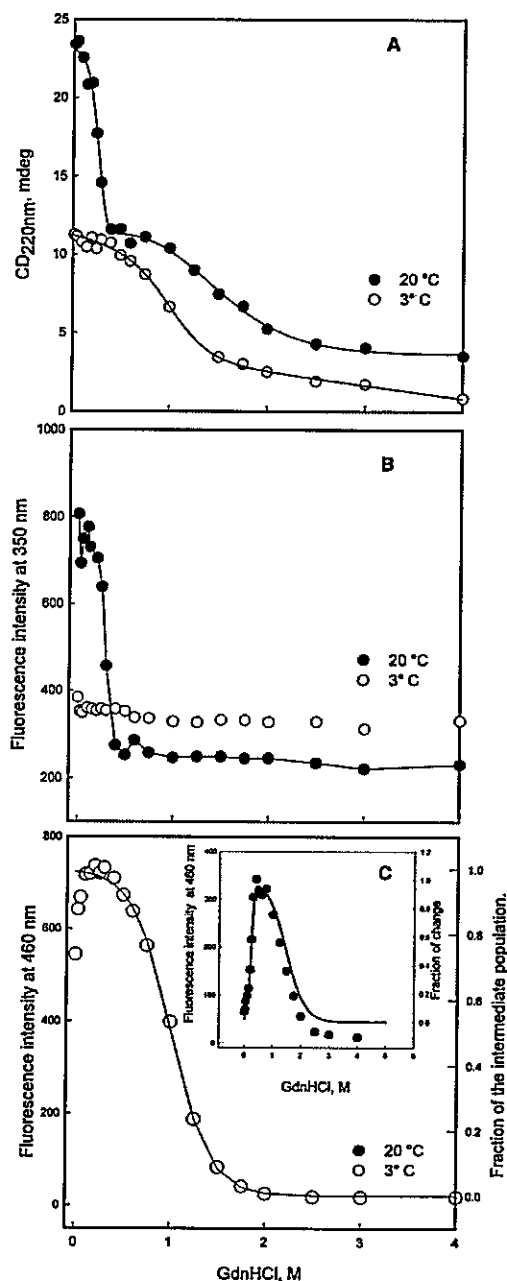


FIGURE 7 Unfolding of Pfk-2 induced by GdnHCl. Protein samples were incubated at 3°C (O) or at 20°C (●) at several GdnHCl concentrations: CD measurements (A), intrinsic fluorescence intensity (B), and fluorescence intensity of the hydrophobic probe ANS (C). (Continuous lines in

TABLE 3 Thermodynamic parameters for the chemical unfolding of Pfk-2 at 20 and 3°C

	20°C		3°C
	$N_2 \leftrightarrow 2I$	$I \leftrightarrow U$	$I \leftrightarrow U$
$\Delta G$	$12.1 \pm 2.1$	$1.7 \pm 1.5$	$2.7 \pm 0.8$
$m$	$18 \pm 7$	$1.3 \pm 0.4$	$2.6 \pm 0.7$

Values of  $\Delta G$  in kcal/mol and  $m$ -values in kcal/mol<sup>2</sup>/K. Standard errors were obtained from the fit procedure.

(4), but there are some differences with respect to the fluorescence obtained upon ANS binding. The amplitude of the ANS fluorescence at 3°C was larger than the amplitude of the  $I \leftrightarrow U$  step measured at 20°C. These changes cannot be attributed to the temperature dependence of the ANS fluorescence because, in absence of protein, the quantum yield and maximum of emission of the probe remained unaltered (data not shown). Hence, because the overall properties of the monomeric state observed at 3°C seem similar to those of the intermediate "I" measured at 20°C, it is proposed that the affinity of the probe for hydrophobic regions of the protein is increased at lower temperatures.

From a thermodynamic point of view, the two-state fit to the thermal unfolding curves of Pfk-2 predicts free energy values of  $\sim 11$  kcal/mol at 20°C, which is similar to that calculated from the dissociation and unfolding step  $N_2 \leftrightarrow 2I$  obtained in the presence of GdnHCl ( $12$  kcal/mol  $\pm 0.6$  (4)). Nevertheless, this analysis should be taken with caution because the unfolding transition induced by heat shows deviations from a two-state mechanism. Although we cannot discard the presence of intermediates upon heat denaturation, it is plausible that side reactions leading to irreversible aggregates could account for the absence of an isosbestic point during the heat transition. Therefore, from a thermodynamic and structural point of view, cold-induced unfolding of Pfk-2 resembles the cooperative dissociation and unfolding represented by the  $N_2 \leftrightarrow 2I$  equilibrium characteristic of the effect of GdnHCl.

#### Why Is Pfk-2 unfolded upon cold treatment?

From the classic thermodynamic point of view, cold denaturation of proteins results as a natural consequence of a large  $\Delta C_p$  value arising from the hydrophobic surfaces exposed to the solvent upon protein unfolding (7,8). This value has been parameterized in terms of the polar and hydrophobic contributions for a particular solvent-exposed

panel A) Fit to a  $N_2 \leftrightarrow 2I \leftrightarrow 2U$  mechanism at 20°C and to a  $I \leftrightarrow U$  mechanism at 3°C. (Continuous lines in panel C) Variation of the intermediate species population predicted at 3°C (main graph) and 20°C (inset) using the thermodynamic values obtained by CD (see Table 3). The protein concentration was 3.6  $\mu$ M. For comparison purposes, the sign of the CD signal was inverted.

protein surface (see Spolar et al. (30), Myers et al. (34), and Whitten et al. (35), and references therein). The Pfk-2 dimer hides a surface area of 56,021 Å<sup>2</sup> upon folding (see Materials and Methods), which corresponds to a  $\Delta C_p$  value between 7.8 and 10 kcal depending on the parameterization used (see Spolar et al. (30), Myers et al. (34), and Whitten et al. (35), and references therein). Parameterized values are in agreement with the observed  $\Delta C_p$  of 7.8 kcal/mol<sup>o</sup>K obtained with the assumption of a two-state mechanism for the thermal unfolding of Pfk-2 (Table 2). Also, there are no unusual thermodynamic values obtained from the cold- and heat-denaturation processes, as the enthalpies calculated at some temperatures (e.g., 110°C, 650 kcal/mol) can be predicted adequately using the parameterization of Freire (110°C, 760 kcal/mol; see Whitten et al. (35) and references therein). Therefore, the simplest explanation for the observed cold unfolding of Pfk-2 is the large size of the cooperative unit, i.e., almost the entire dimeric structure.

The large size of the cooperative unit of Pfk-2 can be rationalized from within the framework of the hierarchic model (see Freire et al. (19) and Whitten et al. (35), and references therein). As pointed out by Freire et al. (19), cold denaturation can be seen as a special kind of perturbation simply because the intrinsic stabilities of different domains or parts of the protein structure and their coupling energy of interaction vary in such a way as a function of temperature. Interaction among their cooperative units is favorable at high temperatures yet unfavorable at low temperatures. Thus, to show unbiased cold unfolding, the unfolding coupling energy among several parts of the Pfk-2 structure should remain larger than their intrinsic stabilities either at low or high temperatures. In the absence of quantitative knowledge about the energies of coupling and intrinsic stability, an unbiased cold-unfolding mechanism can be best explained by a mutual constraint imposed between the  $\beta$ -clasp and the  $\alpha/\beta/\alpha$  domain. Such domain is created by an interrupted chain connectivity of both domains (discontinuous domains) together with the intertwined nature of the  $\beta$ -clasp that stabilizes both subunits (see Fig. 1).

## SUPPORTING MATERIAL

Additional sections and two figures are available at [http://www.biophysj.org/biophysj/supplemental/S0006-3495\(12\)01118-6](http://www.biophysj.org/biophysj/supplemental/S0006-3495(12)01118-6).

We thank Richard C. Garratt for critical reading of the manuscript and Avram Stovic for technical assistance. CD measurements and light-scattering experiments were performed at the Grupo de Cristalografía, Instituto de Física de São Carlos, Universidade de São Paulo, Brazil. Sedimentation velocity experiments were performed at the Biophysics Facility, Department of Chemistry and Biochemistry, University of California, San Diego, CA.

This work was supported by a grant from the Comisión Nacional de Investigación Científica y Tecnológica (CONICYT), Fondo Nacional de Desarrollo Científico y Tecnológico (FONDECYT) No. 1090336, Chile. Christian A. M. Wilson was a recipient of a CONICYT graduate fellowship,

a Beca de Apoyo a Tesis Doctoral award (No. 24090160), and a travel prize from Fermelo Chile S.A. César A. Ramírez-Sarmiento was the recipient of a CONICYT graduate fellowship, a joint Mejoramiento de la Calidad y Equidad de la Educación and Sciences, Technology and Innovation Program for the Americas fellowship No. UCH0717, and was partially supported by a FONDECYT grant.

## REFERENCES

1. Warringer, J., and A. Blomberg. 2006. Evolutionary constraints on yeast protein size. *BMC Evol. Biol.* 6:61.
2. Cabrera, R., J. Babul, and V. Guixé. 2010. Ribokinase family evolution and the role of conserved residues at the active site of the PfkB subfamily representative, Pfk-2 from *Escherichia coli*. *Arch. Biochem. Biophys.* 502:23–30.
3. Cabrera, R., A. L. Ambrosio, ..., J. Babul. 2008. Crystallographic structure of phosphofructokinase-2 from *Escherichia coli* in complex with two ATP molecules. Implications for substrate inhibition. *J. Mol. Biol.* 383:588–602.
4. Baez, M., and J. Babul. 2009. Reversible unfolding of dimeric phosphofructokinase-2 from *Escherichia coli* reveals a dominant role of inter-subunit contacts for stability. *FEBS Lett.* 583:2054–2060.
5. Baez, M., C. A. M. Wilson, and J. Babul. 2011. Folding kinetic pathway of phosphofructokinase-2 from *Escherichia coli*: a homodimeric enzyme with a complex domain organization. *FEBS Lett.* 585:2158–2164.
6. Luque, I., S. A. Leavitt, and E. Freire. 2002. The linkage between protein folding and functional cooperativity: two sides of the same coin? *Annu. Rev. Biophys. Biomol. Struct.* 31:235–256.
7. Privalov, P. L., and S. J. Gill. 1988. Stability of protein structure and hydrophobic interaction. *Adv. Protein Chem.* 39:191–234.
8. Dill, K. A. 1990. Dominant forces in protein folding. *Biochemistry.* 29:7133–7155.
9. Privalov, P. L. 1990. Cold denaturation of proteins. *Crit. Rev. Biochem. Mol. Biol.* 25:281–305.
10. Gursky, O., and D. Atkinson. 1996. High- and low-temperature unfolding of human high-density apolipoprotein A-2. *Protein Sci.* 5:1874–1882.
11. Griko, Y. V., and V. P. Kutysenko. 1994. Differences in the processes of  $\beta$ -lactoglobulin cold and heat denaturations. *Biophys. J.* 67:356–363.
12. Griko, Y. V., P. L. Privalov, ..., S. Y. Venyaminov. 1988. Cold denaturation of staphylococcal nuclease. *Proc. Natl. Acad. Sci. USA.* 85:3343–3347.
13. Griko, Y. V., S. Y. Privalov, and P. L. Venyaminov. 1989. Heat and cold denaturation of phosphoglycerate kinase (interaction of domains). *FEBS Lett.* 244:276–278.
14. Richardson, 3rd, J. M., S. D. Lemaire, ..., G. I. Makhatadze. 2000. Difference in the mechanisms of the cold and heat induced unfolding of thioredoxin H from *Chlamydomonas reinhardtii*: spectroscopic and calorimetric studies. *Biochemistry.* 39:11154–11162.
15. Lopez, C. F., R. K. Darst, and P. J. Rosky. 2008. Mechanistic elements of protein cold denaturation. *J. Phys. Chem. B.* 112:5961–5967.
16. Chen, B. L., and J. A. Schellman. 1989. Low-temperature unfolding of a mutant of phage T4 lysozyme. 1. Equilibrium studies. *Biochemistry.* 28:685–691.
17. Pastore, A., S. R. Martin, ..., P. A. Temussi. 2007. Unbiased cold denaturation: low- and high-temperature unfolding of yeast frataxin under physiological conditions. *J. Am. Chem. Soc.* 129:5374–5375.
18. Kitakuni, E., Y. Kuroda, ..., H. Nakamura. 1994. Thermodynamic characterization of an artificially designed amphiphilic  $\alpha$ -helical peptide containing periodic prolines: observations of high thermal stability and cold denaturation. *Protein Sci.* 3:831–837.
19. Freire, E., K. P. Murphy, ..., P. L. Privalov. 1992. The molecular basis of cooperativity in protein folding. Thermodynamic dissection of

- interdomain interactions in phosphoglycerate kinase. *Biochemistry*. 31:250-256.
20. Mizuguchi, M., D. Hashimoto, ..., K. Nitta. 2000. Cold denaturation of  $\alpha$ -lactalbumin. *Proteins*. 38:407-413.
  21. Kunugi, S., and N. Tanaka. 2002. Cold denaturation of proteins under high pressure. *Biochim. Biophys. Acta*. 1595:329-344.
  22. Azuaga, A. I., M. L. Galisteo, ..., P. L. Mateo. 1992. Heat and cold denaturation of  $\beta$ -lactoglobulin B. *FEBS Lett.* 309:258-260.
  23. Babul, J. 1978. Phosphofructokinases from *Escherichia coli*. Purification and characterization of the nonallosteric isozyme. *J. Biol. Chem.* 253:4350-4355.
  24. Bradford, M. M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72:248-254.
  25. Schuck, P. 2000. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophys. J.* 78:1606-1619.
  26. Bowie, J. U., and R. T. Sauer. 1989. Equilibrium dissociation and unfolding of the Arc repressor dimer. *Biochemistry*. 28:7139-7143.
  27. Reference deleted in proof.
  28. Baez, M., R. Cabrera, ..., J. Babul. 2007. Unfolding pathway of the dimeric and tetrameric forms of phosphofructokinase-2 from *Escherichia coli*. *Biochemistry*. 46:6141-6148.
  29. Thompson, K. S., C. R. Vinson, and E. Freire. 1993. Thermodynamic characterization of the structural stability of the coiled-coil region of the bZIP transcription factor GCN4. *Biochemistry*. 32:5491-5496.
  30. Spolar, R. S., J. R. Livingstone, and M. T. Record, Jr. 1992. Use of liquid hydrocarbon and amide transfer data to estimate contributions to thermodynamic functions of protein folding from the removal of nonpolar and polar surface from water. *Biochemistry*. 31:3947-3955.
  31. Daniel, E., and G. Weber. 1966. Cooperative effects in binding by bovine serum albumin. I. The binding of 1-anilino-8-naphthalenesulfonate. Fluorimetric titrations. *Biochemistry*. 5:1893-1900.
  32. Uversky, V. N. 1993. Use of fast protein size-exclusion liquid chromatography to study the unfolding of proteins which denature through the molten globule. *Biochemistry*. 32:13288-13298.
  33. Li, Y., B. Shan, and D. P. Raleigh. 2007. The cold denatured state is compact but expands at low temperatures: hydrodynamic properties of the cold denatured state of the C-terminal domain of L9. *J. Mol. Biol.* 368:256-262.
  34. Myers, J. K., C. N. Pace, and J. M. Scholtz. 1995. Denaturant  $m$  values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Sci.* 4:2138-2148.
  35. Whitten, S. T., A. J. Kurtz, ..., V. J. Hilser. 2006. Revealing the nature of the native state ensemble through cold denaturation. *Biochemistry*. 45:10163-10174.

## Observation of Solvent Penetration during Cold Denaturation of *E. coli* Phosphofructokinase-2

César A. Ramírez-Sarmiento,<sup>†</sup> Mauricio Baez,<sup>‡</sup> Christian A. M. Wilson,<sup>†</sup> Jorge Babul,<sup>†</sup> Elizabeth A. Komives,<sup>§\*</sup> and Victoria Guixé<sup>†\*</sup>

<sup>†</sup>Departamento de Biología, Facultad de Ciencias, Universidad de Chile, Santiago, Chile; <sup>‡</sup>Departamento de Bioquímica y Biología Molecular, Facultad de Ciencias Químicas y Farmacéuticas, Universidad de Chile, Santiago, Chile; and <sup>§</sup>Department of Chemistry and Biochemistry, University of California San Diego, La Jolla, California

**ABSTRACT** Phosphofructokinase-2 is a dimeric enzyme that undergoes cold denaturation following a highly cooperative  $N_2 \rightleftharpoons 2I$  mechanism with dimer dissociation and formation of an expanded monomeric intermediate. Here, we use intrinsic fluorescence of a tryptophan located at the dimer interface to show that dimer dissociation occurs slowly, over several hours. We then use hydrogen-deuterium exchange mass spectrometry experiments, performed by taking time points over the cold denaturation process, to measure amide exchange throughout the protein during approach to the cold denatured state. As expected, a peptide corresponding to the dimer interface became more solvent exposed over time at 3°C; unexpectedly, amide exchange increased throughout the protein over time at 3°C. The rate of increase in amide exchange over time at 3°C was the same for each region and equaled the rate of dimer dissociation measured by tryptophan fluorescence, suggesting that dimer dissociation and formation of the cold denatured intermediate occur without appreciable buildup of folded monomer. The observation that throughout the protein amide exchange increases as phosphofructokinase-2 cold denatures provides experimental evidence for theoretical predictions that cold denaturation primarily occurs by solvent penetration into the hydrophobic core of proteins in a sequence-independent manner.

### INTRODUCTION

Protein stability can be altered via perturbation of several properties that are required to ensure a properly folded state. Changes in pH, addition of chaotropic agents, high pressure, and increase in temperature have been successfully applied to drive proteins through their global unfolding transitions (1). Interestingly, it has been shown that the Gibbs free energy difference between the native and denatured state for a two-state system reaches zero at two different temperatures, thus suggesting that thermal unfolding transitions of proteins can be induced both on cooling and heating (2). As thermal stability of proteins is dominated by the concave free energy dependence with temperature, it has been predicted that proteins with large and positive  $\Delta C_p$  and small  $\Delta H$  values would permit the observation of cold-induced unfolding/denaturation at temperatures above the freezing point of water (3,4).

Cold denaturation phenomena in proteins have been studied by several techniques, including spectroscopy (5), calorimetry (6), NMR (7), and molecular dynamics (8). However, with the exception of NMR experiments, these techniques only account for global properties and do not allow measurement of local changes during cold-induced denaturation. NMR is still limited by the size of the protein and requirement for specific labeling. Therefore, a better

description of the cold denaturation process may be obtained by analysis of local changes in the structure of proteins using amide hydrogen-deuterium (H/D) exchange. H/D exchange followed by mass spectrometry (HXMS) is a powerful tool for studying protein-protein interactions (9), protein dynamics upon binding, and intrinsically disordered regions in proteins (10,11). The extent of exchange probes the solvent accessibility of amide protons, whereas the approach in which the exchange reaction is quenched, the protein is digested with pepsin, and the resulting peptides are analyzed by mass spectrometry, allows sampling of these changes over local regions of the protein.

*Escherichia coli* phosphofructokinase-2 (Pfk-2) is a homodimer of 66 kDa and one of the largest proteins that shows thermal unfolding transitions upon cooling and heating (12). Reversible thermal unfolding experiments under equilibrium conditions have shown that the cold-induced transition of Pfk-2 proceeds above the freezing point of water without requiring additional perturbations such as chaotropic agents (13,14), high pressure (15), extreme pH (16), or point mutations (17,18). Moreover, it has been shown that cold denaturation of Pfk-2 follows a two-state mechanism to a populated intermediate state ( $N_2 \rightleftharpoons 2I$ ) and it has been suggested that this process is highly cooperative, occurring concomitantly with dimer dissociation and leading to an expanded monomeric conformation that shows striking similarities to the guanidine HCl-induced monomeric intermediate (12).

In this work, we first used analytical ultracentrifugation to determine the amount of dimeric and monomeric species

Submitted December 19, 2012, and accepted for publication April 12, 2013.

\*Correspondence: [ekomives@ucsd.edu](mailto:ekomives@ucsd.edu) or [vguixe@uchile.cl](mailto:vguixe@uchile.cl)

Christian A. M. Wilson's present address is QB3 Institute, University of California at Berkeley, Berkeley, CA 94720.

Editor: Patricia Clark.

© 2013 by the Biophysical Society  
0006-3495/13/05/2254/10 \$2.00

<http://dx.doi.org/10.1016/j.bpj.2013.04.024>



under equilibrium conditions spanning temperatures between 2 and 25°C, followed by intrinsic fluorescence measurements to study the kinetics of the unbiased cold denaturation of Pfk-2. Finally, we probed the amide exchange of the protein along the cold denaturation process with HXMS. We find that cold denaturation of Pfk-2 is a slow process that occurs without buildup of folded monomer. The amide exchange results reveal local increases in amide exchange that occur throughout Pfk-2 during cold challenge, leading to the monomeric intermediate state. We show that amide exchange increases throughout the protein as cold denaturation proceeds. Because the cold denaturation process is slow, the rate at which amide exchange increases at locations throughout the protein core could be ascertained by measuring HXMS in samples removed from the denaturation reaction at different times. Remarkably, the rate of increase in amide exchange was the same for all regions of the protein, and also matched the rate of monomer formation measured by tryptophan fluorescence. The results provide experimental validation of recently proposed theories of solvent penetration throughout the folded core of the protein as the mechanism of cold denaturation (19).

## EXPERIMENTAL PROCEDURES

### Purification and storage of Pfk-2

Pfk-2 was purified and stored as described by Babul (20). Before performing kinetic denaturation experiments, the storage buffer was changed to standard buffer (50 mM Tris pH 8.2, 5 mM MgCl<sub>2</sub>, and 2 mM DTT) by centrifugation using a Micro Bio-Spin 6 column (BioRad Laboratories Inc., Hercules, CA). The enzyme was concentrated using a VivaSpin 3 concentrator (Sartorius Stedim Biotech S.A., Goettingen, Germany) when required. Protein concentration was determined by Bradford assay (21) using bovine serum albumin as standard and is expressed in terms of monomer concentration.

### Sedimentation analysis by analytical ultracentrifugation at different temperatures

To ascertain the temperature at which Pfk-2 was completely monomeric, sedimentation velocity experiments were conducted at different temperatures. A Beckman Optima XL-I analytical ultracentrifuge (Beckman Coulter, Brea, CA) equipped with an An60Ti rotor was used for the experiments. Pfk-2 was exchanged from its storage buffer to 50 mM Tris pH 8.2, 5 mM MgCl<sub>2</sub>, and 2 mM TCEP before the experiment. Pfk-2 samples (~38 μM) were incubated for 24 h at each desired temperature (from 2 to 25°C) to ensure equilibrium conditions and then placed in the analytical cells for centrifugation at a rotor speed of 42,000 rpm until complete sedimentation was achieved. The rotor, centrifuge, and analytical cells were equilibrated at the desired temperature before loading the samples. The moving boundary was monitored by absorbance at 280 nm. The resulting data were analyzed using different models. First, we used SEDFIT and the two-dimensional model of distribution of sedimentation coefficient and frictional ratio  $c(S, f, \rho)$  for analysis of the monomeric and dimeric species free from assumptions about the frictional ratio (22). This analysis was only performed at 4 and 25°C, where single species rather than monomer/dimer mixtures were expected. Distributions were obtained with equidistant  $f$ -grid values between 0.8 and 2.2 in 0.1 steps, and linear  $S$ -grid values ranging from 0.5 to 6.5 in 0.1 steps. Data are presented in terms of  $S_{20,w}$  and  $R_h$ , where  $S_{20,w}$

corresponds to the sedimentation coefficient normalized to 20°C and water, and  $R_h$  is the Stokes radius. Two other models from SEDPHAT were then used: i), a simple global analysis of species to obtain the partial concentration of each monomer and dimer at each temperature (23), and ii), a self-association monomer-dimer model to compare the kinetic rates of dissociation estimated from our different measurements (24).

### Rate of change of intrinsic tryptophan fluorescence over time in the cold

Kinetic denaturation of Pfk-2 was measured by intrinsic fluorescence, using its single tryptophan (W88) located on the interface between subunits as probe. Measurements were done in a Shimadzu PC-5031 spectrofluorometer (Shimadzu Europa GmbH, Duisburg, Germany). The change in intrinsic fluorescence was measured over time at 3°C. The protein was excited at a wavelength of 295 nm and the change was followed at 346 nm. The protein concentration used was 0.3 μM and the decay in signal was adjusted to an exponential decay to a minimum corresponding to the signal of the protein after 24 h at 3°C.

$$F = F_0 e^{-kt} + F_f \quad (1)$$

where  $F$  is the fluorescence intensity as a function of time  $t$ ,  $F_0$  is the fluorescence at  $t = 0$  h,  $F_f$  is the fluorescence intensity recorded after 24 h at 3°C, and  $k$  is the rate of decay of the fluorescence signal.

### Native state backbone amide exchange at 25°C

Pfk-2 backbone amide exchange was performed as previously described (10). The enzyme was incubated at room temperature for 2 h in standard buffer before starting the experiments. The exchange reaction was initiated when 80 μM of Pfk-2 was diluted 10-fold into D<sub>2</sub>O. The protein was allowed to exchange at room temperature for 0–10 min. Exchange was quenched by diluting the sample 10-fold into 0.1% trifluoroacetic acid (TFA) in H<sub>2</sub>O. Quenched protein was digested with pepsin as described and small aliquots of the digested sample were rapidly frozen in liquid N<sub>2</sub> and stored at -80°C.

### Backbone amide exchange during cold denaturation

Pfk-2 was cold denatured by incubating the enzyme at 3°C. At several time points during the 12 h incubation, samples were removed and subjected to exchange in D<sub>2</sub>O as described previously, and the exchange rate was measured between 0 and 40 min. The reaction was quenched into 0.1% TFA in H<sub>2</sub>O and samples were digested and stored as described. To compare the denatured intermediate formed by cold denaturation with that formed at low guanidine concentrations, experiments were also performed on Pfk-2 diluted in 0.85 M guanidine HCl for 48 h, as in Baez and Babul (25). These experiments were performed at room temperature and exchange was monitored for 0–10 min, quenched, digested, and analyzed as described previously.

### Digestion of Pfk-2 and Identification of peptic peptides

After completion of the exchange reaction and quenching of amide exchange, Pfk-2 was digested with immobilized pepsin (Pierce Scientific, Rockford, IL) as previously described (10). Immobilized pepsin slurry (25 μL) was activated by washing twice with 0.1% TFA at pH 2.2. The protein was diluted in 0.1% TFA at pH 2.2 before incubation with immobilized pepsin for 5 min. Peptides were identified by MS/MS using a 4800 Plus

MALDI-TOF-TOF (AB SCIEX, Framingham, MA) and spectral analysis using MASCOT (Matrix Science, Boston, MA) and Protein Prospector.

### Mass spectrometry

Matrix-assisted laser desorption/ionization time of flight (MALDI-TOF) mass spectra were acquired on a Voyager DE-STR instrument (Applied Biosystems, Carlsbad, CA) as previously described (26). The matrix used was 4.5 mg/mL  $\alpha$ -Cyano-4-hydroxycinnamic acid (Sigma-Aldrich, St. Louis, MO) in a solution of 1:1:1 acetonitrile/ethanol/0.1% TFA. The pH of the matrix was adjusted to pH 2.2 using 2% TFA. The matrix solution and MALDI-TOF target plates were equilibrated at 4°C before the experiment as previously described (26).

MALDI-TOF mass spectra were analyzed as previously described (9,10), correcting for both side-chain contributions due to residual deuterium (4.5%) and back-exchange loss (50%). Plots were fit to a one-parameter exponential by using KaleidaGraph version 4.0 (Synergy Software, Reading, PA).

To calculate the kinetic denaturation rates ( $k$ ), the maximum extent of exchange for each peptide as a function of time of exposure at cold temperature was fitted to an exponential rise to maximum:

$$E(t) = E_0 + N(1 - e^{-kt}) \quad (2)$$

where  $E(t)$  is the maximum extent of exchange upon time,  $E_0$  is the maximum extent of exchange at  $t = 0$  h (i.e., extent of exchange for native state Pfk-2),  $N$  is the amplitude of the change in deuterons incorporated (i.e., the amount of backbone amides that begin to exchange due to changes in solvent accessibility), and  $t$  is the time of incubation at cold temperature.

Some samples were also analyzed on a Synapt G2 mass spectrometer (Waters, Milford, MA) following protocols described by Engen and co-workers (27) and these data were analyzed by PLGS and DynamX software packages.

## RESULTS

### Sedimentation analysis of Pfk-2 at different temperatures

To determine the equilibrium distribution of monomer and dimer species at different temperatures between 25 and 4°C, samples were equilibrated for 24 h and then analyzed by sedimentation velocity (Fig. 1). The protein is mainly a dimer at 25°C. Both monomers and dimers are present at temperatures between 8 and 15°C. At temperatures below 6°C almost all of the protein exists as a monomer. Furthermore, a Stokes radius larger than expected for a compact monomer is observed at 4°C (Fig. 1 A), in agreement with previous dynamic light scattering (DLS) experiments (12). This experiment established that cold-dependent dissociation to a monomer occurs (Fig. 1 B) and that to reach the endpoint of this transition, it is necessary to incubate the protein at a temperature of 4°C or lower.

### Rate of Pfk-2 dissociation over time at 3°C

Typically, cold denaturation experiments are performed with the concomitant addition of chaotropic agents to facilitate the denaturation process. Because we knew that Pfk-2

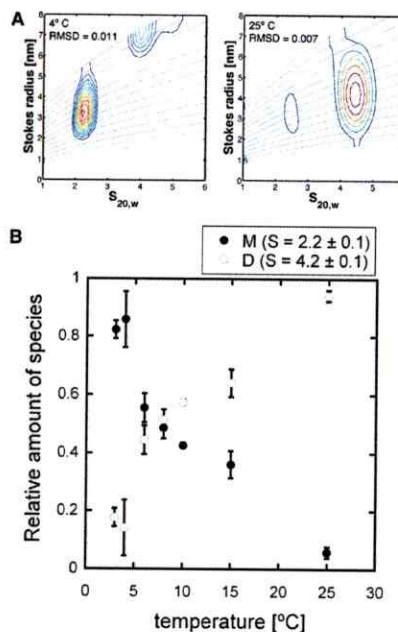


FIGURE 1 (A) Contour plots of the two-dimensional distribution in terms of  $S_{20,w}$  and Stokes radius of Pfk-2 monomers and dimers at 4 and 25°C. Data were obtained by sedimentation velocity experiments and subsequent analysis using a continuous  $c(S_{20,w}, R_h)$  model, here represented as  $c(S_{20,w}, R_h)$ . The magnitude of  $c(S_{20,w}, R_h)$  is indicated by the color of the contours in equidistant intervals of  $c$  from blue (low, outer boundary of each contour) to red (high, center of each contour). The dotted lines indicate lines of constant  $f_r$ . The RMSD resulting from fitting the data is indicated. (B) Plot of the percent of monomer (solid circles) and dimer (open circles) at each temperature measured as calculated from the global analysis of species using SEDPHAT. The sedimentation coefficient values indicated were obtained from fitting the data and are similar between the different models used.

denatures completely to the monomer at 3°C without the addition of chaotropic agents, we measured the rate of dissociation at 3°C over time taking advantage of the single tryptophan residue (W88) located on the dimer interface (Fig. 2). The rate of decay of the fluorescence signal over time of incubation at 3°C was  $1.26 \times 10^{-4} \text{ s}^{-1}$ , which corresponds to a relatively slow conversion to monomer from dimer, on the order of hours. Another estimation of the dissociation rate was obtained from sedimentation velocity experiments using a monomer-dimer self-association model (24), which gave  $k_{off} = 0.8 \times 10^{-4} \text{ s}^{-1} \pm 0.3 \times 10^{-4} \text{ s}^{-1}$ . The fact that conversion to the cold-denatured intermediate state is slow, but can be observed without the addition of chaotropic agents, prompted us to carry out amide exchange experiments over the cold denaturation process to investigate how the Pfk-2 molecules change as cold denaturation occurs.



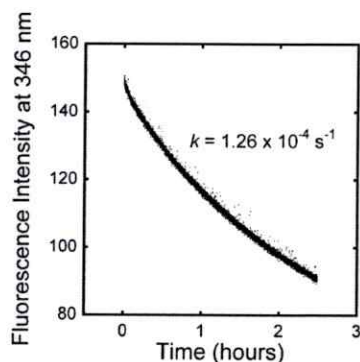


FIGURE 2 Intrinsic fluorescence measurement of Pfk-2 during cold denaturation at 3°C. The kinetic constant for dissociation ( $k$ ) was calculated by fitting the data to an exponential decay curve.

#### Native state amide exchange of Pfk-2 at 25°C

To discover whether there were temperature-dependent differences in the amide exchange of Pfk-2 at 25 and 3°C, we performed HXMS experiments. As a control, Pfk-2 was allowed to exchange under native conditions for 10 min at 25°C. The samples were quenched by dilution into ice cold 0.1% TFA, digested for 5 min with immobilized pepsin, and analyzed on the MALDI-TOF mass spectrometer. Several peptides could be analyzed quantitatively, and these represented ~27% coverage of protein sequence (Table 1, Fig. S1 in the Supporting Material). These peptides are distributed throughout the protein structure, including a single peptide spanning residues 104–113 (MH+ 1124.592) that corresponds to one of the  $\beta$ -strands of the interface, thus working as a reporter of the interface between subunits. At 25°C, Pfk-2 had little solvent accessibility (Fig. S2). Experiments were not carried out to achieve complete exchange, but rather to measure the more rapidly exchanging amides. Most regions of Pfk-2, including the interfacial peptide that reports on association between monomers, exchanged 23% on average after 5 min in D<sub>2</sub>O. Residues 1–7 of the protein (MH+ = 895.507) exchanged 43% (Fig. S2), suggesting that the N-terminus of Pfk-2 is more solvent accessible than the well-packed protein core.

#### Native state amide exchange of Pfk-2 at 3°C

As a control for the planned kinetic analysis of amide exchange during cold denaturation, we performed an experiment in which Pfk-2 was diluted into buffer preequilibrated at 3°C. This important control accounts for any temperature effects of the buffer on the amide exchange rates and also serves as a 0 time point for measuring the rate of change of the amide exchange over the very slow cold denaturation process (occurring over 12 h). According to Bai et al., the exchange is expected to be fourfold slower at 3°C compared to 25°C, so the HXMS time points were measured over 40 min instead of over 10 min (28).

To compare the HXMS of Pfk-2 at 25°C (Fig. S2) and immediately after dilution to 3°C (Fig. S3), we compared the total and backbone solvent accessible surface area (SASA) from Pfk-2 structure (PDB ID 3CQD) to the extent of amide exchange at 25°C (10 min time point) and to the extent of amide exchange of the sample immediately diluted to 3°C (40 min time point). In this manner, we wanted to confirm that the conformation of Pfk-2 at the beginning of the cold denaturation was the same as the conformation at room temperature.

The extent of amide exchange under both conditions was highly correlated with the backbone SASA for all covered regions of Pfk-2 as shown in Fig. 3 for the enzyme at 25°C or immediately diluted into buffer at 3°C (correlation coefficient 0.92 and 0.85, respectively). The results suggest that there is no significant change in Pfk-2 conformation after initial exposure to cold temperature consistent with the very slow rate of conversion to the cold-denatured intermediate observed by tryptophan fluorescence.

#### Cold-induced denaturation and dissociation of Pfk-2

Because the kinetics of cold-induced denaturation of Pfk-2 was so slow, we thought it would be possible to monitor the change in amide exchange at various times during the denaturation process. The experimental setup involved incubating Pfk-2 at 3°C for 12 h, and at intervals during this cold incubation, samples were removed and analyzed for the extent of amide exchange by diluting into deuterated buffer

TABLE 1 List of peptides quantified through MALDI-TOF mass spectrometry

Region of Pfk-2	Peptide mass (m/z)	No. amides	Relative exchange 4°C/25°C	Sequence
1–7	895.507	7	>	MVRIYTL
276–285	934.511	10	>	VRFGVAAGSA
286–294	973.543	9	>	ATLNQGTRL
275–285	1065.551	11	>	MVRFVVAAGSA
104–113	1124.592	9	>	YRFVMPGAAL
118–126	1161.626	9	>	FRQLEEQVL
154–164	1174.636	11	>	SAAKQGIKRCI
44–69	2445.300	25	>	VARAIAHLGGSATAIFPAGGATGEHL



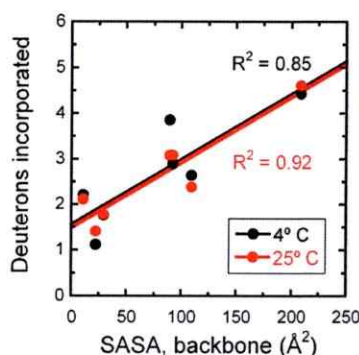


FIGURE 3 Correlation between backbone SASA calculated from Pfk-2 crystal structure (PDB ID 3CQD) and backbone amide exchange reactions at room temperature (red) and immediately after the Pfk-2 was exposed to cold temperature (black).

and monitoring deuterium incorporation into each region of the protein over 40 min as described previously. This experiment follows the change in amide exchange for each peptide representing different regions of the protein and the dimer interface over the entire time course of cold denaturation (Fig. 4).

As shown in Fig. 2, Pfk-2 denatures to a monomeric intermediate over hours of incubation at 3°C. Therefore, we expected (and observed) that the peptide corresponding to

residues 104–113 (MH+ 1124.592), representing one of the  $\beta$ -strands that forms the dimer interface, would exchange more as the protein denatured to the monomeric state. To obtain these data, we plotted the maximum exchange (i.e., the extent of exchange at 40 min) observed for samples removed after 2, 4, 8, and 12 h of incubation of Pfk-2 at 3°C (Figs. S4–S7). The extent of deuteration for these samples represents the change in amide exchange of each region of Pfk-2 during the cold denaturation process. It was not surprising that the dimer interface becomes more solvent accessible as the protein becomes monomeric during the denaturation process, however, the results from the peptides corresponding to other regions of Pfk-2 were very much unexpected. Remarkably, all of the regions of Pfk-2 corresponding to peptides we could analyze by MALDI-TOF mass spectrometry showed increasing amide exchange when Pfk-2 was incubated at 3°C over several hours. The extent of exchange reached a plateau at <100% backbone amide exchange and longer times of deuteration were not explored (11,29). The plateauing amide exchange is consistent with a partially folded state in agreement with previous data showing that Pfk-2 reaches an intermediate state and not a completely unfolded state upon cold denaturation (12). Moreover, the extent of exchange of the cold-induced species was less than the exchange observed for the chemically induced intermediate at room temperature in 0.85 M guanidine HCl (Fig. 5). The increase in solvent accessibility over the cold denaturation at 3°C could be fitted to a single exponential rate

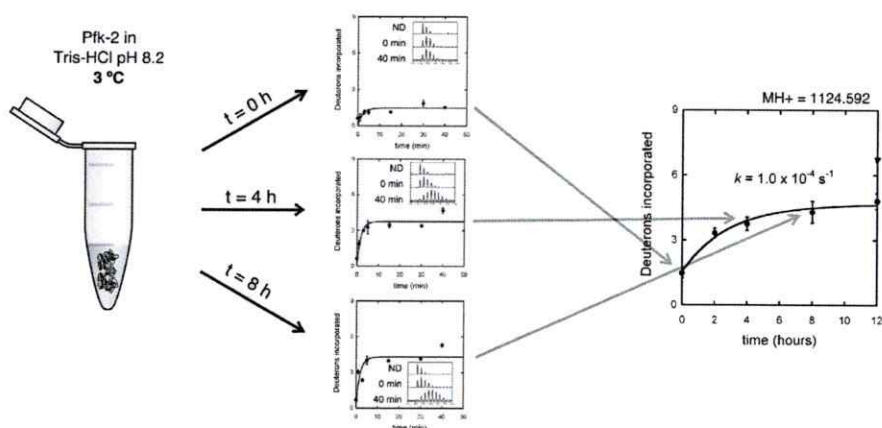


FIGURE 4 Schematic of the experimental approach to monitoring solvent penetration over time during cold denaturation. The protein was equilibrated to 3°C, and samples were removed at times 0, 2, 4, 8, and 12 h. The sampling was based on the slow rate of conversion to the cold-denatured intermediate measured by fluorescence (Fig. 2). After the sample was removed, it was diluted into deuterated buffer and allowed to exchange over 40 min during which time samples of this reaction were removed and the extent of exchange was measured for each time point. The plots of exchange over 40 min immediately after equilibration at 3°C (0 h) are shown in Fig. S3; the plots of exchange after equilibration at 3°C for 2 h are shown in Fig. S4; the plots of exchange after equilibration at 3°C for 4 h are shown in Fig. S5; the plots of exchange after equilibration at 3°C for 8 h are shown in Fig. S6; and the plots of exchange after equilibration at 3°C for 12 h are shown in Fig. S7. From these plots, the extent of deuteration (amount of deuteration at 40 min) was determined and plotted against the time at 3°C in hours. These are the data that are presented in Fig. 5.

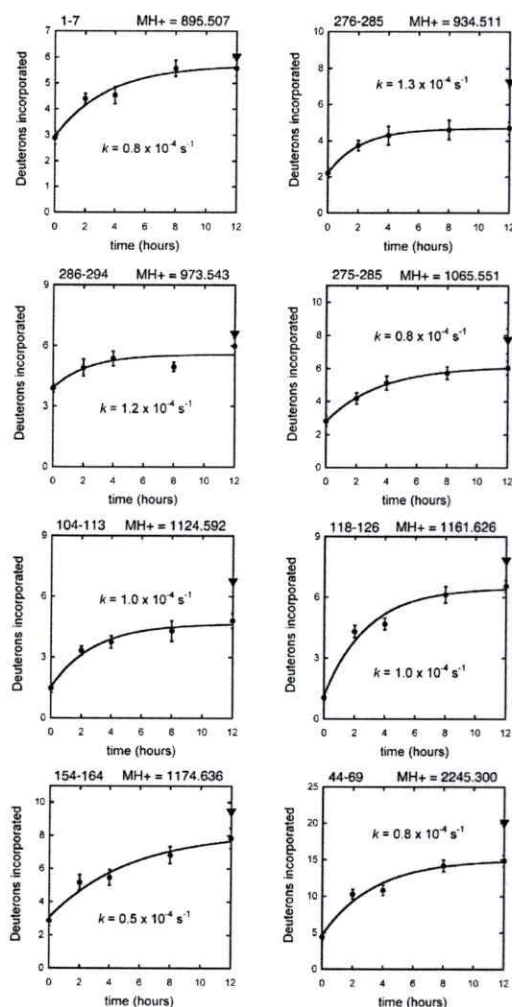


FIGURE 5 Cold denaturation of Pfk-2 followed by backbone amide exchange. Each plot represents the change in deuterons incorporated after 40 min for each peptide as a function of time during exposure to cold temperature. The y axis represents the number of exchangeable amides in the particular peptide. The kinetic constant ( $k$ ) was calculated by fitting the data to an exponential curve. Triangles indicate the deuterons incorporated after 10 min of exchange for Pfk-2 incubated in 0.85 M guanidine HCl. The results suggest that the cold-induced monomeric intermediate is similar to the intermediate induced by guanidine HCl however it may be slightly more compact.

curve, and the rates for each different region of Pfk-2 were very similar for the different regions,  $1.0 \times 10^{-4} \pm 0.25 \times 10^{-4} \text{ s}^{-1}$  (Fig. 5). The peptide that covered the dimer interface also increased solvent accessibility at the same rate,

indicating that dissociation and denaturation are concurrent processes during cold denaturation of the enzyme. The rate of increase in amide exchange during cold denaturation was very similar to the rate of cold denaturation estimated by measuring the decay of the fluorescence signal from the single tryptophan (W88) located on the interface between both subunits (Fig. 2) reinforcing the idea that the dimer dissociation and increased amide exchange throughout Pfk-2, which both occur upon incubation of Pfk-2 for several hours at 3°C, likely represent the same denaturation process. In keeping with the cooperative nature of the denaturation process, some of the mass envelopes (cf. Figs. S5 and S6) eventually appeared bimodal indicating an EX1-type of exchange process that would be expected if the dimer (less exchanging) converted to the expanded monomer (more exchanging) without formation of other intermediates.

The peptides that could be analyzed from the MALDI-TOF HXMS were distributed throughout the protein, but only 27% of the protein sequence was covered. The peptides we analyzed all increased in the extent of deuteration over time in the cold suggesting that the protein globally expanded (Fig. S8). To strengthen this conclusion, additional HXMS data were collected on a Synapt G2 HDX instrument. Due to limitations of instrument time, only one time point of cold denaturation could be studied. The protein was incubated at either 25 or 4°C for 6 h, and amide exchange was measured for time points of 0, 2, 5, and 10 min after incubation. The peptides that could be analyzed quantitatively covered 100% of the protein sequence (Table 2). All regions of Pfk-2 showed increased deuteration after incubation in the cold, recapitulating the MALDI-TOF data, except for two peptides that covered the active site catalytic residues, which showed identical exchange at both temperatures. Thus, these additional data suggest that the entire protein except the catalytic center expands during cold denaturation.

## DISCUSSION

The mechanism of cold denaturation remains controversial. Some reports suggest the process is noncooperative (30), whereas other reports suggest that at least for Yfh1 and the C-terminal domain of L9, it is highly cooperative (7,31). Questions also remain as to whether the cold denatured state is compact, and whether it contains only native structure or also nonnative structure (32,33). These discrepancies arise from several factors: i), commonly, proteins cold denature at temperatures below the freezing point of water, a problem that can be circumvented by an additional perturbation but with the cost of moving away from physiological conditions; ii), the unfolding rate strongly depends on temperature, being several times slower at temperatures close to 0°C and thus making the cold-denatured state kinetically inaccessible (34). Recent results from DSC have

**TABLE 2** List of peptides identified through MS<sup>n</sup> mass spectrometry and relative exchange at 4 and 25°C

Region of Pfk-2	Peptide mass (m/z)	No. amides	Relative exchange		Sequence
			4°C/25°C	Ratio avg. exchange 25°C/4°C	
1-13	1477.844	9	>	0.87	MVRIYTLAPSL
2-13	1346.804	8	>	0.88	VRIYTLAPSL
14-69	5514.822	50	>	0.89	DSATITPQIYPEGKLRCTAPVFEPGGGINVARIAIHL GGSATAIFPAGGATGEHL
33-69	3469.797	34	>	0.89	PVFEPGGGINVARIAIHLGGSATAIFPAGGATGEHL
44-69	2445.299	24	>	0.78	VARIAIHLGGSATAIFPAGGATGEHL
44-72	2744.484	27	>	0.87	VARIAIHLGGSATAIFPAGGATGEHLVSL
73-84	1256.636	10	>	0.88	LADENVPVATVE
73-103	3442.698	29	>	0.89	LADENVPVATVEAKDWTRQNLHVHVEASGEQ
73-113	4548.272	38	>	0.88	LADENVPVATVEAKDWTRQNLHVHVEASGEQYRFVMPGAAL
104-113	1124.592	8	>	0.92	YRFVMPGAAL
104-117	1611.747	12	>	0.95	YRFVMPGAALNEDE
118-126	1161.626	8	>	0.59	FRQLEEQVL
119-126	1014.557	7	>	0.57	RQLEEQVL
127-134	831.446	6	>	0.9	EISGAIL
134-163	3147.828	28		0.94	LVISGSLPPGVKLEKLTQLISAAQKQGIRC
135-163	3034.744	27		0.94	VISGSLPPGVKLEKLTQLISAAQKQGIRC
164-175	1119.552	11	=	1.0	IVDSSGEALSAA
164-176	1232.636	12	=	1.05	IVDSSGEALSAAAL
177-183	729.414	6	=	1.0	AIGNIEL
183-238	5968.203	52	>	0.95	LVKPNQKELSAVNRELTQDDVRKAAQEVNSGKAKRV VVS LGPQQALGVDSENC
239-260	2153.127	18	=	1.06	IQVVPPPVKSQSTVAGDSMVG
239-263	2456.252	21	=	1.0	IQVVPPPVKSQSTVAGDSMVGAMT
239-264	2569.337	22	=	0.96	IQVVPPPVKSQSTVAGDSMVGAMTL

shown that the unfolding rates are, in fact, highly important to ensure an adequate definition of the energetic and structural features of the unfolding process. The slow unfolding rate at low temperatures due to a high folding/unfolding free energy barrier can lead to kinetic trapping in the initial distribution of states and therefore to misleading interpretations about the cooperativity of the cold denaturation process (34). Pfk-2 forms a cold-denatured intermediate that is monomeric and somewhat expanded at temperatures below 4°C. Because no nonphysiological perturbants are required, and because the denaturation process occurs over several hours (Figs. 2 and 5), Pfk-2 provides an excellent model system in which to study the cold denaturation process.

#### Backbone amide exchange and mass spectrometry to probe cold-induced protein denaturation

Backbone amide exchange measured by mass spectrometry (35) has been used to describe protein dynamics (26,36), protein-protein interactions (9,37), and intrinsically disordered regions (11,29). Here, we apply this technique to discover how amide exchange changes over the course of the slow cold denaturation of Pfk-2 that occurs without requiring additional perturbation (12). Although HXMS is a mid-resolution technique, it is not limited by the molecular weight of a particular protein and, in principle, it can be applied to study proteins with biased cold-induced denatur-

ation or proteins that dissociate upon exposure to cold temperatures (38,39). Therefore, HXMS constitutes a powerful technique to describe changes in protein association and amide exchange upon denaturation. The combination of global probes (such as intrinsic fluorescence) with HXMS to describe local structure changes constitutes a powerful combined approach to analyze changes in protein association and amide exchange upon denaturation.

#### Rate of monomer formation during cold-induced protein denaturation

In previous work, the global probes of intrinsic tryptophan fluorescence and circular dichroism spectroscopy revealed the formation of an intermediate with residual secondary structure during kinetic unfolding of Pfk-2 in guanidine HCl (40). Incubation of Pfk-2 at 3°C resulted in the formation of a cold denatured species, which according to intrinsic tryptophan fluorescence, DLS, and circular dichroism resembled the guanidine HCl-induced unfolding intermediate (12). DLS revealed that the cold-denatured species was monomeric. Here, we confirmed by sedimentation velocity experiments that equilibration of Pfk-2 at 3°C results in complete formation of a monomeric species. The rate of formation of this cold-induced monomeric species of Pfk-2 as indicated by the global probe of intrinsic tryptophan fluorescence change, was slow;  $k = 1.26 \times 10^{-4} \pm 0.04 \times 10^{-4} \text{ s}^{-1}$  (Fig. 2). Another estimation of the dissociation rate was obtained from sedimentation velocity experiments using a

monomer-dimer self-association model, which was similar to the rate calculated by intrinsic fluorescence ( $k_{off} = 0.8 \times 10^{-4} \text{ s}^{-1} \pm 0.3 \times 10^{-4} \text{ s}^{-1}$ ) (24). We were able to monitor the change in amide exchange over time of incubation of Pfk-2 at 3°C for a peptide representing the dimer interface, and the rate of increasing amide exchange was  $1.0 \times 10^{-4} \pm 0.26 \times 10^{-4} \text{ s}^{-1}$ . Given that this peptide was at the dimer interface, we can conclude that the increased exchange is reporting on the monomer formation. The fact that the rates are within error suggests that the tryptophan at the interface and the change in amide exchange are monitoring the same phenomenon; monomer formation.

### Observation of solvent penetration during cold denaturation

Because the cold denaturation of Pfk-2 occurs over hours of incubation at 3°C, we were able to probe changes in amide exchange throughout the protein during the denaturation process. As shown in Figs. 5 and 6, all of the peptides from regions throughout Pfk-2 showed an increase in the extent of backbone amide exchange as the protein is undergoing the slow process of cold denaturation. Thus, as Pfk-2 cold denatures, it adopts a monomeric, globally expanded structure. The C-terminal domain of L9 protein also adopts an expanded structure upon cold denaturation, as indicated

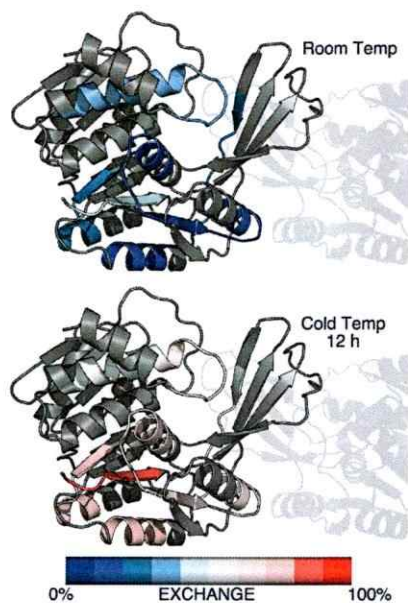


FIGURE 6 Change in solvent accessibility in the Pfk-2 structure upon cold denaturation. Peptides are colored according to their maximum exchange at room temperature (*top*) and after 12 h at cold temperature (*bottom*).

by pulse-field gradient NMR diffusion experiments (32). As pointed out by these authors, this type of protein expansion is most likely due to the weakening of hydrophobic interactions upon cold denaturation (32). Consistent with the idea that the expanded cold denatured state becomes more solvent penetrated, we observed more and more amide exchange throughout the folded protein core over the course of Pfk-2 cold denaturation. The fact that all regions of the protein showed similar rates of increasing amide exchange is consistent with a model in which monomer formation and expansion are concurrent. No buildup of a folded monomeric species was detected. In addition, the fact that amide exchange increases throughout the protein suggests that the increased amide exchange is due to solvent penetration occurring as the protein cold denatures.

The HXMS experiments allowed us to monitor probes throughout the Pfk-2 structure over time. Remarkably, for all of the different regions of the protein covered in the HXMS experiments, the rate of increased amide exchange over time in the cold was similar, with an average rate of increase being  $1.0 \times 10^{-4} \pm 0.25 \times 10^{-4} \text{ s}^{-1}$ . These results strongly suggest that cold denaturation of this enzyme is accompanied by solvent penetration that occurs globally throughout the structure. The only exception to this assertion is the active site, which appears to remain compact according to additional HXMS data (Table 2).

Theoretical studies have also suggested that cold denaturation involves destabilization of hydrophobic contacts within the protein core in favor of solvent-separated configurations (41,42). This process is similar to that observed for pressure-induced unfolding (43). The theoretical studies reveal that at physiological temperatures, water molecules close to the protein have at least one nonsaturated H-bond, which is pointing toward the protein and they form a thin solvation shell. As the temperature decreases, the energy of the water H-bonds increases and cages of water molecules are formed, thus increasing the solvation shell. Coupled with the destabilization of hydrophobic interactions within the protein, hydration of the protein core is predicted to occur in a sequence-independent manner. Recent experimental studies on cold denaturation of proteins using water- $^{17}\text{O}$  magnetic relaxation dispersion to characterize the single-molecule water dynamics in the hydration layer, provided information about the number of water molecules that interact directly with the protein and about their rotational correlation times. For a destabilized mutant of apomyoglobin and for  $\alpha$ -lactoglobulin in 4 M urea, these authors concluded that the protein molecules remain "relatively compact and are better described as solvent-penetrated than as unfolded" (44).

### CONCLUSION

Using HXMS, we have now been able to probe the local solvent accessibility of protein backbone NH groups and for

the first time, to our knowledge, obtain a view from the perspective of the protein, of solvent penetration during cold denaturation in the absence of other perturbants. Our results show a similar increase in solvent accessibility of the amides across the entire protein to a state that is apparently solvent penetrated, but still retains some degree of compaction. The proposed theoretical mechanism does not involve specific amino acid contacts, but rather appears to be a phenomenon of weakening of hydrophobic contacts and mostly a rearrangement of the water structure (45). One prediction of the theory is therefore that cold denaturation would affect the entire cooperatively folded core of the protein equally. The fact that local regions of Pfk-2 all increase solvent accessibility at equal rates during the cold denaturation process provides strong support for this theoretical prediction.

## SUPPORTING MATERIAL

Eight figures, supporting legends, and references (46–47) are available at [http://www.biophysj.org/biophysj/supplemental/S0006-3495\(13\)00457-8](http://www.biophysj.org/biophysj/supplemental/S0006-3495(13)00457-8).

This work was partially supported by Fondo Nacional de Desarrollo Científico y Tecnológico (Fondecyt) grants 1110137 and 1090336. César A. Ramírez-Sarmiento was supported by a Comisión Nacional de Investigación Científica y Tecnológica (Conicyt) graduate fellowship, a joint Mejoramiento de la Calidad y Equidad de la Educación (MECESUP) and Sciences, Technology and Innovation Program for the Americas (STIPAs) fellowship UCH0717 and a MECESUP fellowship UCH0713. We gratefully acknowledge Dr. Jon Fitchett from Lilly Biotechnology Center of San Diego for providing access to the Synapt G2 mass spectrometer.

## REFERENCES

- Matthews, C. R. 1993. Pathways of protein folding. *Annu. Rev. Biochem.* 62:653–683.
- Privalov, P. L. 1990. Cold denaturation of proteins. *Crit. Rev. Biochem. Mol. Biol.* 25:281–305.
- Privalov, P. L., Griko YuV, ..., V. P. Kutysenko. 1986. Cold denaturation of myoglobin. *J. Mol. Biol.* 190:487–498.
- Privalov, P. L., and S. J. Gill. 1988. Stability of protein structure and hydrophobic interaction. *Adv. Protein Chem.* 39:191–234.
- Damaschun, G., H. Damaschun, ..., D. Zirwer. 1993. Cold denaturation-induced conformational changes in phosphoglycerate kinase from yeast. *Biochemistry.* 32:7739–7746.
- Richardson, 3rd, J. M., S. D. Lemaire, ..., G. I. Makhatadze. 2000. Difference in the mechanisms of the cold and heat induced unfolding of thioredoxin h from *Chlamydomonas reinhardtii*: spectroscopic and calorimetric studies. *Biochemistry.* 39:11154–11162.
- Adrover, M., V. Esposito, ..., P. A. Temussi. 2010. Understanding cold denaturation: the case study of Yfh1. *J. Am. Chem. Soc.* 132:16240–16246.
- Lopez, C. F., R. K. Darst, and P. J. Rossky. 2008. Mechanistic elements of protein cold denaturation. *J. Phys. Chem. B.* 112:5961–5967.
- Mandell, J. G., A. M. Falick, and E. A. Komives. 1998. Identification of protein-protein interfaces by decreased amide proton solvent accessibility. *Proc. Natl. Acad. Sci. USA.* 95:14705–14710.
- Hughes, C. A., J. G. Mandell, ..., E. A. Komives. 2001. Phosphorylation causes subtle changes in solvent accessibility at the interdomain interface of methylesterase CheB. *J. Mol. Biol.* 307:967–976.
- Truhlar, S. M., J. W. Torpey, and E. A. Komives. 2006. Regions of IkappaBalpha that are critical for its inhibition of NF-kappaB-DNA interaction fold upon binding to NF-kappaB. *Proc. Natl. Acad. Sci. USA.* 103:18951–18956.
- Baez, M., C. A. M. Wilson, ..., J. Babul. 2012. Expanded monomeric intermediate upon cold and heat unfolding of phosphofructokinase-2 from *Escherichia coli*. *Biophys. J.* 103:2187–2194.
- Griko YuV, and V. P. Kutysenko. 1994. Differences in the processes of beta-lactoglobulin cold and heat denaturations. *Biophys. J.* 67:356–363.
- Mizuguchi, M., D. Hashimoto, ..., K. Nitta. 2000. Cold denaturation of alpha-lactalbumin. *Proteins.* 38:407–413.
- Kunugi, S., and N. Tanaka. 2002. Cold denaturation of proteins under high pressure. *Biochim. Biophys. Acta.* 1595:329–344.
- Azuaga, A. I., M. L. Galisteo, ..., P. L. Mateo. 1992. Heat and cold denaturation of beta-lactoglobulin B. *FEBS Lett.* 309:258–260.
- Chen, B. L., W. A. Baase, and J. A. Schellman. 1989. Low-temperature unfolding of a mutant of phage T4 lysozyme. 2. Kinetic investigations. *Biochemistry.* 28:691–699.
- Chen, B. L., and J. A. Schellman. 1989. Low-temperature unfolding of a mutant of phage T4 lysozyme. 1. Equilibrium studies. *Biochemistry.* 28:685–691.
- Dias, C. L., T. Ala-Nissila, ..., M. Grant. 2008. Microscopic mechanism for cold denaturation. *Phys. Rev. Lett.* 100:118101.
- Babul, J. 1978. Phosphofructokinases from *Escherichia coli*. Purification and characterization of the nonallosteric isozyme. *J. Biol. Chem.* 253:4350–4355.
- Bradford, M. M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72:248–254.
- Brown, P. H., and P. Schuck. 2006. Macromolecular size-and-shape distributions by sedimentation velocity analytical ultracentrifugation. *Biophys. J.* 90:4651–4661.
- Houtman, J. C., P. H. Brown, ..., P. Schuck. 2007. Studying multisite binary and ternary protein interactions by global analysis of isothermal titration calorimetry data in SEDPHAT: application to adaptor protein complexes in cell signaling. *Protein Sci.* 16:30–42.
- Schuck, P. 2003. On the analysis of protein self-association by sedimentation velocity analytical ultracentrifugation. *Anal. Biochem.* 320:104–124.
- Baez, M., and J. Babul. 2009. Reversible unfolding of dimeric phosphofructokinase-2 from *Escherichia coli* reveals a dominant role of inter-subunit contacts for stability. *FEBS Lett.* 583:2054–2060.
- Mandell, J. G., A. M. Falick, and E. A. Komives. 1998. Measurement of amide hydrogen exchange by MALDI-TOF mass spectrometry. *Anal. Chem.* 70:3987–3995.
- Wales, T. E., K. E. Fadgen, ..., J. R. Engen. 2008. High-speed and high-resolution UPLC separation at zero degrees Celsius. *Anal. Chem.* 80:6815–6820.
- Bai, Y., J. S. Milne, ..., S. W. Englander. 1993. Primary structure effects on peptide group hydrogen exchange. *Proteins.* 17:75–86.
- Croy, C. H., S. Bergqvist, ..., E. A. Komives. 2004. Biophysical characterization of the free IkappaBalpha ankyrin repeat domain in solution. *Protein Sci.* 13:1767–1777.
- Pometun, M. S., R. W. Peterson, ..., A. J. Wand. 2006. Cold denaturation of encapsulated ubiquitin. *J. Am. Chem. Soc.* 128:10652–10653.
- Luan, B., B. Shan, ..., D. P. Raleigh. 2013. Cooperative cold denaturation: the case of the C-terminal domain of the ribosomal protein L9. *Biochemistry.* 52:2402–2409.
- Li, Y., B. Shan, and D. P. Raleigh. 2007. The cold denatured state is compact but expands at low temperatures: hydrodynamic properties of the cold denatured state of the C-terminal domain of L9. *J. Mol. Biol.* 368:256–262.
- Shan, B. B., S. S. McClendon, ..., D. P. Raleigh. 2010. The cold denatured state of the C-terminal domain of protein L9 is compact and

- contains both native and non-native structure. *J. Am. Chem. Soc.* 132:4669–4677.
34. Romero-Romero, M. L., A. Inglés-Prieto, ..., J. M. Sanchez-Ruiz. 2011. Highly anomalous energetics of protein cold denaturation linked to folding-unfolding kinetics. *PLoS ONE*. 6:e23050.
  35. Zhang, Z., and D. L. Smith. 1993. Determination of amide hydrogen exchange by mass spectrometry: a new tool for protein structure elucidation. *Protein Sci.* 2:522–531.
  36. Engen, J. R., T. E. Smithgall, ..., D. L. Smith. 1997. Identification and localization of slow, natural, cooperative unfolding in the hematopoietic cell kinase SH3 domain by amide hydrogen exchange and mass spectrometry. *Biochemistry*. 36:14384–14391.
  37. Engen, J. R., W. H. Gmeiner, ..., D. L. Smith. 1999. Hydrogen exchange shows peptide binding stabilizes motions in Hck SH2. *Biochemistry*. 38:8926–8935.
  38. Weber, G. 1993. Thermodynamics of the association and the pressure dissociation of oligomeric proteins. *J. Phys. Chem.* 97:7108–7115.
  39. Foguel, D., and J. L. Silva. 1994. Cold denaturation of a repressor-operator complex: the role of entropy in protein-DNA recognition. *Proc. Natl. Acad. Sci. USA*. 91:8244–8247.
  40. Baez, M., C. A. Wilson, and J. Babul. 2011. Folding kinetic pathway of phosphofructokinase-2 from *Escherichia coli*: a homodimeric enzyme with a complex domain organization. *FEBS Lett.* 585:2158–2164.
  41. Oshima, H., T. Yoshidome, ..., M. Kinoshita. 2009. A theoretical analysis on characteristics of protein structures induced by cold denaturation. *J. Chem. Phys.* 131:205102.
  42. Das, P., and S. Matysiak. 2012. Direct characterization of hydrophobic hydration during cold and pressure denaturation. *J. Phys. Chem. B*. 116:5342–5348.
  43. Panick, G., R. Malessa, ..., C. A. Royer. 1998. Structural characterization of the pressure-denatured state and unfolding/refolding kinetics of staphylococcal nuclease by synchrotron small-angle X-ray scattering and Fourier-transform infrared spectroscopy. *J. Mol. Biol.* 275:389–402.
  44. Davidovic, M., C. Mattea, ..., B. Halle. 2009. Protein cold denaturation as seen from the solvent. *J. Am. Chem. Soc.* 131:1025–1036.
  45. Levy, Y., and J. N. Onuchic. 2006. Water mediation in protein folding and molecular recognition. *Annu. Rev. Biophys. Biomol. Struct.* 35:389–415.
  46. Parducci, R. E., R. Cabrera, ..., V. Guixé. 2006. Evidence for a catalytic Mg<sup>2+</sup> ion and effect of phosphate on the activity of *Escherichia coli* phosphofructokinase-2: regulatory properties of a ribokinase family member. *Biochemistry*. 45:9291–9299.
  47. Bork, P., C. Sander, and A. Valencia. 1993. Convergent evolution of similar enzymatic function on different protein folds: the hexokinase, ribokinase, and galactokinase families of sugar kinases. *Protein Sci.* 2:31–40.

**CAPITULO DOS:**  
**Caracterización biofísica del monómero  
aislado inducido por disociación del  
homodímero de Pfk-2 y Kdgk**



## 2.1. INTRODUCCION

El estudio del desplegamiento cinético y al equilibrio de la enzima dimérica fosfofructoquinasa-2 (Pfk-2) mediante la aplicación de diversas perturbaciones (adición de agentes caotrópicos [114, 126], aumento y disminución de temperatura [128, 148]) describe la acumulación de un intermediario monomérico cuyas propiedades estructurales (radio hidrodinámico; contenido de estructura secundaria; unión de anilino-8-naftaleno, ANS) son diferentes a las esperadas para un monómero compacto, sugiriendo que la disociación del dominio bimolecular compuesto por el dominio menor de ambas subunidades [94] sucede de manera concertada con el desplegamiento parcial de los monómeros. Mediante intercambio isotópico de hidrógeno por deuterio seguido de mediciones de espectrometría de masas, se ha logrado describir estructuralmente el estado final de la desnaturación a 3 °C de Pfk-2, el cual corresponde a una especie monomérica relativamente compacta (un estado intermediario, mas no completamente desplegado) altamente solvatada [148]. Sin embargo, existen ciertas diferencias entre este estado intermediario monomérico final observado mediante desplegamiento en frío y el estado intermediario observado mediante la adición de concentraciones pequeñas de cloruro de guanidinio (GndHCl), en particular en cuanto a la unión de la sonda ANS: para el intermediario inducido químicamente a 20° C, la intensidad de la señal de fluorescencia de la sonda ANS es significativamente menor que la observada para el intermediario inducido por incubación a 3° C [128], sugiriendo que la afinidad de la especie intermediaria de Pfk-2 por esta sonda hidrofóbica es mayor para el intermediario descrito en este último caso. Esto ocurre a pesar que el contenido de estructura secundaria medido por dicroísmo circular para ambas especies intermediarias es similar [128]. Por otro lado, el análisis del desplegamiento al equilibrio inducido por GndHCl del estado intermediario observado a 3° C para alcanzar el estado completamente desnaturado, muestra que tanto el cambio de energía libre ( $2.7 \pm 0.8$  kcal/mol) como el valor  $m$  ( $2.6 \pm 0.7$  kcal/mol/K), el cual se correlaciona con el área que se expone al

solvente producto de la desnaturación [150], son mayores que en el caso del desplegamiento del intermediario monomérico inducido por GndHCl a 20° C para alcanzar el estado desnaturado ( $1.7 \pm 1.5$  kcal/mol y  $1.3 \pm 0.4$  kcal/mol/K, respectivamente) [114, 128]. Esto sugiere que existen leves diferencias estructurales entre los estados intermediarios observados mediante distintos tipos de perturbaciones.

Al respecto, existen diversos modelos de estudio en los cuales se han observado diferencias en cuanto al ensamble conformacional observado en el punto final del desplegamiento, e incluso en cuanto a la vía a través de la cual se alcanza este estado desnaturado [151], de acuerdo al tipo de perturbación utilizada [152, 153, 154]. El trabajo pionero de Tanford y colaboradores [155] sugirió que los estados desnaturados inducidos mediante el aumento de la temperatura presentaban un mayor contenido de estructura secundaria que los estados inducidos por GndHCl. Consistentemente, la adición de concentraciones crecientes de GndHCl sobre el estado final del desplegamiento térmico de lisozima, quimotripsina y ribonucleasa A induce a una pérdida cooperativa de estructura secundaria, demostrando así la presencia de estructura secundaria residual [156]. Estas diferencias estructurales del estado final del desplegamiento también pueden encontrarse al inducir la desnaturación química con agentes caotrópicos diferentes, como urea y GndHCl, como ha sido determinado para lisozima,  $\alpha$ -quimotripsina,  $\beta$ -lactoglobulina [157], ribonucleasa A y ribonucleasa T1 [158]. Estudios similares sobre distintas variantes mutagénicas del residuo Lis73 de iso-1-citocromo C mediante dicroísmo circular y calorimetría de escaneo diferencial (DSC), demostraron que el área que se expone al solvente producto de la desnaturación inducida por GndHCl para las distintas mutantes es significativamente menor en comparación a la proteína silvestre, de acuerdo a lo estimado por el cambio en los valores  $m$ , lo que sugiere que los estados desnaturados de estas variantes son más compactos [159]. Por otra parte, los resultados del análisis del cambio en calor específico ( $\Delta C_p$ ) de estas mutantes, que también se correlaciona con el área expuesta al solvente [160], presenta un comportamiento dispar, es decir, algunas proteínas con sustituciones en el residuo

Lis73 exhiben una disminución del  $\Delta C_p$  (estado desnaturalado compacto) y otras un aumento. Por consiguiente, estos resultados sugieren diferencias en el grado de compactación entre las configuraciones estructurales del estado desnaturalado inducido por cada perturbación [159]. Por ende, dado que los valores  $m$  del desplegamiento a partir del intermediario obtenido por desnaturalación en frío y del intermediario químico de Pfk-2 a 20° C hasta el estado completamente desnaturalado fueron determinados usando la misma perturbación (adición de concentraciones crecientes de GndHCl) [128], la separación entre estos valores sugiere que existen diferencias estructurales en el grado de compactación de los estados intermediarios descritos y no del estado desnaturalado.

Uno de los estudios más interesantes respecto al grado de compactación de los estados finales del desplegamiento de proteínas, corresponde al análisis acucioso realizado sobre especies desnaturaladas inducidas químicamente y térmicamente de la ribonucleasa bovina pancreática mediante cambios en compresibilidad y volumen, demostrando que el punto final del desplegamiento térmico posee estructura secundaria residual y una alta concentración local de grupos no polares expuestos al solvente, lo que permite mantener una importante cavidad interna (en otras palabras, describe un estado denaturado compacto), mientras que el estado desnaturalado inducido por GndHCl se caracteriza por una exposición de un gran número de residuos aminoacídicos, resultando en un aumento en la hidratación pero en una reducción de la cavidad interna de la proteína [161]. En el caso de Pfk-2, esto puede explicar tanto las diferencias observadas para la unión de la sonda ANS como las diferencias en cuanto al área accesible al solvente expuesta por el desplegamiento químico de cada estado intermediario [128].

Una estrategia experimental que ha permitido determinar de manera directa las diferencias en el grado de compactación del estado desnaturalado en solución, corresponde a la difracción de rayos X de ángulo pequeño, la cual permite obtener una caracterización estructural gruesa del tamaño y la forma de la proteína en solución, en presencia o en ausencia de diversas perturbaciones [162]. La caracterización de los

estados desnaturalados inducidos por urea, alta y baja temperatura para la proteína inhibidora de subtilisina de *Streptomyces*, demostró que el radio de giro del estado desnaturalado depende de la perturbación utilizada, siendo el estado desnaturalado en frío más compacto que el estado desnaturalado por incremento de la temperatura, y que la inducción de la desnaturalación mediante agentes caotrópicos conduce a un estado altamente desordenado y expandido a un nivel esperado para una cadena polipeptídica completamente extendida en conformaciones arbitrarias [163].

Estas diferencias en el grado de compactación también son observadas en el caso de perturbaciones locales, como la introducción de mutaciones puntuales, en la interfaz de asociación de enzimas multiméricas. Este es el caso del dominio Apple 4 (A4) del factor de coagulación XI [164], el cual es un homodímero de 160 KDa presente en el plasma sanguíneo a concentraciones de 30 nM dada la presencia de un puente disulfuro intermolecular [165]. La mutación por serina del residuo responsable de este enlace covalente (C231S) provoca la leve acumulación de un intermediario monomérico cuyo radio hidrodinámico es menor que el de la especie dimerica y que alcanza un máximo del 75% del total de especies proteicas en solución a 4.3 M de GndHCl [166]. En este caso, el valor  $m$  para el proceso de disociación y desplegamiento del dímero de C231S A4 es menor que en el caso del dímero de la proteína silvestre, contrario a lo esperado dada la incapacidad de disociación de ésta última. Sin embargo, mediante experimentos de resonancia magnética nuclear (NMR) se determinó que las vueltas entre las cuales se establece el puente disulfuro intermolecular son altamente flexibles [167], por lo que la mutación de cisteína por serina provocaría eventos de desplegamiento local de estas regiones y, en consecuencia, una diferencia estructural importante entre el monómero de la proteína silvestre y mutante, dando cuenta de un aumento de 30% en el área accesible al solvente [166]. Un comportamiento similar se observa para las mutantes W19F y W99F del aporepresor de triptófano de *Escherichia coli* [48], las cuales promueven una desestabilización del dímero, la acumulación de un intermediario monomérico y el consecuente cambio de mecanismo de dos estados a tres

estados [168]. Asimismo, en el caso de la quinasa dimérica de creatina de músculo de conejo (MM-CK), ésta sigue un mecanismo de plegamiento similar al de Pfk-2 en presencia de GndHCl, detectándose la formación de intermediarios monoméricos durante este proceso [169]. El remplazo del residuo triptófano 210 por tirosina produce una desestabilización del dímero, causando la consiguiente disociación en dos monómeros inactivos a concentraciones bajas (0.2 a 0.4 M) de GndHCl. Estos monómeros presentan un radio de Stokes de 3.1 nm, el cual es menor al observado para el dímero (3.8 nm), y que solo es posible observar en el caso de la mutante W210Y, mientras que en el caso del dímero solo se observa un aumento del radio hidrodinámico a concentraciones crecientes de GndHCl. Este monómero compacto y estructurado aumenta su radio hidrodinámico a medida que aumenta la concentración de agente caotrópico sobre 0.4 M. Adicionalmente, en el rango de concentración de GndHCl de 0.2 a 0.4 M, la intensidad de fluorescencia extrínseca de la sonda hidrofóbica ANS es aproximadamente 50% mayor para la mutante W210Y que para la quinasa silvestre. Todos estos antecedentes sugieren que esta especie monomérica es inestable y que su estabilización depende de la formación del dímero [169].

Dado que la estabilidad energética del intermediario monomérico de Pfk-2 inducido térmicamente y químicamente es menor a 2.7 kcal/mol [114, 128], su estabilidad parece depender fuertemente de la formación del dímero. Además, las especies intermediarias observadas mediante estas perturbaciones tienen una diferencia de 1 kcal/mol en su estabilidad [128]. Por otro lado, los resultados observados para el dominio A4 del factor de coagulación XI y para la quinasa dimérica de creatina de músculo de conejo sugieren que, mientras las perturbaciones de solvente y temperatura tienen consecuencias globales sobre la estructura proteica, las consecuencias de mutaciones en las regiones de la interfaz de asociación en proteínas diméricas serían locales o de rango medio [166, 169]. En consecuencia, distintas perturbaciones sobre nuestro modelo de estudio pueden conducir a la formación de subunidades aisladas con

distintos grados de compactación y de accesibilidad al solvente. Es de nuestro interés poder describir estos cambios con el mayor detalle posible.

Sin embargo, si bien la enzima fosfofructoquinasa-2 (Pfk-2) de *Escherichia coli* es uno de los miembros mejor estudiados de la familia riboquinasa [97], sus características estructurales y su estado oligomérico no son propiedades compartidas por todos los miembros de esta agrupación. En concreto, dentro de la familia riboquinasa, a la cual pertenece Pfk-2, es posible encontrar al menos dos tipos de variaciones respecto a la estructura cuaternaria de la enzima Pfk-2: i) miembros de esta familia que corresponden a enzimas monoméricas, como la quinasa humana de adenosina [101]; ii) miembros diméricos que, pese a que establecen las interacciones proteína-proteína mediante la misma superficie de asociación contenida en la sábana  $\beta$  que compone el dominio menor de estas enzimas, exhiben rotaciones respecto a la disposición de los monómeros en el contexto oligomérico, lo cual modifica la forma general del dímero (Figura 2). La consecuencia más notoria de esta diferencia en la asociación proteína-proteína es el cambio de orientación de los sitios de unión a ligando para cada subunidad, los cuales se encuentran hacia el mismo lado en Pfk-2, mientras en enzimas como 2-ceto-3-desoxigluconato quinasa (Kdgk) de *Thermus thermophilus* [110] se encuentran en lados opuestos. Al analizar en detalle la estructura cristalográfica de las enzimas diméricas y monoméricas de esta familia de quinasa [97], es posible sugerir que estos cambios en el estado de agregación y en la estrategia de asociación se deben a modificaciones estructurales a nivel del dominio menor (Figura 2). En el caso de las quinasa monoméricas, se observa que su dominio menor contiene inserciones de elementos de estructura secundaria (una hebra  $\beta$  adicional y dos hélices  $\alpha$  contra la cara exterior de la sábana  $\beta$ ) que ocluyen la superficie de interacción proteína-proteína y estabilizan al monómero aislado [97, 101]. Para las quinasa diméricas cuyos sitios activos se encuentran en lados opuestos del oligómero, la hebra  $\beta_3$  (segunda de cuatro hebras  $\beta$  presentes en el dominio menor) no se encuentra intervenida por la presencia de un " $\beta$ -bulge" [111], por lo que esta hebra  $\beta$  preserva su estructura continua y no pueden

ser compartida con otra subunidad, como en el caso de Pfk-2, en donde los residuos Tre31, Ala32 y Pro33 introducen una interrupción que permite un giro de 90° que facilita la interacción de la parte distal de la hebra  $\beta$ 3 del dominio menor adyacente, formando un barril aplanado [94]. En consecuencia, la interacción entre los dominios menores de dos subunidades se asemeja a una sábana  $\beta$  antiparalela, extendida pero pronunciadamente curvada [110, 112].

En este contexto, es importante determinar si los cambios inducidos en la estructura de Pfk-2 por GndHCl, temperatura o por mutaciones de la interfaz de asociación son atribuibles a las características de la interfaz de asociación proteína-proteína localizada en el dominio menor y, por consiguiente, si las estructuras de las subunidades aisladas obtenidas mediante estas distintas perturbaciones son similares entre distintos modelos de enzimas díméricas de la familia riboquinasa que difieren en la topología de la interfaz de asociación y en la forma general del dímero.

Para poder determinar las características estructurales del monómero de Pfk-2 bajo distintos tipos de perturbaciones, en primer lugar se realizó un escaneo de mutaciones por alanina de la interfaz de asociación de esta enzima basado en el análisis *in silico* de la interacción proteína-proteína, obteniendo una mutante monomérica (L93A) a concentraciones de proteína de 20  $\mu$ M en ausencia de sal. A continuación, se comparó el grado de compactación de este monómero respecto al observado en el punto final del desplegamiento en frío de la enzima mediante experimentos de velocidad de sedimentación en ultracentrífuga analítica. Para finalizar, se comparó en detalle la accesibilidad al solvente de las subunidades aisladas obtenidas mediante los distintos tipos de perturbación (GndHCl, temperatura y mutación de L93 por alanina) mediante experimentos de intercambio isotópico de hidrógeno-deuterio medidos a través de espectrometría de masas (HXMS), con el fin de determinar las consecuencias de cada tipo de perturbación sobre distintas regiones locales de la proteína.

Adicionalmente, para determinar si existen similitudes estructurales entre subunidades aisladas obtenidas mediante diversos tipos de perturbaciones para otras

enzimas diméricas de la familia riboquinasa con distintas topologías de asociación, se realizaron experimentos *in silico* e *in vitro* de escaneo de mutaciones por alanina de la interfaz de asociación y experimentos de velocidad de sedimentación en ausencia y presencia de GndHCl en la enzima Kdgk de *Escherichia coli*, obteniendo dos mutaciones (Y101A y W103A) que conducen a la desestabilización parcial y total de la estructura del dímero, respectivamente. La enzima silvestre fue también cristalizada en presencia del sustrato MgADP y su estructura fue resuelta mediante remplazo molecular [170]. Además, la forma y tamaño en solución de la enzima silvestre y la mutante W103A fueron determinados mediante experimentos de difracción de rayos X de ángulo pequeño, demostrando que esta mutante presenta propiedades hidrodinámicas similares a las observadas para la mutante L93A de Pfk-2 mediante otras estrategias.

## 2.2. MATERIALES Y METODOS

### 2.2.1. Purificación y almacenamiento de Pfk-2 de *Escherichia coli*.

La enzima Pfk-2 fue purificada y almacenada de acuerdo a lo descrito por Babul [118]. Previo a la realización de los experimentos, la solución de almacenamiento de la enzima fue intercambiada por amortiguador Tris 50 mM pH 8.2, MgCl<sub>2</sub> 5 mM, DTT 2 mM mediante centrifugación utilizando las columnas Micro Bio-Spin 6 (Bio-Rad Laboratories Inc.). En función de las condiciones experimentales, la enzima fue concentrada utilizando un concentrador por centrifugación VivaSpin 3 (Sartorius Stedim Biotech S.A.). La concentración de proteína fue determinada mediante el método de Bradford [171], utilizando albumina de suero bovino (BSA) como estándar de calibración. La concentración de proteína se expresa en términos de la concentración de monómero.

### 2.2.2. Clonamiento, purificación y almacenamiento de Kdgk de *Escherichia coli*.

El gen codificante para la enzima Kdgk fue amplificado mediante la reacción en cadena de polimerasa (PCR), utilizando la polimerasa Pfu Ultra II Fusion (Agilent



Technologies, Inc.) y partidores directos y reversos apropiados (Integrated DNA Technologies), a partir de ADN cromosómico extraído mediante *kit* de purificación Masterpure (Epicentre Biotechnologies) de un cultivo crecido por 12 h a 37° C en medio Luria-Bertani (LB) de la cepa de *Escherichia coli* O157:H7.

El producto de PCR, cuyos extremos son romos debido al uso de Pfu, fue ligado a través de la estrategia independiente de ligasa TOPO [172], la cual se realiza sobre el vector de almacenamiento pENTR/TEV/TOPO (Invitrogen Corp.) que cuenta con la topoisomerasa covalentemente unida a los extremos libres de la hebra de ADN, para facilitar la ligación direccional del producto de PCR.

La enzima Kdgk fue subclonada a partir de este vector de almacenamiento en un vector de expresión pDEST17 (Invitrogen Corp.) por recombinación homóloga, el cual fue introducido en cepas de *E. coli* BL21. El constructo proteico final corresponde a una proteína con un epítipo de histidina N-terminal y un sitio de reconocimiento para la proteasa TEV que permite escindir este epítipo. La sobre-expresión se llevó a cabo una vez que los cultivos en medio LB alcanzaron una densidad óptica de 0.4-0.5 (longitud de onda de 600 nm). La inducción de la sobre-expresión se realizó a 25 °C durante 16 h, utilizando 1 mM IPTG, posteriormente las células se sedimentaron mediante centrifugación. El extracto crudo se obtuvo mediante resuspensión en amortiguador Tris 50 mM pH 8.2, 5 mM MgCl<sub>2</sub>, 2 mM beta-mercaptoetanol, 30 mM imidazol, 400 mM NaCl, seguido de sonicación y centrifugación, y luego la enzima fue purificada mediante cromatografía de afinidad por níquel inmovilizado, utilizando una columna HisTrap (GE Healthcare), eluyendo la proteína mediante un gradiente de concentración de imidazol entre 100 y 500 mM. La remoción del epítipo de histidina se realizó añadiendo la proteasa TEV (que posee un epítipo de histidina no escindible) en relación 1:50 molar/molar e incubando a temperatura ambiente por tiempos mayores a 2 h. La proteína fue dializada contra solución amortiguadora sin imidazol y separada de la proteasa TEV mediante una segunda ronda de cromatografía en la columna HisTrap, obteniendo la proteína sin epítipo en el frente de carga. Previo a la realización de los experimentos, la

solución de almacenamiento de la enzima fue intercambiada por amortiguador Tris 50 mM pH 8.2, MgCl<sub>2</sub> 5 mM, DTT 2 mM mediante centrifugación utilizando las columnas Micro Bio-Spin 6 (Bio-Rad Laboratories Inc.). En función de las condiciones experimentales, la enzima fue concentrada utilizando un concentrador por centrifugación VivaSpin 3 (Sartorius Stedim Biotech S.A.).

Para la sobre-expresión de las mutantes de la interfaz de asociación de Kdgk (ver sección 2.2.5), se utilizó una cepa BL21 pLys en remplazo de la cepa BL21.

### 2.2.3. Cristalización de la enzima Kdgk de *Escherichia coli* en presencia de MgADP

La cristalización de la enzima Kdgk fue realizada por el método de difusión de vapor en gota colgante [173]. La enzima fue concentrada hasta 5 mg/mL en amortiguador HEPES 50 mM pH 7.5, 5 mM ADP, 10 mM MgCl<sub>2</sub>. Luego, 2  $\mu$ L de proteína fueron mezclados con 2  $\mu$ L de solución de cristalización. Cristales difractables se obtuvieron en 100 mM HEPES pH 7.5, 22% PEG 4000, 10% isopropanol.

Para la difracción, los cristales fueron cubiertos con solución de crio-protección previo a la exposición. Los patrones de difracción con un desplazamiento angular de 0.5° fueron colectados a 100K en una fuente de ánodo rotatorio Rigaku Micromax007 con un detector Raxis IV++. Las reflexiones observadas fueron indexadas e integradas mediante el programa iMosflm [174], escalados con el programa SCALA [175] y las intensidades fueron interpretadas como amplitudes mediante CTRUNCATE [176], ambos del paquete de programas CPP4i [177]. La estructura de Kdgk en presencia de MgADP fue obtenida mediante remplazo molecular [170] en el servidor BALBES [178], utilizando la estructura apo de Kdgk de *S. flexineri* (PDB ID 3LHX). Los mapas de densidad electrónica previos al refinamiento de la estructura de Kdgk de *Escherichia coli* mostraron diferencias de densidad respecto al modelo de búsqueda, consistentes con la presencia de MgADP en el sitio de unión de sustratos. A partir de este punto, la reconstrucción de la estructura y el subsecuente refinamiento fueron realizados de manera iterativa mediante Coot [179, 180] y REFMAC5 [181], respectivamente.

#### 2.2.4. Escaneo in silico de mutaciones de residuos de la interfaz por alanina en Pfk-2 y Kdgk de *Escherichia coli*.

Para determinar los residuos de la interfaz a modificar mediante mutagénesis sitio-dirigida en las enzimas Pfk-2 y Kdgk de *Escherichia coli*, se realizó un análisis del efecto *in silico* del remplazo de los residuos de la superficie de interacción proteína-proteína de estas enzimas y de otras enzimas homólogas por alanina [182] mediante el uso del programa Robetta [183]. Para determinar si los sitios con efectos significativos en la estabilidad de la asociación del homodímero son conservados entre las distintas estructuras homólogas, se elaboró un alineamiento de secuencias basado en superposición estructural mediante el programa PROMALS3D [184].

Finalmente, para determinar qué tan favorables son los contactos nativos establecidos en la interfaz de asociación del homodímero de estas proteínas para la estabilidad de la proteína, se determinó el grado de frustración mutacional y configuracional mediante el programa Protein Frustratometer [185].

#### 2.2.5. Mutagénesis sitio-dirigida de Pfk-2 y Kdgk de *Escherichia coli*.

Las mutaciones sitio-específicas por alanina sobre la interface de asociación se realizaron mediante amplificación por PCR utilizando, para Pfk-2, el kit Genetailor (Invitrogen) y, para Kdgk, el kit QuickChange (Stratagene). Las mutantes generadas fueron L93A y V95A para Pfk-2 y Y101A y W103A para Kdgk. Los protocolos de purificación son similares a los realizados para las enzimas silvestres. La mutante V95A no pudo ser purificada con éxito.

#### 2.2.6. Sedimentación por velocidad de las enzimas Pfk-2 y Kdgk.

La enzima Pfk-2 fue sedimentada en una ultracentrífuga analítica Beckman Coulter Optima XLI equipada con ópticas de absorbancia en interferencia de Rayleigh, utilizando un rotor An60Ti de 4 espacios y celdas de centrifugación de 2 canales con ventanas de zafiro. Para la preparación de las muestras, el agente reductor DTT fue

remplazado por TCEP a una concentración final de 2 mM. Las muestras fueron incubadas a la temperatura y a las composiciones de la solución escogidas para el experimento de sedimentación por 24 h previo a su realización, de manera de alcanzar condiciones de equilibrio. El rotor, las celdas analíticas y la ultracentrifuga fueron equilibradas a la temperatura de trabajo previo a cargar las muestras, las que fueron posteriormente sometidas a centrifugación a una velocidad de 42000 rpm hasta conseguir sedimentación completa. La frontera móvil fue seguida mediante absorbancia a 280 nm y los datos resultantes fueron analizados mediante el programa SEDFIT, utilizando el modelo de distribución continua de coeficientes de sedimentación ( $c(S)$ , [186]), el cual permite obtener una aproximación del coeficiente de sedimentación de las especies en solución. Para el primero, las distribuciones se obtuvieron utilizando grillas de valores lineales de coeficiente de sedimentación entre 0.2 y 8.2 S, separados en pasos de 0.05 S. La abundancia de las especies en solución a cada temperatura se determinó mediante ajustes globales utilizando un modelo de análisis de especies disponible en el programa SEDPHAT [187]. La densidad y viscosidad de la solución amortiguadora utilizada en los experimentos de sedimentación y el volumen parcial específico de Pfk-2 y Kdgk fueron calculados mediante el programa SEDNTERP.

Para la enzima Kdgk silvestre y la mutante Y101A, se realizaron experimentos adicionales de sedimentación en presencia de GndHCl, para lo cual la enzima fue incubada a las distintas concentraciones de agente caotrópico al menos por 24 h previo a la realización de los experimentos de ultracentrifugación analítica.

#### 2.2.7. Intercambio isotópico para amidas de la cadena principal del intermediario monomérico inducido por GndHCl de Pfk-2 y de la mutante L93A.

El intercambio de hidrógeno por deuterio para las amidas de la cadena principal de la enzima Pfk-2 fue realizado de acuerdo al protocolo de Hughes [188]. Para el intermediario inducido por GndHCl, la enzima fue incubada en solución amortiguadora con 0.85 M de GndHCl durante 24 h a 25° C previo a la realización de los experimentos.

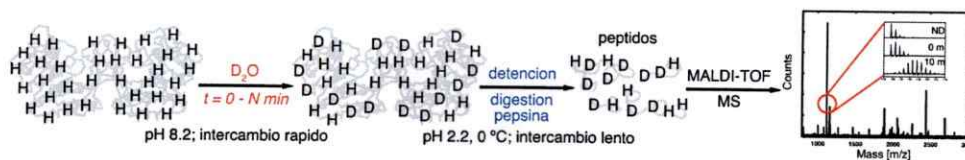
En el caso de la mutante L93A, la enzima fue incubada durante 2 h a 25°C previo a la realización de los experimentos.

Las enzimas, a una concentración de 40  $\mu\text{M}$  (20  $\mu\text{M}$  para L93A), fueron diluidas 10 veces en  $\text{D}_2\text{O}$  para iniciar la reacción de intercambio, la cual fue realizada durante 0-10 min. Una vez transcurrido el tiempo de intercambio, la reacción fue detenida diluyendo la muestra 10 veces en ácido trifluoroacético (TFA) al 0.1% en  $\text{H}_2\text{O}$ , pH 2.2 y 0 °C. La enzima fue proteolizada mediante el uso de pepsina inmovilizada en partículas de agarosa (Pierce Scientific), la cual se separó de la fracción soluble mediante centrifugación rápida (3 s) a 0 °C (Figura 4). Los detalles de la fragmentación proteolítica se indican más adelante (sección 2.2.8). El sobrenadante fue congelado rápidamente utilizando nitrógeno líquido y almacenado a una temperatura de -80 °C.

Ni la enzima Kdgk ni las mutantes Y101A y W103A fueron analizadas mediante esta estrategia experimental debido a la baja abundancia de fragmentos peptídicos luego de la digestión con pepsina (datos no mostrados).

#### 2.2.8. Digestión proteolítica de Pfk-2 y secuenciación de los fragmentos peptídicos.

Para evaluar distintas regiones de la enzima Pfk-2 de manera independiente en los experimentos de intercambio isotópico, es necesario fragmentar la macromolécula en péptidos de tamaño variable mediante el uso de proteasas. En este caso, la proteólisis fue realizada mediante el uso de pepsina inmovilizada en agarosa (Pierce Scientific). Las condiciones de proteólisis son equivalentes a las indicadas para la digestión proteolítica de Pfk-2 en condiciones nativas [148], donde una dilución de factor 10 de la enzima Pfk-2 en  $\text{D}_2\text{O}$  es nuevamente diluida 10 veces en TFA al 0.1% y pH 2.2 en  $\text{H}_2\text{O}$  y sometida a digestión con pepsina a 0°C. La pepsina (25  $\mu\text{L}$ ) fue activada mediante dos lavados con 1 mL de ácido trifluoroacético al 0.1% y pH 2.2. El tiempo de digestión utilizado fue de 5 min, similar al utilizado para la determinación de los fragmentos proteolíticos de la enzima Pfk-2 en condiciones nativas [148].



**Figura 4. Esquema del experimento de intercambio isotópico hidrógeno/deuterio (HXMS) de las amidas de la cadena principal de la enzima Pfk-2 medido por espectrometría de masas.**

En este experimento, la reacción de intercambio isotópico es iniciada mediante la dilución de la muestra proteica en agua deuterada a 25 °C. La reacción es permitida durante tiempos entre 0 y 10 min, y luego de transcurrido este tiempo, la reacción es detenida mediante la dilución en una solución de ácido trifluoroacético (TFA) al 0.1% a pH 2.2 y 0° C, y sometida a digestión proteolítica mediante el uso de pepsina inmovilizada pre-activada con solución TFA al 2%. El número de deuterones incorporados para cada péptido es cuantificado mediante espectrometría de masas, de acuerdo al método de desorción/ionización por láser asistida por matriz.

Para identificar a qué segmentos de la proteína corresponde cada uno de los péptidos obtenidos luego de la digestión con pepsina, se procedió a la secuenciación de estos fragmentos por espectrometría de masas tándem (MS/MS), de acuerdo al método de desorción/ionización por láser asistida por matriz (*matrix-assisted laser desorption/ionization*, MALDI), utilizando un equipo 4800 Plus MALDI-TOF-TOF (AB SCIEX) que cuenta con dos detectores en serie de tiempo de vuelo (*time-of-flight*, TOF-TOF) y realizando el análisis automatizado y manual de los espectros obtenidos con el servidor MASCOT (Matrix Science) y Protein Prospector [189], respectivamente.

#### 2.2.9. Espectrometría de masa.

Para cuantificar la incorporación de deuterio en las amidas de la cadena principal de los fragmentos peptídicos de Pfk-2, se obtuvieron espectros de masa mediante MALDI-TOF en un equipo Voyager DE-STR (Applied Biosystems) como se ha descrito con anterioridad [190]. La matriz utilizada correspondió a 4.5 mg/mL de ácido  $\alpha$ -ciano-4-hidroxicinámico en una solución 1:1:1 acetonitrilo/etanol/TFA al 0.1%. El pH de la matriz se ajustó a pH 2.2 mediante el uso de TFA al 2%. Las muestras fueron descongeladas rápidamente y sometidas a espectrometría mediante mezcla en proporción 1:1 con la matriz, y luego dispuestas en las placas de MALDI-TOF y secadas rápidamente en frío utilizando una bomba de vacío. Tanto la matriz como las placas de MALDI-TOF fueron equilibradas a 4°C previo a la realización del experimento, como se ha descrito anteriormente [190].

Los espectros de MALDI-TOF fueron analizados como se ha descrito anteriormente [144, 188], calculando el centroide de los picos isotópicos obtenidos para cada péptido analizado. Además, se realizó la corrección de estos valores de acuerdo a la contribución de las cadenas laterales debido al deuterio residual (4.5%) y a la pérdida por intercambio inverso (~50%). Estos datos corregidos fueron ajustados a una exponencial simple utilizando el programa KaleidaGraph v4.0 (Synergy Software, Reading, PA).

2.2.10. Seguimiento del desplegamiento inducido por GndHCl de la mutante L93A de Pfk-2 y de Kdgk silvestre y las mutantes Y101A y W103A mediante dicroísmo circular.

Las distintas enzimas fueron diluidas en distintas concentraciones de GndHCl en solución amortiguadora (50 mM Tris pH 8.2, 5 mM MgCl<sub>2</sub>, 2 mM DTT) e incubadas durante 24 h (mutante L93A de Pfk-2) ó 48 h (Kdgk) previo a la realización de las mediciones.

Las mediciones espectroscópicas de dicroísmo circular se realizaron en un equipo Jasco J-810 (Jasco Inc.) Los espectros fueron medidos en la región de ultravioleta lejano, entre 250 y 210 nm, utilizando una cuveta de paso óptico de 1 mm. Cada espectro obtenido corresponde a la acumulación de 4 escaneos a una temperatura de 20° C (mutante L93A de Pfk-2) ó 25° C (Kdgk).

2.2.11. Difracción de rayos X de ángulo pequeño de la enzima Kdgk de *Escherichia coli* y de la mutante W103A.

Todos los datos de difracción fueron colectados en la línea D11A-SAXS1 [191] en el Laboratorio Nacional de Luz Sincrotrón (Campinas, Sao Paulo, Brazil). Para las mediciones, se utilizó una longitud de onda de 1.55 Å y una distancia del detector a la muestra de 942 mm. Los datos de difracción fueron colectados utilizando un detector Pilatus. Las muestras (aprox. 100 µL) de las enzimas Kdgk y la mutante W103A fueron inyectadas en un capilar situado en la celda contenedora a una concentración de ~58 µM. Las muestras fueron irradiadas por tiempos variables (5 segundos a 5 min), de manera de monitorear el posible daño por la radiación y la estabilidad del láser. La difracción de la solución amortiguadora fue sustraída de cada uno de los espectros colectados y la curva de intensidades resultante fue utilizada para los análisis posteriores.

El radio de giro de la enzima silvestre y la mutante W103A de Kdgk fueron determinados mediante la aproximación lineal de Guinier [192], utilizando el programa PRIMUS [193]. La función de distribución de distancia P(r) fue generada mediante el



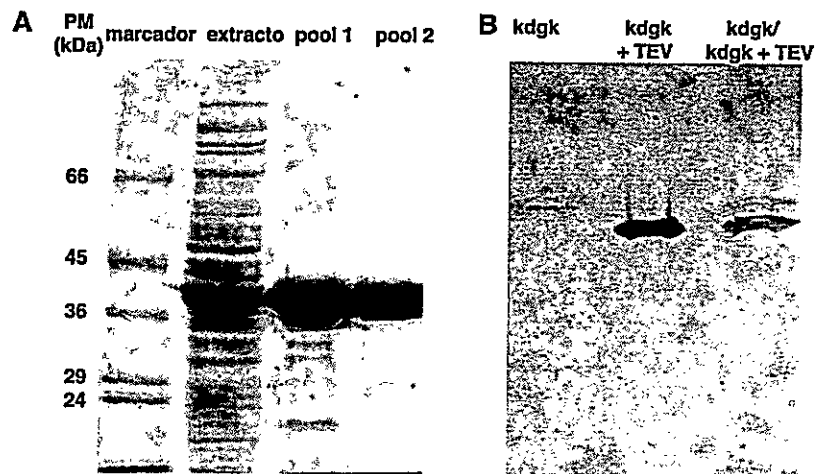
programa GNOM [194], y esta información fue utilizada para crear modelos *ab initio* de la forma y tamaño de cada una de las enzimas mediante el uso de los programas DAMMIN [195] y GASBOR [196], ambos disponibles en el paquete de análisis ATSAS [197]. Las curvas experimentales fueron comparadas a la curva teórica generada a partir de la estructura resuelta de Kdgk usando el programa CRY SOL [198].

## 2.3. RESULTADOS

### 2.3.1. Purificación y remoción del epítipo de histidina de la enzima Kdgk mediante proteólisis.

La purificación de la enzima Kdgk de *Escherichia coli* fue realizada a partir de un extracto crudo obtenido por lisado de células BL21 que sobre-expresan esta proteína unida a un epítipo de histidinas en posición amino-terminal mediante inducción por IPTG. La utilización de cromatografía de afinidad por níquel inmovilizado es suficiente para la purificación de la enzima Kdgk a partir de estos extractos, mostrando dos poblaciones de elución en función del aumento de concentración del competidor imidazol, de acuerdo a mediciones de concentración proteica por Bradford (Figura 5A).

Esta enzima fue sometida a proteólisis por la proteasa TEV recombinante, la cual también presente un epítipo de histidinas pero carece de la secuencia reconocida por esta proteasa para la digestión proteolítica. Mediante un segundo paso de cromatografía de afinidad por níquel inmovilizado, posterior a la digestión por uso de la proteasa TEV y diálisis contra solución amortiguadora sin imidazol, se observa que la enzima Kdgk sin el epítipo de histidinas eluye en el frente de cromatografía (es decir, en ausencia de imidazol). La remoción del epítipo de histidina y la homogeneidad de la purificación de la enzima Kdgk son evidentes, de acuerdo a lo observado mediante electroforesis en geles de poliacrilamida de las muestras antes y después de la digestión con la proteasa TEV (Figura 5B).



**Figura 5. Purificación y remoción del epítipo de histidinas en el extremo amino-terminal de la enzima recombinante Kdgk de *Escherichia coli*.**

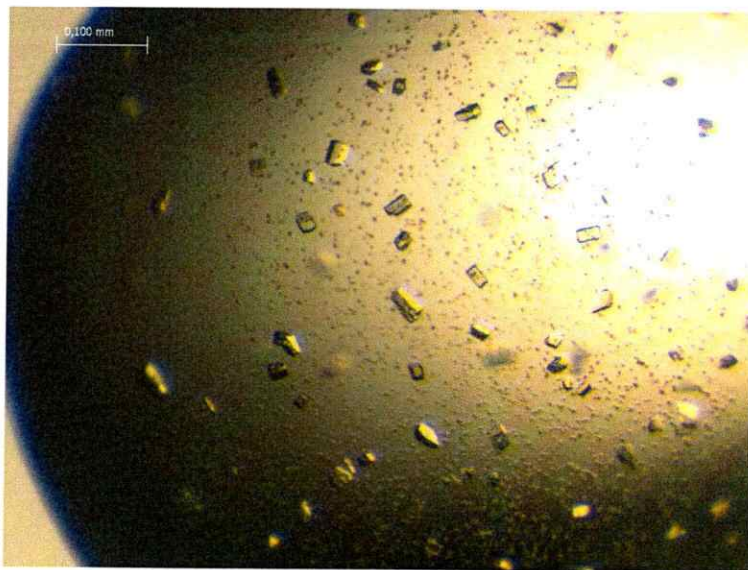
A) Purificación mediante cromatografía de afinidad por níquel inmovilizado. De izquierda a derecha, los carriles corresponde al marcador de peso molecular (BioRad), extracto crudo, primer y segundo pico de elución de la cromatografía de afinidad por níquel mediante uso de un gradiente de imidazol (*pool 1*, *pool 2*). Además, se indica el peso molecular (en KDa) de cada muestra de proteína contenida en el marcador.

B) Remoción del epítipo de histidina de Kdgk mediante digestión proteolítica utilizando la proteasa TEV. La enzima con el epítipo (primer carril, *Kdgk*) fue incubada con la proteasa por 2 h a temperatura ambiente, dializada y sometida a una segunda ronda de cromatografía (segundo carril, *Kdgk + TEV*). Se observa la diferencia en peso molecular entre la enzima con y sin este marcador de afinidad (carril 3, *Kdgk/Kdgk + TEV*).

### 2.3.2. Resultados preliminares de la estructura resuelta por remplazo molecular de la enzima Kdgk de *Escherichia coli*.

Las quinasas de 2-ceto-3-desoxigluconato resueltas hasta la fecha provienen, en su mayoría, de organismos extremófilos, como *Thermus thermophilus* [110], *Thermotoga marítima* [112] y *Sulfolobus solfataricus* [113]. Recientemente se ha depositado la estructura de un homólogo cercano a la enzima Kdgk de *Escherichia coli*, proveniente de *Shigella flexineri*, el cual comparte con ésta un porcentaje de identidad de secuencia del 98%. Desafortunadamente, esta estructura no presenta todos los elementos topológicos del dominio menor que permiten la interacción entre subunidades; en particular, no se observan 10 residuos por subunidad, que configuran las vueltas entre las hebras  $\beta$  que componen el dominio menor. Además, este cristal fue realizado en ausencia de sustratos, por lo que se desconocen las características del sitio de unión de esta enzima en organismos mesófilos, así como también si es que esta enzima es capaz de acomodar uno o dos cationes divalentes junto con el nucleótido, como en otras enzimas de la superfamilia [122]. Para abordar esto, se realizó la cristalización de la enzima Kdgk de *Escherichia coli* en presencia del producto MgADP, de acuerdo a lo indicado en la sección de Materiales y Métodos (sección 2.2.3). Los cristales obtenidos (Figura 6) fueron difractados y la estructura cristalográfica fue resuelta a 2.09 Å de resolución por remplazo molecular (Figura 7A y Tabla 1).

Los resultados preliminares del procesamiento de datos y refinamiento de la estructura de Kdgk de *Escherichia coli* sugieren que rondas adicionales de refinamiento permitirán resolver la estructura de las vueltas que conectan las hebras del dominio menor, dado el grado de completitud de las reflexiones indexadas (Tabla 1) y la presencia de abundante densidad electrónica en el dominio menor de la proteína, particularmente en las conexiones entre las vueltas del dominio menor (Figura 7B). Adicionalmente, se observó densidad electrónica para el ligando MgADP dentro del sitio de unión de ambas subunidades (Figura 7C).

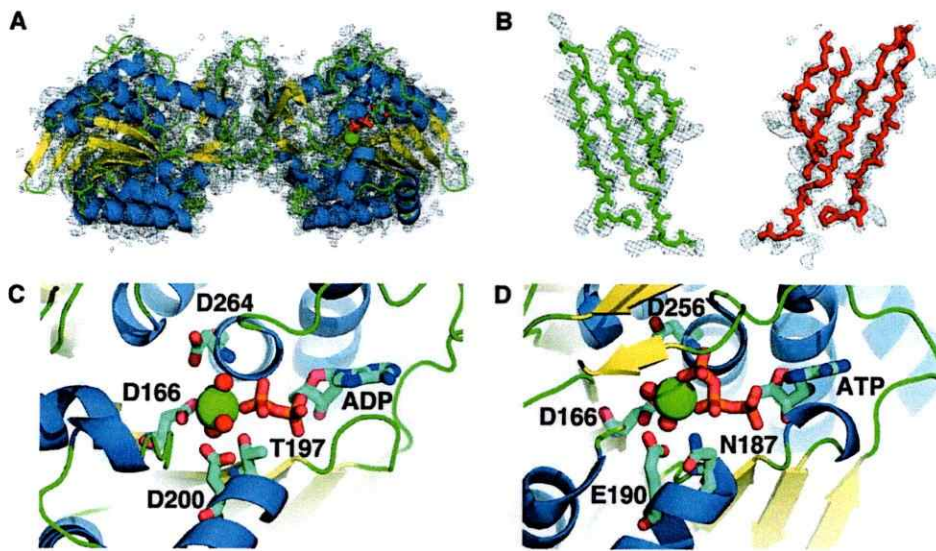


**Figura 6. Cristales obtenidos para la enzima Kdgk de *Escherichia coli* en presencia del ligando MgADP.**

Las condiciones de cristalización se indican en la sección de Materiales y Métodos (sección 2.2.3). Los cristales observados exhiben una apariencia que sugiere un sistema ortorrómbico, de un tamaño en el eje de mayor largo de aproximadamente 50  $\mu\text{m}$ .

**Tabla 1. Parámetros obtenidos del procesamiento de los patrones de difracción colectados y posterior refinamiento de la estructura de Kdgk de *Escherichia coli* en presencia de MgADP.**

Grupo Espacial	P212121		
Dimensiones Celda (Å) (a = b = c = 90°)	39.100	77.060	169.230
Resolución Máxima (Å)	2.09		
Resolución Mínima (Å)	34.15		
Compleitud (%)	99.25		
Numero de Reflexiones Únicas	30982		
Rmerge	0.07		
I/σ	3.3		
R-factor	0.2305		
R-free	0.3085		



**Figura 7. Estructura resuelta de la enzima Kdgk de *Escherichia coli* en presencia del ligando MgADP.**

A) Representación de cintas de la estructura del dímero de Kdgk de *Escherichia coli*, resuelta mediante cristalografía y difracción de rayos X. Se muestra también el mapa de densidad electrónica ( $2F_o - F_c$ ,  $\sigma = 1.0$ ) por sobre la representación de cintas. El ligando MgADP se muestra en representación de barras.

B) Densidad electrónica ( $2F_o - F_c$ ,  $\sigma = 1.0$ ) del dominio menor de la enzima Kdgk de *Escherichia coli*. Se muestran en representación de barras la subunidad A (verde) y B (rojo). Para la subunidad A, se observa una densidad electrónica que permite la resolución de las vueltas que conectan las hebras  $\beta$  del dominio menor.

C) Sitio de unión de MgADP de la enzima Kdgk de *Escherichia coli*. El ligando y las cadenas laterales se muestran en representación de barras, y se indica la posición e identidad de los residuos aminoacídicos que participan de la coordinación del catión divalente (magnesio) del complejo metal-nucleótido. El catión de magnesio se muestra como una esfera de color verde, coordinada por 4 moléculas de agua (esferas rojas).

D) Sitio de unión de MgATP de la enzima Pfk-2 de *Escherichia coli*. La orientación es equivalente a la figura 7C. En la figura se muestra la posición e identidad los residuos que participan de la coordinación del complejo metal-nucleótido. La representación de las cadenas laterales, el catión divalente y el nucleótido son equivalentes a las presentadas en la Figura 7C.

La orientación y configuración rotacional del ligando MgADP, así como también la posición del catión de magnesio, son similares a las observadas para la estructura cristalográfica de Pfk-2 (PDB ID 3CQD, Figura 7D). Además, solamente se observa un catión de magnesio en el sitio de unión (Figura 7C). Por otro lado, no existe una conservación de los residuos de asparragina y glutámico que permiten la coordinación de cationes divalentes en el sitio de unión (motivo <sup>187</sup>NXXE<sup>190</sup> en Pfk-2) (Figuras 7C y 7D). Los residuos de Kdgc que provienen del elemento de estructura secundaria equivalente al del motivo <sup>187</sup>NXXE<sup>190</sup> en Pfk-2 y que permiten la coordinación mediada por moléculas de agua del catión de magnesio, corresponden a los residuos Tre197 y Asp200. Pese a esto, el residuo aspártico involucrado en la transferencia del grupo fosforilo desde ATP al sustrato aceptor, perteneciente al motivo altamente conservado GXGD, sí es detectable en esta estructura y corresponde al residuo Asp264.

### 2.3.3. Determinación de los residuos claves para la estabilidad del homodímero de Pfk-2 y Kdgc.

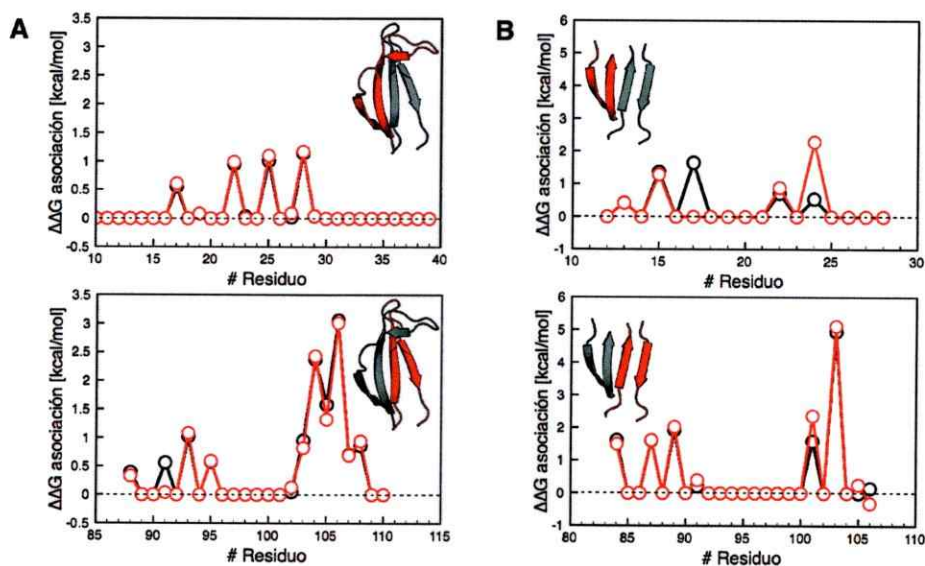
La identificación de los residuos importantes para la estabilidad de la asociación proteína-proteína de manera experimental es compleja, dado que a medida que el tamaño de la proteína y de la superficie de asociación entre subunidades es mayor, el número de residuos que deben ser sometidos a esta perturbación aumenta. Si consideramos el tamaño del dominio menor de Pfk-2 como ejemplo (PDB ID 3N1C), solo esta región de la proteína se compone de 52 residuos aminoacídicos. Es por esta razón que recurrimos a herramientas bioinformáticas para predecir los *hot-spots* de estabilidad de la asociación entre subunidades en las enzimas Kdgc y Pfk-2.

La proposición de los residuos claves para estabilidad de la interfaz de asociación monómero-monómero se realizó a través del escaneo *in silico* del efecto del remplazo de cada uno de estos residuos por alanina, utilizando el programa Robetta [183]. De acuerdo a esta estrategia, los residuos más importantes para la estabilidad del

homodímero (es decir, aquellos que provocan el mayor cambio en energía libre) de la enzima Pfk-2 de *Escherichia coli* (PDB ID 3NIC [104]) y Kdgk de *Shigella flexineri* (PDB ID 3LHX), el cual corresponde a un homólogo muy cercano de la enzima Kdgk de *Escherichia coli*, se localizan en la última hebra del dominio menor de estas proteínas (Figura 8).

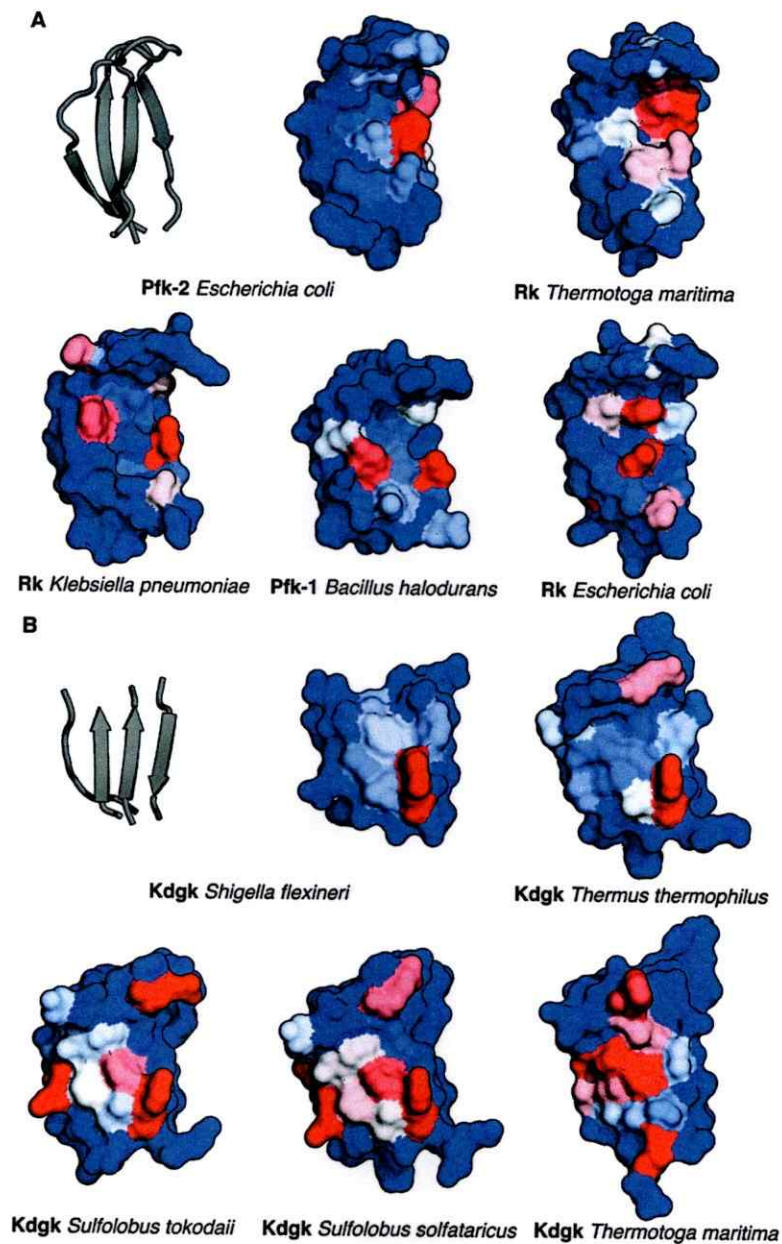
Dada la abundancia de estructuras cristalográficas de enzimas de la familia riboquinasa, nos preguntamos si la posición de los residuos que explican la estabilidad de la asociación entre subunidades está restringida en base la topología estructural del dominio menor. En otras palabras, si es que la posición de los *hot-spots* es conservada entre enzimas con interfaces similares a Pfk-2 o Kdgk. Esto se ilustra en la figura 9, en donde se representa el cambio de energía libre calculado para cada una de las mutaciones por alanina *in silico* realizadas para las siguientes enzimas: Pfk-2 de *Escherichia coli* (PDB ID 3NIC [104]), riboquinasa (Rk) de *Thermotoga maritima* (PDB ID 1VM7), Rk de *Escherichia coli* (PDB ID 1GQT [199]), Rk de *Klebsiella pneumoniae* (PDB ID 3I3Y), fructosa-1-fosfato quinasa (Pfk-1) de *Bacillus halodurans* (PDB ID 2ABQ), Kdgk de *Shigella flexineri* (PDB ID 3LHX), Kdgk de *Thermus thermophilus* (PDB ID 1V19 [110]), Kdgk de *Sulfolobus tokodaii* (PDB ID 2DCN), Kdgk de *Sulfolobus solfataricus* (PDB ID 2V78 [113]) y Kdgk de *Thermotoga maritima* (PDB ID 2AFB [112]). Se observa que, en el caso de las enzimas con topología similar a Pfk-2, los residuos más importantes para la estabilidad del homodímero se encuentran en los costados del dominio menor (Figura 9A), posiblemente para fortalecer la interacción de las hebras  $\beta$  laterales en la conformación cerrada de barril aplanado [94]. Para Pfk-2, este *hot-spot* se encuentra en los residuos Tir104, Arg105 y Fen106 (Figura 9A), y esta hebra  $\beta$  establece interacciones con la hebra proveniente de la subunidad adyacente [94]. En el caso de las enzimas Kdgk de diversos organismos, el único *hot-spot* observable en el dominio menor de la enzima Kdgk de *Shigella flexineri* es conservado en prácticamente todas las estructuras evaluadas (Figura 9B), el cual corresponde al residuo Trp103 en la enzima Kdgk de *Escherichia coli* y de *Shigella flexineri*, y corresponde a





**Figura 8. Cambios de estabilidad de la asociación de subunidades estimados mediante mutaciones sitio-específicas *in silico*.**

El escaneo *in silico* de los residuos claves para la estabilidad del homodímero fue realizado mediante remplazo puntual de los residuos del dominio menor por alanina sobre la estructura resuelta de Pfk-2 de *Escherichia coli* (Figura 8A, PDB ID 3N1C) y Kdgc de *Shigella flexineri* (Figura 8B, PDB ID 3LHX). Los gráficos en la parte superior representan el primer par de hebras  $\beta$  que emergen desde el dominio Rossmann [94], y los gráficos de la parte inferior corresponden al segundo par de hebras  $\beta$ , las cuales se destacan en color rojo en la representación de cintas del dominio menor incluida en el inserto de cada gráfico. En el eje  $x$  se indica la posición de cada residuo en la estructura primaria. Los símbolos en negro corresponden a la subunidad A y en rojo a la subunidad B.



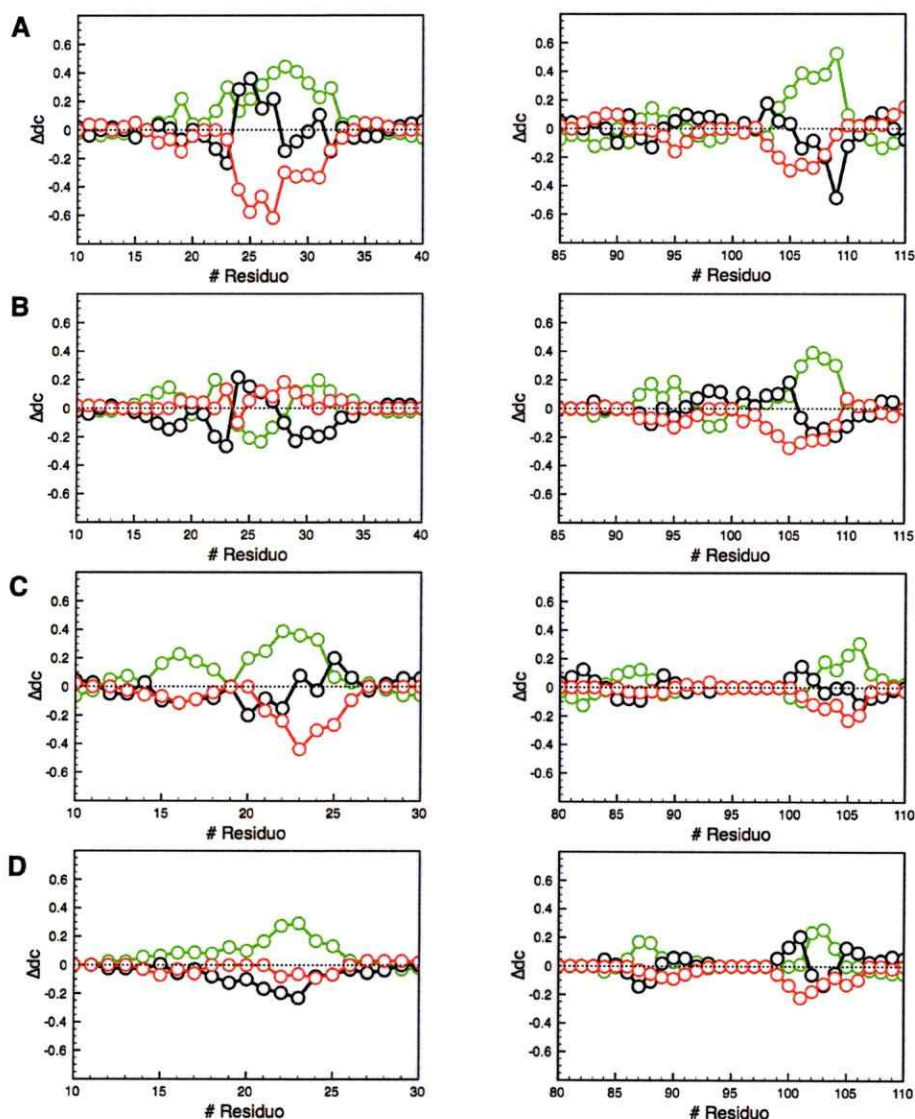
**Figura 9. Conservación de la posición de los residuos claves para la estabilidad del homodímero en la estructura del dominio menor**

A) Estructuras cuya topología de asociación es similar a Pfk-2 de *Escherichia coli*. El dominio menor se representa como un diagrama de superficie, cuya orientación es equivalente a la representación de cintas para el dominio menor de Pfk-2, pudiendo observar la cara de la sábana  $\beta$  que se ocluye producto de la asociación. El gradiente de color representa el cambio en energía libre producto del remplazo *in silico* por alanina, yendo desde ausencia de cambio de estabilidad (azul) hasta mayor cambio en estabilidad (rojo).

B) Estructuras cuya topología de asociación es similar a Kdgk de *Escherichia coli*. La representación es equivalente a la Figura 9A. La orientación se ilustra en la representación de cintas del dominio menor de Kdgk de *Shigella flexineri*.

tirosina en otros miembros provenientes de organismos termófilos. Además, se observa un segundo *hot-spot* localizado en la hélice 3<sub>10</sub> del dominio menor, por lo que esta modificación topológica también es importante para el establecimiento de la interfaz de asociación.

La estabilización provocada por estos residuos puede ser explicada en función del principio de la mínima frustración [12], ya que las regiones de la proteína que son claves para la estabilidad de su plegamiento están enlazadas por una red de interacciones locales que individualmente son mínimamente frustradas [200], mientras que el aumento de la frustración energética local tiene relación con regiones que participan de reconfiguraciones estructurales que permiten la estabilización de intermediarios de plegamiento [201] o que están asociados a sitios relevantes para la función biológica que desempeñan, como cambios conformacionales para la unión de ligandos [7, 87] y reestructuraciones entre estados múltiples de baja energía [202]. Asimismo, se ha demostrado que cerca del 25% de los contactos reducen su frustración como producto de la oligomerización [200]. Motivados por esta propuesta, decidimos determinar el índice de frustración local del dominio menor de Pfk-2 y Kdgk tanto para las subunidades aisladas como en el contexto homodimérico, de manera de determinar si los cambios en estabilidad predichos por las mutaciones *in silico* por alanina pueden ser explicados en base al principio de mínima frustración [12]. Tanto para Pfk-2 como para Kdgk, se observa un aumento importante en la densidad de contactos nativos mínimamente frustrados como consecuencia de la asociación del homodímero en regiones específicas del dominio menor (Figura 10), las cuales son ocupadas por los residuos que explican el cambio de estabilidad energética calculado mediante el escaneo *in silico* por mutaciones por alanina (Figura 8). Esta afirmación es válida tanto para el cambio en frustración configuracional (grado de frustración energética para las interacciones nativas entre dos residuos, en comparación con otras interacciones que estos residuos puedan efectuar en otras estructuras compactas en donde la distancia entre residuos y la densidad de contactos es diferente) y mutacional (grado de frustración energética de los



**Figura 10. Cambio en el grado de frustración local como consecuencia de la oligomerización de Pfk-2 y Kdgk.**

Se muestra el cambio en la fracción de contactos mínimamente frustrados en verde, altamente frustrados en rojo y neutralmente frustrados en negro, para las enzimas Pfk-2 (A, B) y Kdgk (C, D), de acuerdo a los criterios de frustración configuracional (A, C) y mutacional (B, D). La diferencia en la fracción de contactos ( $\Delta dc$ ) en cada categoría se calculan como  $\Delta dc = dc(\text{dímero}) - dc(\text{monómero})$ . Valores positivos de  $\Delta dc$  indican que este tipo de contactos es más abundante en el dímero que en el monómero. Los gráficos a la izquierda representan el primer par de hebras  $\beta$  que emergen desde el dominio Rossmann [94], y los gráficos a la derecha corresponden al segundo par de hebras  $\beta$ .

residuos que participan de la interacción nativa en comparación con otros residuos en esa misma posición) [185], lo que sugiere que tanto la identidad de estos residuos como la posición que ocupan en el dominio menor son una consecuencia de la topología de la interfaz de asociación proteína-proteína. De esta manera, los residuos claves para la estabilidad del oligómero ocupan posiciones donde se reduce la frustración energética como consecuencia de la interacción proteína-proteína (Figura 10)

En base a estos antecedentes, se realizaron mutaciones puntuales experimentales para reemplazar los residuos de la interfaz de asociación por alanina, obteniendo una mutante de estudio para Pfk-2 de *Escherichia coli* (mutante L93A) y dos mutantes de Kdgk de *Escherichia coli* (mutantes Y101A y W103A).

#### 2.3.4. La mutante L93A de Pfk-2 y la mutante W103A de Kdgk son monoméricas.

Las mutantes de interfaz L93A de Pfk-2, Y101A y W103A de Kdgk fueron purificadas de la misma manera en que fue realizado para las enzimas silvestres. Para determinar el estado de agregación de estas enzimas en solución, realizamos experimentos de velocidad de sedimentación por ultracentrifugación analítica, utilizando concentraciones de proteína de ~20  $\mu\text{M}$  para la enzima Pfk-2, y de ~13  $\mu\text{M}$  para la enzima Kdgk. Los resultados muestran que la mutación del residuo de Leu93 por alanina en Pfk-2 y Trp103 por alanina en Kdgk provoca la disociación de la enzima (Figura 11A y 11B), mientras que la mutante Y101A de Kdgk presenta una mezcla de dímero y monómero en solución (Figura 11B y 11C). Estos resultados están en buen acuerdo con los cambios de energía libre de asociación calculados mediante el procedimiento *in silico* (Figura 8); el cambio energético esperado para la mutante L93A es de 1.02 kcal/mol, mientras que los reemplazos de tirosina y triptófano por alanina en Kdgk exhiben un cambio energético teórico de 1.59 y 4.95 kcal/mol, respectivamente. Esto presupone que el reemplazo de Y101A tiene un efecto menor en la estabilidad del homodímero de Kdgk.

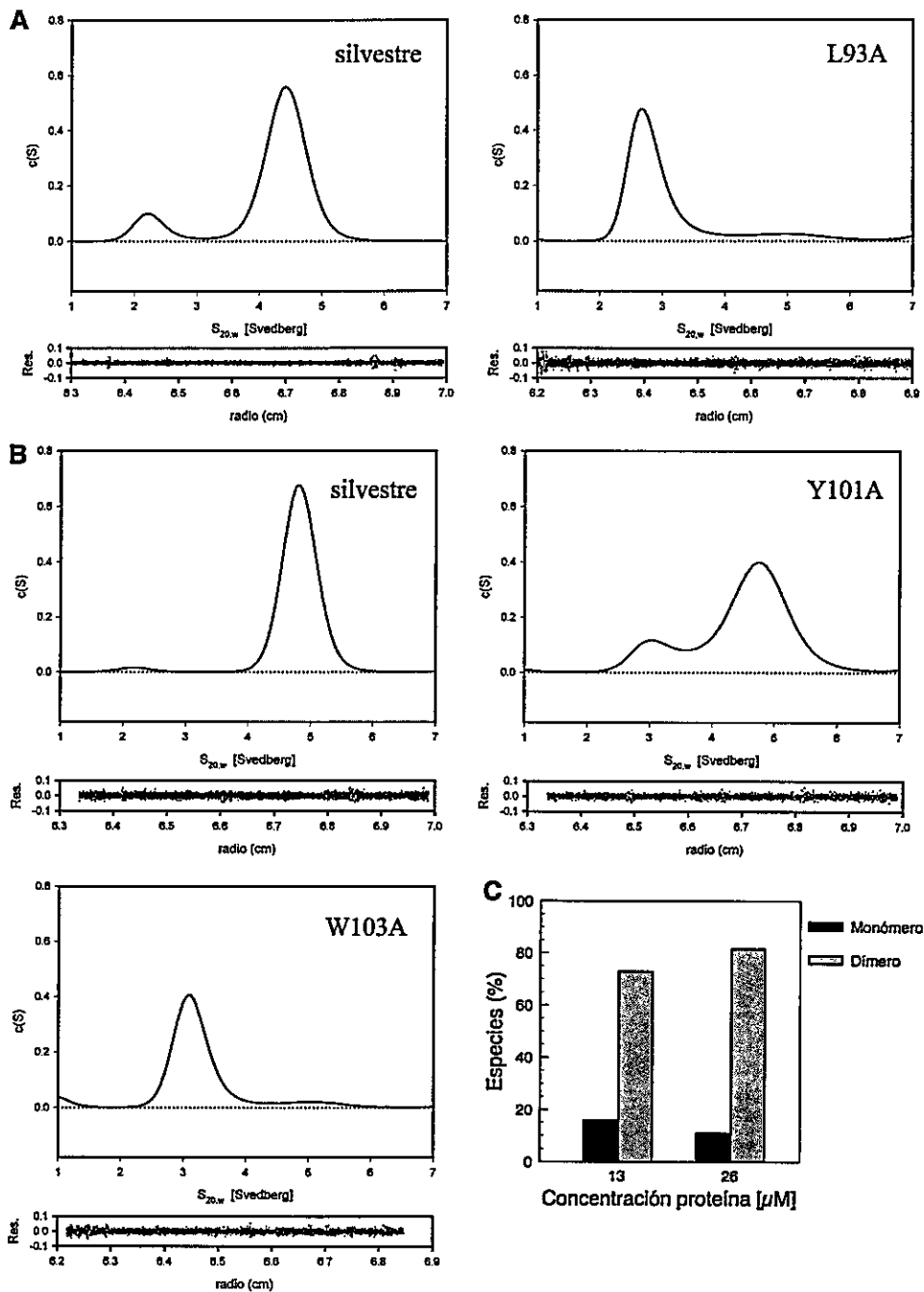


Figura 11. Disociación de las enzimas Pfk-2 y Kdgk de *Escherichia coli* mediante mutagénesis sitio-dirigida de la interfaz de asociación proteína-proteína.

A) Distribución del coeficiente de sedimentación y residuales del ajuste al modelo de distribución para la enzima Pfk-2 silvestre y la mutante L93A a una concentración de proteína de 20  $\mu\text{M}$ .

B) Distribución del coeficiente de sedimentación y residuales del ajuste para la enzima Kdgk silvestre y las mutantes Y101A y W103A a una concentración de proteína de 13  $\mu\text{M}$ .

C) Abundancia de las especies monomérica y dimerica para la mutante Y101A a dos concentraciones de proteína.

Para determinar si efectivamente la mutante Y101A introduce algún cambio menor en la estabilidad del homodímero, realizamos un segundo experimento de velocidad de sedimentación utilizando el doble de concentración de proteína (~26  $\mu\text{M}$ ). En este caso, la abundancia de la especie dimérica aumenta, en desmedro de la abundancia de la especie monomérica (Figura 11C). Adicionalmente, realizamos experimentos de ultracentrifugación analítica en presencia de concentraciones crecientes de GndHCl, de manera de determinar si, para el caso de la mutante Y101A, la disociación de las subunidades ocurre a concentraciones menores de agente caotrópico en comparación a la enzima silvestre, para una misma concentración de proteína (~13  $\mu\text{M}$ ). De estos resultados se desprende que, en efecto, la mutación debilita la interacción entre subunidades, dado que la disociación ocurre a concentraciones menores del agente caotrópico al cual ha sido enfrentada (Figura 12).

Tomando en consideración que el enfrentamiento de Pfk-2 tanto a agentes caotrópicos [114] como a bajas temperaturas por tiempos prolongados [128, 148] provoca la disociación del homodímero en subunidades parcialmente desplegadas cuyo radio hidrodinámico es mayor al esperado para el monómero compacto, es posible que los radios hidrodinámicos de los estados disociados inducidos por las mutantes presenten comportamientos similares, dando pistas acerca del acoplamiento entre los procesos de disociación y desplegamiento cuando la perturbación es local. Utilizando los datos de velocidad de sedimentación, calculamos el radio hidrodinámico de las especies monoméricas de Pfk-2 y Kdgk generadas por mutación sitio-dirigida y de las especies monoméricas inducidas por GndHCl. De estos datos se desprende que, efectivamente, la especie monomérica observada mediante perturbaciones locales (mutagénesis sitio-dirigida) es más compacta que las especies inducidas mediante la adición de GndHCl (Tabla 2). El radio hidrodinámico de la especie monomérica observada para la mutante L93A es un 30% menor que el observado para el intermediario monomérico inducido por agentes caotrópicos o desnaturalación en frío [114, 128]. Un comportamiento similar se observa para la mutante W103A de Kdgk, la cual es ~12% más compacta que la

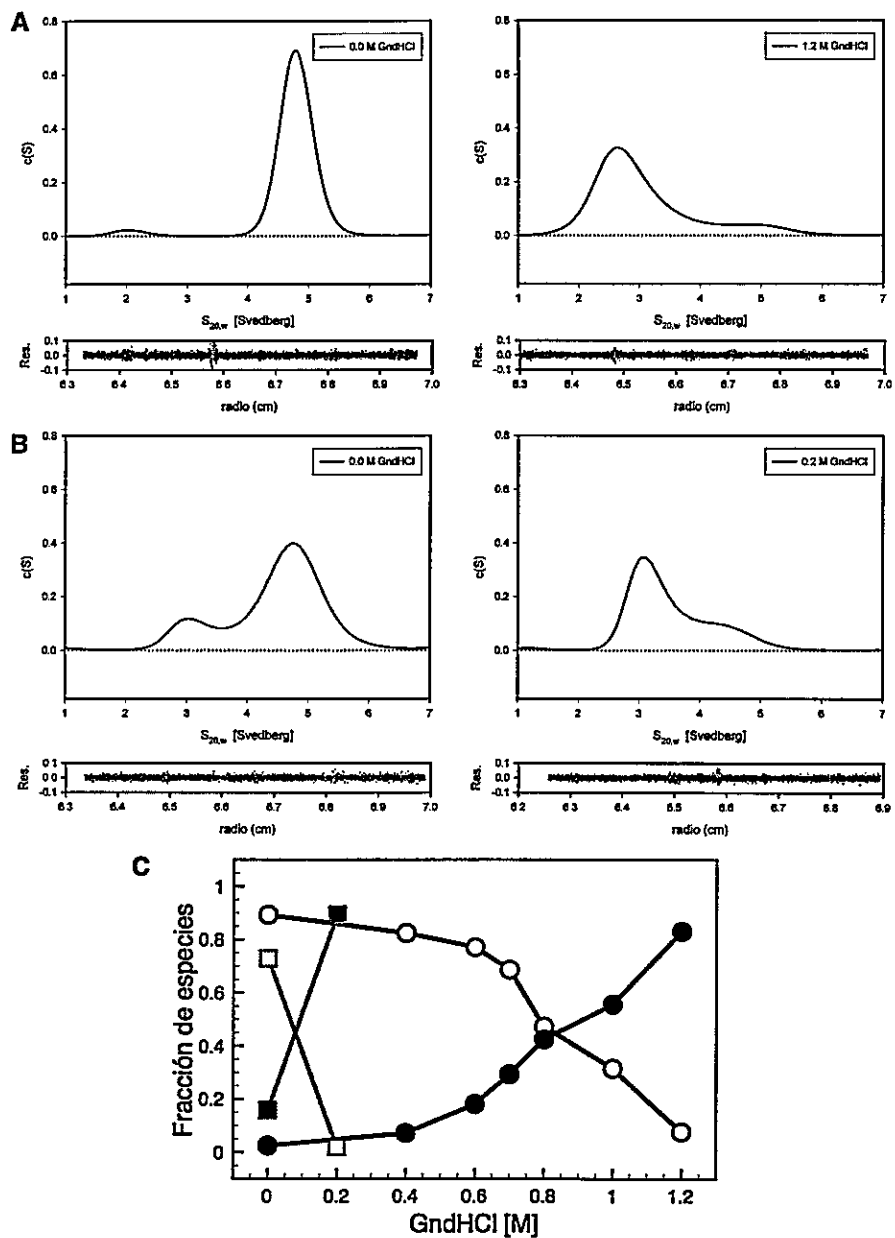


Figura 12. Efectos de la adición de agentes caotrópicos sobre la distribución de especies oligoméricas en la enzima Kdgk silvestre y la mutante Y101A.

A) Distribución del coeficiente de sedimentación y residuales del ajuste al modelo de distribución para la enzima Kdgk silvestre a concentraciones de 0 y 1.2 M de GndHCl.

B) Distribución del coeficiente de sedimentación y residuales del ajuste a modelo de distribución para la mutante Y101A de la enzima Kdgk a concentraciones de 0 y 0.2 M de GndHCl.

C) Distribución de dímeros (símbolos vacíos) y monómeros (símbolos llenos) en función de la concentración de GndHCl para la enzima Kdgk (círculos) y la mutante Y101A (cuadrados).



**Tabla 2. Propiedades hidrodinámicas de las especies observadas mediante ultracentrifugación analítica para las enzimas silvestres y mutantes de Kdgk y Pfk-2 \***

A) Propiedades hidrodinámicas para especies observadas por mutagénesis sitio-especifica						
	S20w, mon (S)	S20w, dim (S)	Rh, mon (nm)	Rh, dim (nm)	Mon (%)	Dim (%)
<b>Kdgk</b>	-	4.83 ± 0.01	-	3.7 ± 0.1	-	95 ± 1
<b>Y101A</b>	3.20 ± 0.02	4.79 ± 0.02	3.1 ± 0.4	3.8 ± 0.5	18 ± 1	73 ± 1
<b>W103A</b>	3.34 ± 0.2	5.51 ± 0.22	3.0 ± 0.1	3.8 ± 0.1	84 ± 5	6 ± 1
<b>Pfk-2</b>	2.40 ± 0.06	4.52 ± 0.1	2.9 ± 0.4	3.5 ± 0.1	13 ± 3	75 ± 6
<b>L93A</b>	2.86 ± 0.01	5.08 ± 0.06	2.9 ± 0.1	3.8 ± 0.2	75 ± 1	6 ± 1

B) Propiedades hidrodinámicas para especies observadas por adición de GndHCl						
<b>Kdgk<sup>§</sup></b>	2.91 ± 0.03	5.18 ± 0.01	3.3 ± 0.3	4.2 ± 0.7	79 ± 6	6 ± 2
<b>Y101A<sup>†</sup></b>	3.58 ± 0.01	5.04 ± 0.10	2.8 ± 0.2	3.4 ± 0.3	82 ± 3	15 ± 1

<sup>§</sup>Calculadas a partir de experimentos de velocidad de sedimentación a 1.2 M de GndHCl.

<sup>†</sup>Calculadas a partir de experimentos de velocidad de sedimentación a 0.2 M de GndHCl.

\*Otras especies observadas en la sedimentación (componentes de la solución y/o agregados) los cuales pueden alcanzar un máximo de ~10% de la composición total, no se muestran en la tabla.

especie monomérica observada por la adición de GndHCl.

Si bien los radios hidrodinámicos de las especies generadas por mutagénesis sitio-dirigida son más compactos que los de las especies inducidas químicamente, no necesariamente coinciden con los radios hidrodinámicos esperados para las especies dimericas y monoméricas compactas observadas en la estructura cristalográfica. Para poder determinar el radio hidrodinámico teórico para las especies dimerica y monomérica compacta de Kdgk y Pfk-2, utilizamos el programa HYDROPRO [203], el cual calcula el coeficiente de sedimentación teórico de estas especies considerando las condiciones de viscosidad y densidad de las soluciones a la temperatura a la cual se realizaron los experimentos de ultracentrifugación analítica. A partir de este coeficiente de sedimentación, el radio hidrodinámico puede ser calculado como:

$$Rh = PM(1 - v\rho)/6N\pi\eta S_{20,w}$$

donde Rh es el radio hidrodinámico, PM es el peso molecular, v es el volumen parcial específico,  $\rho$  es la densidad de la solución a la temperatura de trabajo, N es el número de Avogrado,  $\eta$  es la viscosidad de la solución a la temperatura de trabajo y  $S_{20,w}$  es el coeficiente de sedimentación calculado a partir de la información estructural.

A través de esta relación, se determinó que el radio hidrodinámico teórico para el dímero de Pfk-2 y Kdgk es de 3.4 nm, mientras que para la subunidad aislada compacta de Pfk-2 y Kdgk es de 2.6 nm. De esto se desprende que tanto la mutante L93A de Pfk-2 como la mutante W103A de Kdgk representan especies monoméricas cuyos radios hidrodinámicos son mayores a los esperados para el monómero compacto, pero menores a los observados para los estados intermediarios inducidos mediante perturbaciones globales [114, 128, 148].

#### 2.3.5. La mutante L93A de Pfk-2 representa los primeros rearrreglos estructurales producto de la disociación.

Para determinar las características estructurales locales de la mutante monomérica de Pfk-2 en comparación con las propiedades determinadas para la enzima

desnaturada por agentes caotrópicos y por baja temperatura [148], realizamos estudios de intercambio isotópico medidos por espectrometría de masas (HXMS) de la mutante L93A y del intermediario monomérico inducido por GndHCl de la enzima Pfk-2. Los péptidos que fueron cuantificados a través de esta metodología representan distintas regiones de la proteína, incluyendo la interfaz de asociación entre subunidades, y representan una cobertura de ~38% de la secuencia primaria de la proteína (Tabla 3).

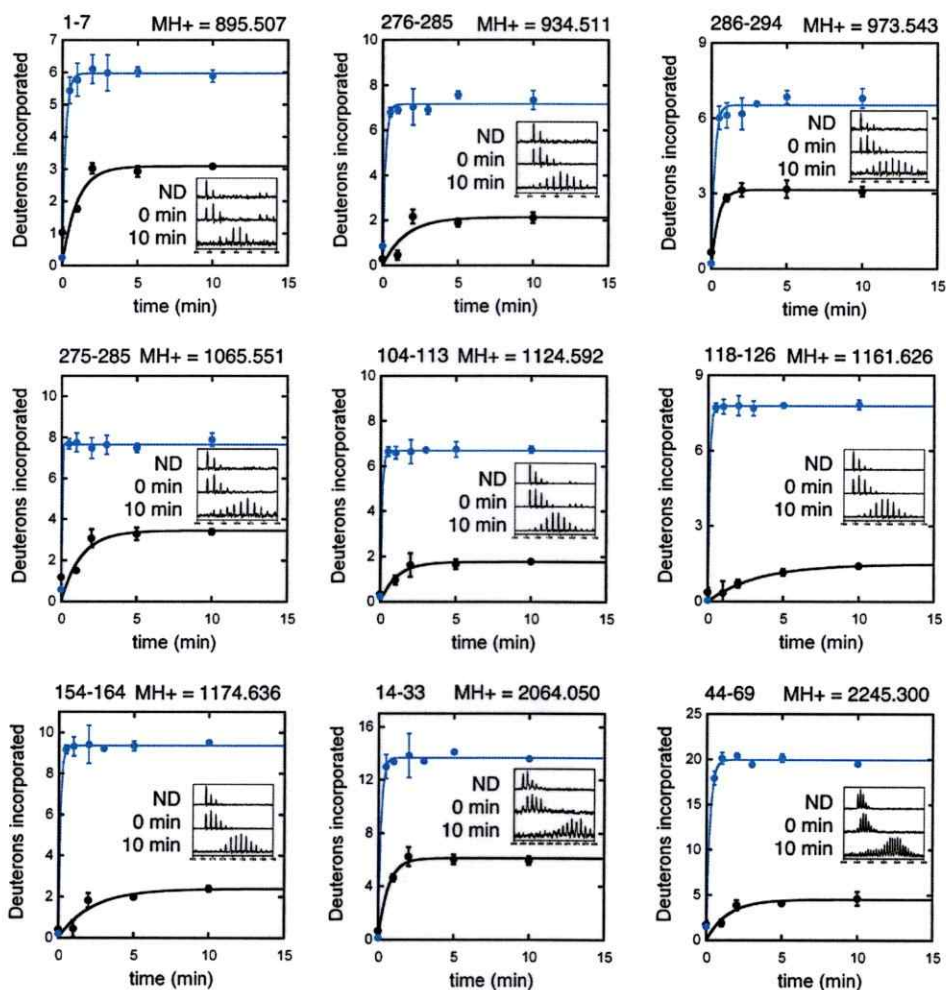
En primer lugar, realizamos estudios de intercambio de hidrógeno por deuterio de la enzima Pfk-2 silvestre incubada en 0.85 M de GndHCl por 24 h para determinar si existen diferencias estructurales entre el intermediario monomérico inducido por bajas temperaturas y el intermediario monomérico inducido por GndHCl, dado que exhiben diferencias en la energía libre de estabilización del intermediario y en las afinidades por la sonda de fluorescencia extrínseca ANS [128]. Los resultados obtenidos demuestran que el estado intermediario inducido por agentes caotrópicos presenta una mayor accesibilidad al solvente que el estado dímérico de Pfk-2 silvestre a 25 °C y que el estado inducido por bajas temperaturas (Figura 13). Esto es consistente con las diferencias en la energía libre de estabilización del intermediario inducido por las distintas perturbaciones [128] y sugiere que, si bien las perturbaciones aplicadas causan la disociación de las subunidades, los cambios estructurales a nivel de las subunidades dependen fuertemente de las propiedades del solvente. Esto, a su vez, es consistente con recientes teorías acerca del desplegamiento por presión o temperatura en donde se considera la solvatación de contactos observados en el estado nativo de proteínas, que sugieren que, luego de ocurrida la penetración del solvente en dichos contactos, la proteína hidratada puede ser expandida, y por ende su accesibilidad al solvente puede aumentar, de manera no cooperativa [204].

A continuación, determinamos los rearrreglos estructurales locales de la enzima monomérica L93A mediante la estrategia de HXMS. Como se observa en la figura 14, para la mayoría de los péptidos analizados se observa un leve incremento en el número de amidas de la cadena principal de la proteína que están disponibles para el intercambio

**Tabla 3. Fragmentos peptídicos de la enzima Pfk-2 y la mutante L93A analizados mediante espectrometría de masas.**

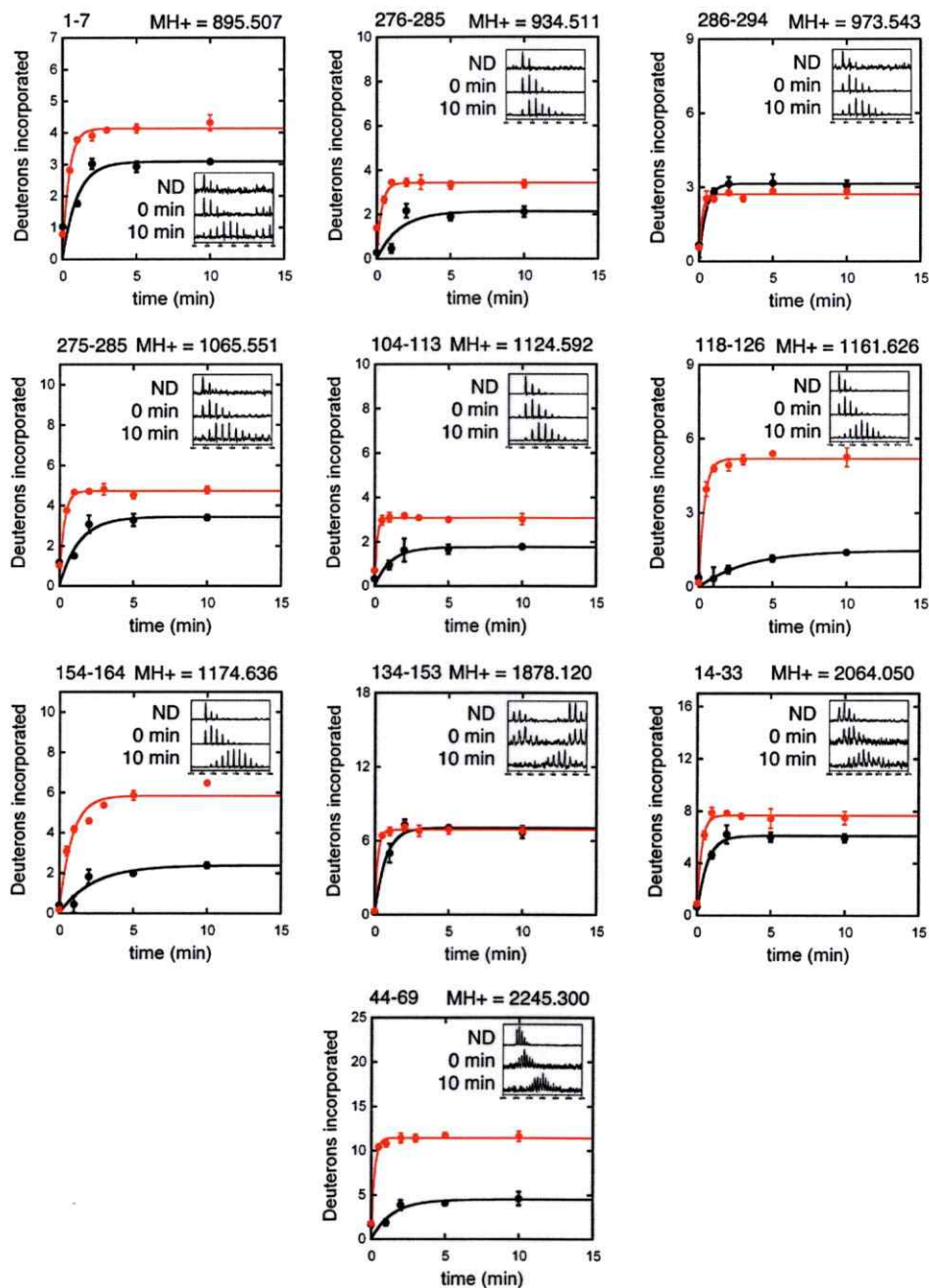
Region de Pfk-2	Masa peptídica MH+ (m/z)	No. amidas	Secuencia del péptido
1-7	895.507	7	MVRIYTL
276-285	934.511	10	VFRFGVAAGSA
286-294	973.543	9	ATLNQGTRL
275-285	1065.551	11	MVFRFGVAAGSA
<b>104-113</b>	<b>1124.592</b>	<b>9</b>	<b>YRFVMPGAAL</b>
118-126	1161.626	9	FRQLEEQL
154-164	1174.636	11	SAAQKQGIRCI
134-153	1878.120	18	LVISGSLPPGVKLEKLTQLI
<b>14-33</b>	<b>2064.050</b>	<b>17</b>	<b>DSATITPQIYPEGKLRCTA</b>
44-69	2445.300	25	VARAIAHLGGSATAIFPAGGATGEHL

\*Los péptidos en negrita pertenecen a la interfaz de asociación entre subunidades.



**Figura 13. Caracterización estructural del intermediario monomérico de Pfk-2 inducido por GndHCl mediante HXMS.**

Comparación de la incorporación de deuterones observada para los distintos péptidos cuantificados mediante MALDI-TOF (Tabla 3) de la enzima Pfk-2 silvestre en condiciones nativas (negro) y la enzima Pfk-2 silvestre incubada por 24 h en 0.85 M GndHCl (azul), ambas a 25 °C. En la parte superior de cada gráfico se indica la masa/carga de cada péptido y los residuos que representa de la estructura de Pfk-2. El máximo del eje  $y$  representa el número de amidas totales para cada péptido y el eje  $x$  representa el tiempo de incubación en  $D_2O$ . El inserto en cada gráfico corresponde a los espectros observados para la muestra no deuterada (ND), y sometida a 0 y 10 min de intercambio en  $D_2O$  para cada péptido de la enzima incubada en GndHCl.



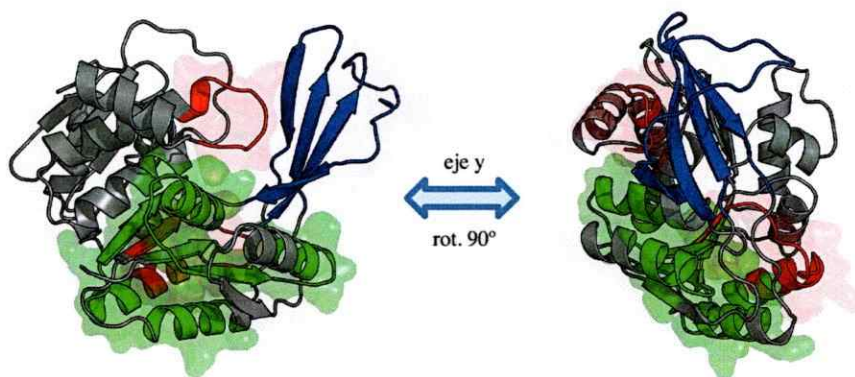
**Figura 14. Caracterización estructural de la mutante L93A de Pfk-2 mediante HXMS.**

Comparación de la incorporación de deuterones observada para los distintos péptidos cuantificados mediante MALDI-TOF (Tabla 3) de la enzima Pfk-2 silvestre en condiciones nativas (negro) y la mutante L93A (rojo) a 25 °C. En la parte superior de cada gráfico se indica la masa/carga de cada péptido y los residuos que representa de la estructura de Pfk-2. El máximo del eje  $y$  representa el número de amidas totales para cada péptido y el eje  $x$  representa el tiempo de incubación en  $D_2O$ . El inserto en cada gráfico corresponde a los espectros observados para la muestra de la mutante L93A no deuterada (ND), y sometida a 0 y 10 min de intercambio en  $D_2O$ .

isotópico con el solvente circundante, con respecto a lo observado para la enzima Pfk-2 silvestre en condiciones nativas. El aumento en el número de deuterones incorporados por péptido en comparación a la enzima silvestre, sugiere que 1-2 amidas se exponen al solvente como producto de la mutación del residuo Leu93. Sin embargo, existen tres regiones en las que se observa un aumento significativo en la incorporación de deuterio en comparación con la enzima silvestre, representadas por los péptidos correspondientes a los residuos 44 – 69 (MH+ = 2445.300), 118 – 126 (MH+ = 1161.626) y 154 – 164 (MH+ = 1174.636), los cuales presentan una diferencia de 4, 4 y 7 amidas disponibles respecto a la enzima silvestre en condiciones nativas, respectivamente. Estas regiones se encuentran confinadas al dominio Rossmann del dominio de mayor tamaño de la enzima Pfk-2, y se conectan al dominio menor a través de la discontinuidad topológica reentrante de la cadena polipeptídica [74, 94]. Esto sugiere fuertemente que los primeros eventos significativos de reconfiguración estructural en la enzima Pfk-2 que acompañan al proceso de disociación ocurren en el dominio Rossmann (Figura 15), sugiriendo fuertemente que la estabilidad del monómero depende de la formación del dímero.

#### 2.3.6. La mutante L93A de Pfk-2 exhibe un desplegamiento no cooperativo en presencia de GndHCl.

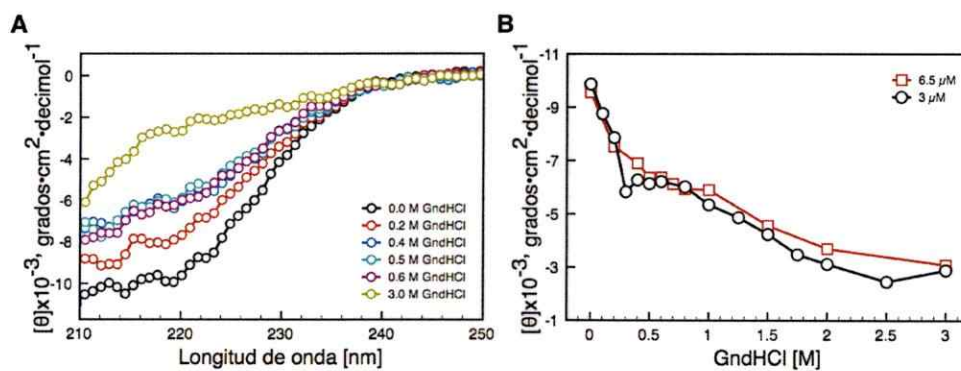
Para determinar si efectivamente la mutante monomérica L93A de la enzima Pfk-2 sigue un mecanismo de desplegamiento no cooperativo, realizamos experimentos de dicroísmo circular en presencia de GndHCl. Para este experimento, se utilizaron dos concentraciones de proteína (3 y 6.5  $\mu$ M), las cuales fueron incubadas por 24 h a distintas concentraciones de GndHCl, entre 0 y 3 M. Como se observa en la figura 16, la mutante monomérica de Pfk-2 exhibe una transición no cooperativa hacia el estado desplegado. Esta transición no cooperativa se observa a dos concentraciones de proteína diferentes (Figura 16B), pese a que la acumulación de los espectros en el rango de 0.5-0.8 M sugieren la posible presencia del intermediario monomérico observado para la enzima silvestre a concentraciones de proteína y agente caotrópico similares [114].



**Figura 15. Mapeo de las regiones afectadas por la disociación inducida por mutagénesis sitio-dirigida del residuo L93 en la estructura de la enzima Pfk-2.**

Se muestra el monómero compacto de la enzima Pfk-2, proveniente de la estructura resuelta por cristalografía de rayos X (PDB 3CQD), en representación de cintas. Las regiones en verde representan a los péptidos que exhiben el mayor aumento en la incorporación de deuterones (residuos 44-69, 118-126 y 154-164), en comparación con la enzima Pfk-2 silvestre en condiciones nativas. En rojo, se indican aquellas regiones sin cambios entre el dímero nativo y el monómero inducido por la mutación L93A (residuos 134-153 y 286-294). En azul se muestran las hebras  $\beta$  del dominio menor que participa de la asociación homodimérica.





**Figura 16. Pérdida no cooperativa de estructura secundaria de la mutante monomérica de Pfk-2 en presencia de GndHCl.**

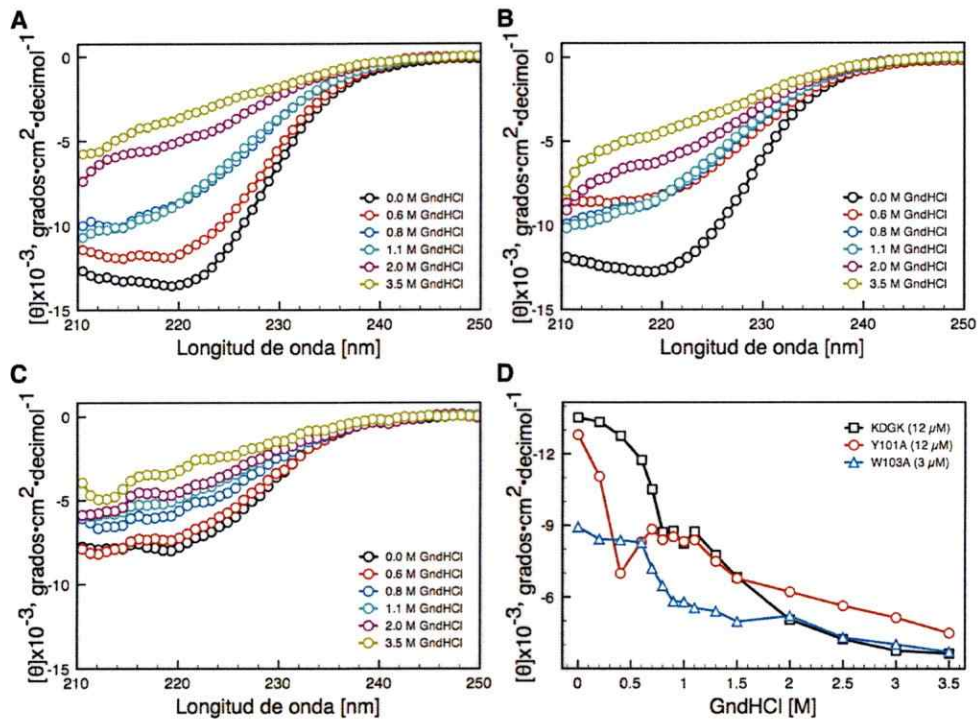
A) Espectros de dicroísmo circular de la mutante L93A de Pfk-2 luego de incubar la proteína por 24 h en solución amortiguadora con diferentes concentraciones de GndHCl, entre 0 (espectro inferior, negro) y 3 M (espectro superior, amarillo). La concentración de proteína utilizada fue de 3  $\mu$ M. La señal dicroica se expresa como elipticidad molar por residuo, y cada espectro fue obtenido por la acumulación de al menos 4 escaneos a la temperatura de trabajo (20  $^{\circ}$ C).

B) Cambios en la elipticidad molar por residuo, a una longitud de onda de 220 nm y una temperatura de trabajo de 20  $^{\circ}$ C, para la mutante L93A de Pfk-2 en función de la concentración de GndHCl. Este experimento fue realizado utilizando concentraciones de proteína de 3  $\mu$ M (negro) y 6.5  $\mu$ M (rojo).

Esta transición no cooperativa es consistente con las teorías acerca del desplegamiento de subunidades cuyos contactos nativos son penetrados por el solvente [204], y con la energía de estabilización del intermediario monomérico de la enzima Pfk-2 silvestre observado cuando la proteína es enfrentada a agentes caotrópicos [114] o baja temperatura [128]. Estos resultados, junto con las observaciones derivadas de los experimentos de HXMS, sugieren fuertemente que la mutante L93A recapitula las características estructurales del intermediario monomérico metastable en ausencia de perturbaciones globales (en otras palabras, en condiciones que favorecen el estado nativo) y que representa los primeros cambios ocurridos luego de la disociación, como consecuencia de la topología discontinua del dominio menor.

#### 2.3.7. Observación de un intermediario monomérico en el desplegamiento químico de la enzima Kdgk.

De acuerdo a los resultados obtenidos mediante ultracentrifugación analítica para la enzima Kdgk silvestre, el aumento de la concentración de GndHCl sobre 1 M conduce a la acumulación de una especie monomérica cuyo radio hidrodinámico es mayor que el esperado para el monómero compacto (Figura 12 y Tabla 2). Para corroborar la detección de este intermediario realizamos mediciones de dicroísmo circular de la enzima Kdgk a 25 °C luego de incubarla en solución amortiguadora a distintas concentraciones de GndHCl (entre 0 y 3.5 M) durante 48 h. Los resultados para la enzima Kdgk silvestre y para la mutante Y101A, las cuales fueron estudiadas a la misma concentración de proteína (12 µM), exhiben la presencia de un intermediario (Figura 17A, 17B y 17D). Este intermediario es detectado a concentraciones de GndHCl en el rango de 0.7 – 1.1 M para la enzima Kdgk silvestre, mientras que en el caso de la mutante Y101A, se alcanza a concentraciones de 0.4 – 1.1 M. La transición desde el estado nativo hasta el estado intermediario para la enzima Kdgk silvestre y para la mutante Y101A es consistente con la disociación de las subunidades del dímero de Kdgk en función de la concentración de GndHCl, seguida mediante ultracentrifugación



**Figura 17. Presencia de un intermediario monomérico en el desplegamiento de la enzima Kdgk.**

Se muestran los espectros de dicroísmo circular a distintas concentraciones de GndHCl, entre 0 (curva inferior, negro) y 3.5 M (curva superior, amarillo), para la enzima Kdgk silvestre (Figura 17A) y las mutantes Y101A (Figura 17B) y W103A (Figura 17C). La enzima fue incubada al menos 48 h previo a las mediciones, las cuales se realizaron a 25 °C. La señal dicroica se expresa como elipticidad molar por residuo, y cada espectro fue obtenido por la acumulación de al menos 4 escaneos. Para la enzima silvestre (Figura 17A) y la mutante Y101A (Figura 17B) se utilizó una concentración de proteína de 12  $\mu\text{M}$ , mientras que para la mutante W103A (Figura 17C) se utilizó una concentración de proteína de 3  $\mu\text{M}$ . El cambio en la elipticidad molar por residuo a una longitud de onda de 220 nm en función de la concentración de GndHCl para la enzima Kdgk y las mutantes se muestra en la figura D.

analítica (Figura 12). Asimismo, la señal dicroica de la mutante W103A en condiciones nativas a una concentración de proteína de 3  $\mu\text{M}$  sugiere que el contenido de estructura secundaria de esta mutante es similar al observado para el intermediario monomérico inducido por la adición de GndHCl (Figura 17C) y que la adición de GndHCl lleva al desplegamiento del monómero sin evidencias de un intermediario durante este proceso (Figura 17D). Además, la transición hacia el estado desplegado comienza a concentraciones menores de GndHCl (0.6 M), lo que sugiere que el remplazo de el residuo de triptófano 103 por alanina puede tener incidencias no solo en la estabilidad del homodímero, sino que también en la estabilidad de la subunidad aislada.

De esta manera, se corrobora que la disociación de las subunidades también tiene efectos sobre el contenido de estructura secundaria y, por ende, sobre la estabilidad de la especie monomérica aislada, de manera similar a lo observado para la enzima Pfk-2 de *Escherichia coli* mediante HXMS (Figura 13 y 14) y dicroísmo circular (Figura 16).

#### 2.3.8. Confirmación del estado parcialmente desplegado de la mutante W103A de Kdgk mediante difracción de rayos X de ángulo pequeño.

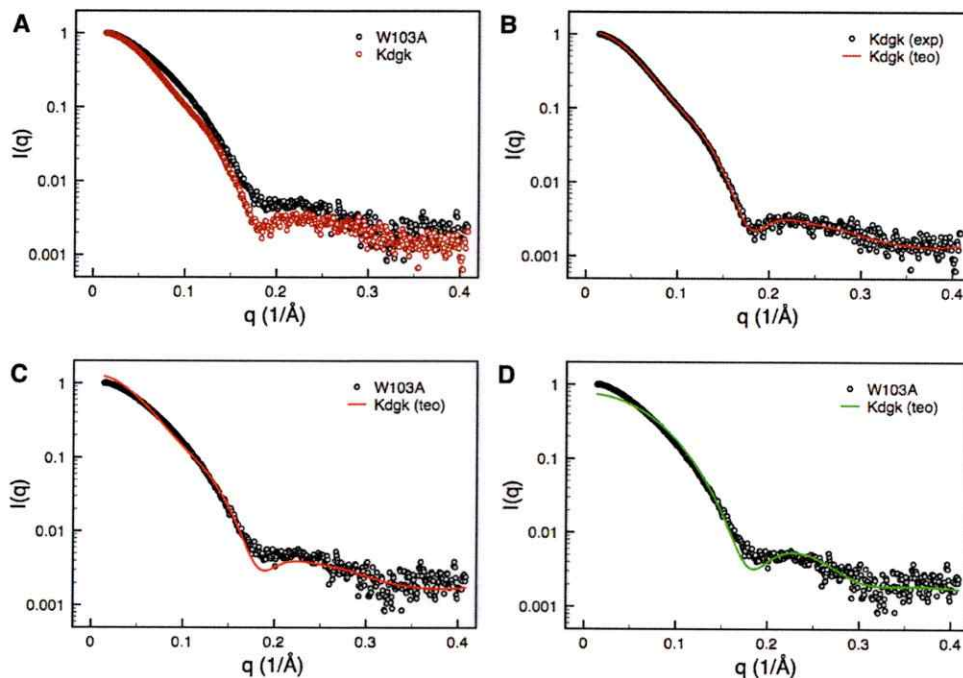
Para confirmar que la especie monomérica de Kdgk exhibe propiedades hidrodinámicas y estructurales características de una especie parcialmente desplegada, realizamos experimentos de difracción de rayos X de ángulo pequeño, técnica que permite determinar de manera gruesa la conformación macromolecular en solución. Estos experimentos fueron realizados utilizando la mutante W103A de Kdgk como representante de la especie monomérica aislada, y comparamos los resultados obtenidos con la enzima Kdgk silvestre en condiciones nativas como control. Para ambas muestras, los experimentos fueron realizados a temperatura ambiente y la concentración de proteína utilizada fue de  $\sim 2$  mg/mL. En estas condiciones, tanto la enzima silvestre como la mutante W103A permanecen como especies monodispersas y representativas de las especies dimericas y monoméricas, de acuerdo a lo observado mediante análisis de

dispersión dinámica de luz a distintas concentraciones de proteína entre 1 y 4 mg/mL (datos no mostrados).

La forma de la curva de difracción de la enzima Kdgk silvestre y de la mutante monomérica W103A son visiblemente diferentes para los ángulos en el rango de  $0.05 \text{ \AA}^{-1}$  a  $0.25 \text{ \AA}^{-1}$  (Figura 18A); para estos ángulos, la intensidad observada para la especie monomérica es levemente mayor que para la especie dimérica.

Para corroborar que la especie dimérica observada en solución es similar a la resuelta cristalográficamente, determinamos la forma teórica de la curva de intensidad en función del ángulo de difracción a partir de la estructura cristalográfica de la enzima Kdgk de *Escherichia coli* mediante el programa CRY SOL [198]. La curva teórica de intensidades se ajusta a los datos experimentales ( $\chi = 2.919$ ), por lo que la especie en solución es estructuralmente similar al dímero resuelto mediante cristalografía de rayos X (Figura 18B). Al repetir esta operación sobre la curva de intensidades experimental obtenida para la mutante W103A, vemos que el ajuste es peor que en el caso de la especie dimérica ( $\chi = 7.438$ ), en buen acuerdo con la diferencia en estructura cuaternaria entre el homodímero de la enzima silvestre y la especie en solución de esta mutante (Figura 18C). Finalmente, para demostrar que la especie monomérica en solución no corresponde a lo esperado para un monómero compacto, se repitió el procedimiento de ajuste, esta vez utilizando una curva de intensidades teórica calculada a partir de una de las subunidades del homodímero resuelto mediante cristalografía de rayos X. En este caso, nuevamente se observa un mal ajuste entre las intensidades teóricas y experimentales ( $\chi = 7.127$ ), aunque levemente mejor que el determinado a partir de la curva de intensidades teórica para el homodímero (Figura 18D). Estos resultados sugieren fuertemente que la enzima Kdgk obtenida mediante mutagénesis sitio-dirigida es una especie monomérica parcialmente desplegada en solución.

A partir de estos datos de difracción en ángulo pequeño, se estimó el radio de giro de la especie en solución a partir de la aproximación lineal de Guinier, en un rango



**Figura 18. Difracción de rayos X de ángulo pequeño para las especies dimericas y monoméricas de Kdgk.**

A) Curvas de intensidad en función del ángulo de difracción para la enzima homodimérica silvestre Kdgk (círculos rojos) y la mutante monomérica W103A (círculos negros). Los datos están normalizados respecto al máximo de intensidad para efectos comparativos.

B) Ajuste de la curva de intensidad experimental de la enzima Kdgk silvestre (círculos negros) versus la curva de intensidad calculada a partir de la estructura resuelta del homodímero de Kdgk de *Escherichia coli* (línea roja).

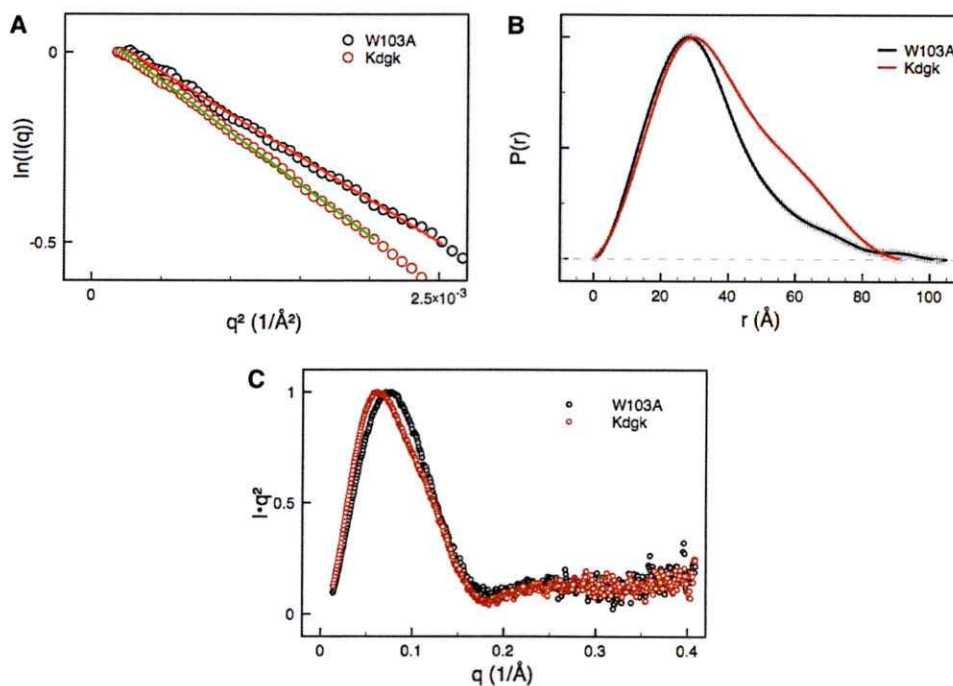
C) Ajuste de la curva de intensidad experimental de la mutante monomérica W103A (círculos negros) versus la curva de intensidad calculada a partir de la estructura resuelta del homodímero de Kdgk de *Escherichia coli* (línea roja).

D) Ajuste de la curva de intensidad experimental de la mutante monomérica W103A (círculos negros) versus la curva de intensidad calculada a partir de uno de los monómeros aislados extraído a partir de la estructura resuelta de Kdgk de *Escherichia coli* (línea verde).

de datos tal que  $q \cdot R_g < 1.3$ . A partir de este análisis, se determinó que el radio de giro de la especie dimerica es  $2.850 \pm 0.005$  nm, mientras que el de la especie monomérica, representada en este caso por la mutante W103A, es de  $2.590 \pm 0.007$  nm (Figura 19A). También calculamos el radio de giro a partir de la función de distribución de distancias, la cual fue determinada utilizando los datos de difracción y el programa GNOM [194]. El radio de giro calculado a partir de este análisis es de  $2.673 \pm 0.009$  nm para el monómero aislado, y de  $2.948 \pm 0.001$  nm para la enzima silvestre (Figura 19B).

Para corroborar que la especie monomérica es relativamente compacta y que posee desplegamiento parcial local y no global, confeccionamos un gráfico de Kratky a partir de las curvas de intensidad de difracción. Este tipo de gráfico permite analizar cualitativamente el grado de flexibilidad de una macromolécula o el "grado de plegamiento" [205]. Para proteínas flexibles,  $I \cdot q^2$  define una primera fase ascendente con un pico parabólico en función de  $q$ , y luego se define una línea base elevada, siendo el caso límite (conformaciones polipeptídicas estocásticas) cuando los valores de  $I \cdot q^2$  no descienden a medida que  $q$  aumenta [205]. Como se observa en la figura 19C, tanto para la enzima dimerica silvestre como para la mutante monomérica, el gráfico de Kratky exhibe una curva característica de proteínas globulares, sin embargo el área invariante de Porod definida para  $0 < q < 0.18 \text{ \AA}^{-1}$  para la mutante W103A es levemente mayor que para la enzima silvestre. Estos resultados refuerzan la idea que los cambios inducidos en la estabilidad estructural de la subunidad aislada producto de la perturbación local de la interfaz de asociación entre subunidades, son locales.

Finalmente, utilizamos los datos de difracción y la función de distribución de distancias  $P(r)$  derivada a partir de éstos para generar modelos *ab initio* de la especie dimerica y de la especie monomérica mediante los programas DAMMIN [195] y GASBOR [196]. Como se observa en la figura 20, el modelo *ab initio* del homodímero de Kdgc es similar, en forma y tamaño, al observado en la estructura cristalográfica. En efecto, la superposición del mejor modelo obtenido para la enzima Kdgc en solución con la estructura resuelta por cristalografía de rayos X mediante el programa SUPCOMB



**Figura 19. Propiedades hidrodinámicas de la enzima Kdgk y la mutante W103A en solución.**

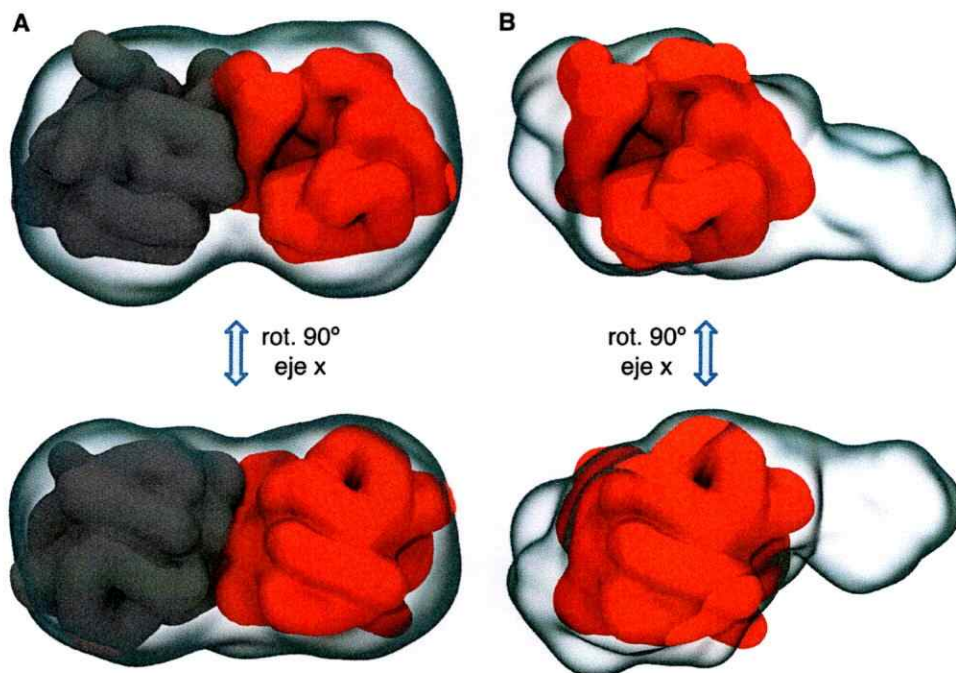
A) Ajuste de Guinier para la región lineal ( $q^2 < 2.5 \times 10^{-3}$ ) de la curva de intensidades de la enzima Kdgk silvestre (círculos negros) y de la mutante W103A (círculos rojos) a concentración de  $\sim 2$  mg/mL. El ajuste lineal se muestra como una línea continua en rojo para la enzima Kdgk y en verde para la mutante W103A. El radio de giro calculado a partir del ajuste de Guinier es de  $2.850 \pm 0.005$  nm para el dímero y de  $2.590 \pm 0.007$  nm para el monómero.

B) Función de distribución de distancias para la enzima Kdgk silvestre (rojo) y la mutante W103A (negro). Las pequeñas líneas verticales en cada curva corresponden al error asociado. El radio de giro calculado a partir de este análisis es de  $2.673 \pm 0.009$  nm para el monómero aislado, y de  $2.948 \pm 0.001$  nm para la enzima silvestre.

C) Gráfico de Kratky para la enzima Kdgk silvestre (círculos rojos) y la mutante W103A (círculos negros).







**Figura 20. Determinación *ab initio* de la forma del homodímero y monómero aislado de la enzima Kdgk de *Escherichia coli*.**

A) Modelo *ab initio* de la forma de la enzima homodimérica Kdgk de *Escherichia coli* y superposición con la estructura cristalográfica. En rojo y gris se muestran, en representación de superficie, la cadena principal de las dos subunidades de la enzima Kdgk de acuerdo a la estructura cristalográfica. El contorno que rodea a la estructura cristalográfica corresponde a la representación de superficie del mejor modelo *ab initio* (de entre diez modelos totales) de la enzima Kdgk en solución. El modelo fue generado a partir de los datos de difracción de rayos X de ángulo pequeño mediante el programa DAMMIN [195].

B) Modelo *ab initio* de la forma de la mutante monomérica W103A de la enzima Kdgk de *Escherichia coli*, superpuesta a una de las dos subunidades de la estructura cristalográfica del homodímero. Las representaciones de superficie son equivalentes a las de la figura 20A. El modelo *ab initio* escogido es el mejor de cinco modelos generados a partir de los datos de difracción de rayos X de ángulo pequeño mediante el programa GASBOR [196].

[206] es buena (discrepancia espacial normalizada (NSD) = 0.8172). En el caso de la mutante monomérica, el modelo *ab initio* corresponde a una estructura relativamente compacta, pero cuyo grado de compactación es menor al observado para cada subunidad en el contexto dimérico (Figura 20), sin embargo sus características de forma sugieren la búsqueda de otras estrategias con la incorporación de regiones flexibles para generar un modelo más realístico de la subunidad aislada de Kdgc [207].

## 2.4. DISCUSION

El fuerte acoplamiento entre los procesos de disociación y desplegamiento de la enzima homodimérica Pfk-2 de *Escherichia coli* conduce a la formación de un intermediario monomérico parcialmente desplegado, en el cual la accesibilidad al solvente y el radio hidrodinámico es mayor que en el caso de un monómero compacto, pero menor que el esperado para una cadena polipeptídica completamente extendida [148]. Este intermediario monomérico, observado tanto en el mecanismo de (des)plegamiento de la enzima Pfk-2 a concentraciones de GndHCl de 0.5 – 0.9 M [114] como en el estado final de la desnaturación en frío [128], presenta diferencias en cuanto a la energía libre de estabilización y exposición de área al solvente (valor  $m$  [150]) respecto al estado desplegado que, si bien pequeñas (~1 kcal/mol y 1.3 kcal/mol/K), apuntan a posibles diferencias en cuanto a las características estructurales de los intermediarios obtenidos cuando esta enzima es enfrentada a distintas perturbaciones.

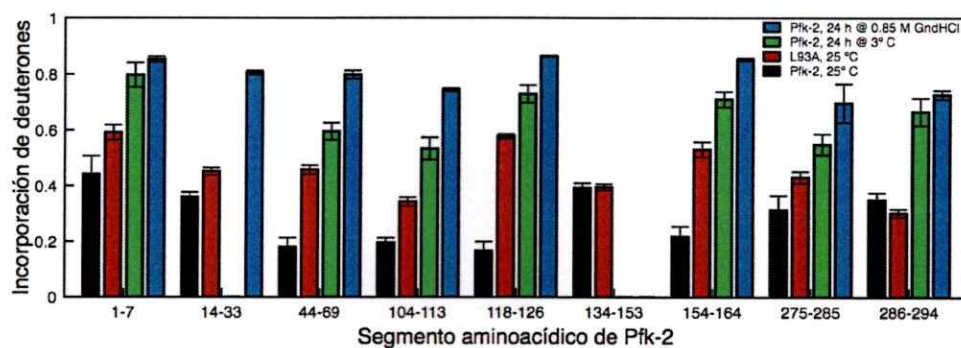
De acuerdo a los experimentos de HXMS sobre la enzima Pfk-2 silvestre sometida a incubación por 24 h en 0.85 M de GndHCl, el intermediario monomérico obtenido mediante agentes caotrópicos exhibe una mayor accesibilidad al solvente que el estado final del proceso de desplegamiento en frío (Figura 13 y 21), confirmando que estos intermediarios son semejantes, pero no equivalentes. Otra evidencia experimental al respecto se encuentra en las diferencias en la intensidad de la señal de fluorescencia de la sonda de cavidades hidrofóbicas ANS para cada uno de estos intermediarios [128], la cual es mayor para la enzima desnaturada en frío y, por ende, sugiere un mayor grado

de compactación. Una explicación para estas diferencias en compactación puede encontrarse en las diferencias de viscosidad entre la solución amortiguadora a diferentes temperaturas: la viscosidad de la solución amortiguadora utilizada a 25 °C es de 0.91 cP, y aumenta a 1.65 cP cuando la temperatura es de 3 °C; en presencia de 0.85 M GndHCl, la viscosidad de la solución en la cual la enzima Pfk-2 es incubada es de 0.93 cP. En otras proteínas como apoflavodoxina o la proteína ribosomal pequeña S16, el uso de agentes de *crowding* (elevada concentración de macromoléculas en solución [208]) como dextrano, el cual provoca cambios en la viscosidad y en el volumen excluido del solvente, causa la compactación de los estados desnaturalados en presencia de GndHCl y el aumento del valor  $m$  [150] y de la energía libre de estabilización [209, 210]. Otra explicación está relacionada con los cambios inducidos por la temperatura en la energía de interacción entre las moléculas de agua y entre los residuos no polares. A medida que la temperatura disminuye, los contactos hidrofóbicos son desestabilizados y la energía de los puentes de hidrógeno entre las moléculas de agua es más favorable, generando arreglos ordenados en capas de moléculas de agua alrededor de los residuos que forman el centro hidrofóbico de la proteína, los cuales son de baja energía, por lo que permiten la existencia de los contactos hidrofóbicos en configuraciones separadas por el solvente [132, 149, 211]. En el caso de GndHCl, el grupo guanidinio induce una sobre-reestructuración de la red de interacciones del agua para poder solvatar a este grupo cargado, lo que reduce la compensación entrópica que permite el colapso hidrofóbico y permite, en consecuencia, la solvatación de los residuos no polares [212]. Sin embargo, la hidratación del grupo guanidino es débil, por lo que este efecto es secundario y la mayor contribución al efecto desnaturalante es la interacción directa de éste con la proteína [213].

Estas diferencias en las propiedades hidrodinámicas y estructurales de los estados desnaturalados inducidos mediante agentes caotrópicos y exposición a bajas temperaturas ya han sido previamente determinadas en otras proteínas, como el dominio carboxilo-terminal de la proteína ribosomal L9 [136] y el inhibidor de subtilisina [163].

Pese a estas leves diferencias estructurales, los intermediarios observados se asemejan en cuanto a las consecuencias de la perturbación global en la desestabilización de la proteína completa.

Por otra parte, la perturbación local de la interfaz de asociación mediante la introducción de mutaciones puntuales, en base al análisis *in silico* de los residuos determinantes de la estabilidad del homodímero (Figura 8), y la utilización de ultracentrifugación analítica, HXMS y dicroísmo circular, revelaron que el monómero aislado, en condiciones que favorecen el estado nativo de la enzima Pfk-2 silvestre, es relativamente compacto en comparación a los estados intermediarios inducidos mediante temperatura o presencia de agentes desnaturantes (Tabla 2), solo presenta cambios importantes en la flexibilidad estructural del dominio Rossmann (Figura 14 y 15) y se despliega de manera no cooperativa (Figura 16). Estas diferencias en compactación producto de perturbaciones locales de la interfaz de asociación ya han sido demostrados para otros modelos proteicos, como el dominio Apple 4 (A4) del factor de coagulación XI [166] y la quinasa dimérica de creatina de músculo de conejo (MM-CK) [169]. Considerando que la topología reentrante solo conecta al dominio Rossmann con el dominio menor, mientras que el dominio  $\beta$ -meandro carboxilo-terminal no está involucrado en este cruce de la cadena polipeptídica entre dominios [74, 94], nuestros resultados sugieren fuertemente que la especie monómerica obtenida a partir de la mutante L93A representa los primeros cambios estructurales que ocurren en la subunidad aislada de la enzima Pfk-2 como consecuencia de la disociación y de la naturaleza discontinua de la topología del dominio menor. De este modo, la transición no cooperativa apreciada durante el desplegamiento inducido por GndHCl de la mutante L93A correspondería al incremento en el radio de hidratación del monómero aislado (el cual tiene una energía de estabilización marginal, dado que el paso de disociación de subunidades del homodímero de Pfk-2 representa la mayor parte de la energía de estabilización del estado nativo [114, 128]), producto de la penetración del solvente en



**Figura 21. Diferencias en la accesibilidad al solvente del dímero nativo de Pfk-2 y del monómero aislado obtenido mediante diversas perturbaciones globales y locales.**

La incorporación de deuterones se indica en el eje  $y$  como el máximo número de deuterones incorporados respecto al total de amidas disponibles para el péptido correspondiente. Las regiones de la proteína (posiciones aminoacídicas de la secuencia primaria) representadas por cada péptido se indican en el eje  $x$ . En negro, se muestra el cambio en accesibilidad al solvente del dímero nativo de Pfk-2; en rojo, la mutante monomérica L93A; en verde, el estado desnaturado por exposición a 3 °C; en azul, el estado intermediario monomérico inducido por GndHCl. Para cada barra, se indica el error asociado al máximo de incorporación calculado, obtenido a partir del ajuste de los datos de intercambio para distintos tiempos de intercambio en D<sub>2</sub>O para cada péptido a una exponencial simple, como en las figuras 13 y 14.

otras regiones de la estructura de la proteína que no están acopladas al dominio menor a través de la topología reentrante de la cadena polipeptídica (Figura 14). De esta manera, los estados inducidos a través de las distintas perturbaciones aplicadas representan distintos estados del monómero aislado con distintos grados de compactación e hidratación de acuerdo a las características del solvente circundante (Figura 21).

La posición de los residuos claves para la estabilidad del homodímero no es azarosa; al contrario, aquellos residuos responsables del mayor cambio en la energía libre de asociación *in silico* se localizan en la última hebra  $\beta$  (dirección amino a carboxilo-terminal) del dominio menor (Figura 8). Esta localización de los *hot-spots* de la interacción proteína-proteína ocurre tanto para Pfk-2 como para Kdgk, pese a que sus topologías de asociación son diferentes [94, 110]. Para enzimas de la familia riboquinasa cuya topología de interfaz de asociación es semejante a Pfk-2, la última hebra  $\beta$  participa de la interacción directa con la hebra proveniente de la subunidad adyacente para conformar el barril  $\beta$  aplanado [88, 94]. Para enzimas similares a Kdgk, la topología de la interfaz entre subunidades exhibe una conformación estructural similar a la de una sábana  $\beta$  antiparalela pronunciadamente curvada [110, 112]. En este caso, los representantes de enzimas Kdgk estudiados exhiben una alta conservación del residuo aromático que constituye el *hot-spot* indicado para *Shigella flexineri* (Trp103, Figura 9), el cual puede ser remplazado por tirosina en los miembros termófilos (datos no mostrados). Por consiguiente, la conservación estructural de los *hot-spots* de estabilidad de los homodímeros de la familia riboquinasa están fuertemente ligada a la topología de la interfaz de asociación. Esta idea es reforzada por la disminución de la frustración mutacional y configuracional del dominio menor en estas regiones producto de la interacción entre subunidades (Figura 10), lo que sugiere que tanto la identidad de estos residuos como la posición que ocupan en el dominio menor son una consecuencia de la topología de la interfaz de asociación proteína-proteína.

Pese a las diferencias topológicas, nuestros experimentos de desplegamiento por adición de GndHCl y de disociación inducida mediante mutaciones puntuales sobre las

enzimas Pfk-2 y Kdgk de *Escherichia coli* sugieren que las consecuencias de la disociación del homodímero sobre la estabilidad del monómero son equivalentes: el alcance de una subunidad aislada parcialmente desplegada, cuyas propiedades dependen del tipo de perturbación utilizada. Esta afirmación es corroborada a través de la combinación de diversas estrategias experimentales que permiten el análisis de los cambios en las propiedades hidrodinámicas, en el estado de agregación y en el contenido de estructura secundaria de los distintos modelos de estudio al ser enfrentados a los distintos tipos de perturbación.

## 2.5. AGRADECIMIENTOS

El autor de este trabajo doctoral agradece a la Dra. Elizabeth A. Komives, Universidad de California, San Diego, Estados Unidos, por su invaluable aporte para la estandarización de la cristalización de la enzima Kdgk; al Dr. Jon Fitchett, de Lilly Biotechnology Center, San Diego, California, Estados Unidos, por facilitar el acceso al espectrómetro de masas Synapt G2; al Dr. Richard Garratt, Universidad de Sao Paulo, Brasil, por facilitar el acceso a la instrumentación del Instituto de Física de Sao Carlos, Universidad de Sao Paulo, Brasil; al Dr. Humberto Pereira, Universidad de Sao Paulo, Brasil, y Víctor Castro, Universidad de Chile, por la colección de los datos de difracción de los cristales de la enzima Kdgk de *Escherichia coli* en el Instituto de Física de Sao Carlos; al Dr. Ricardo Aparicio, Universidad de Campinas, por su asesoramiento para la evaluación de los datos de difracción de rayos X de ángulo pequeño; y a Pablo Villalobos y Ricardo Zamora, Universidad de Chile, por su contribución en las etapas iniciales del trabajo con la mutante L93A de Pfk-2 y la enzima Kdgk, respectivamente. Los experimentos de ultracentrifugación analítica fueron realizados en el Centro de Biofísica, Departamento de Química y Bioquímica, Universidad de California, San Diego. Los experimentos de dicroísmo circular fueron realizados en el Instituto de Física de Sao Carlos, Universidad de Sao Paulo, Brasil. La resolución preliminar de la estructura cristalográfica de la enzima Kdgk de *Escherichia coli* fue realizada durante el

curso *Macromolecular Crystallography School 2013*", impartido en el Institut Pasteur de Montevideo, Uruguay, y coordinado por el Dr. Alejandro Buschiazzo. Los experimentos de difracción de rayos X de ángulo pequeño fueron realizados durante el curso "*SAXS Workbench*", impartido en la línea SAXS1 del Laboratorio Nacional de Luz Sincrotrón, Campinas, Brasil, y coordinado por los Drs. Mateus Cardoso y Vesna Stanic. Este trabajo recibió financiamiento del Fondo Nacional de Desarrollo Científico y Tecnológico (FONDECYT) No. 1110137. CAR-S fue beneficiario de una beca de estudios doctorales de la Comisión Nacional de Investigación Científica y Tecnológica (CONICYT), de una beca de asistencia del Institut Pasteur de Montevideo, Uruguay, y de una beca de asistencia del Laboratorio Nacional de Luz Sincrotrón.



CAPITULO TRES:  
Estudios *in silico* del plegamiento de  
topologías complejas y efecto de la  
penetración del solvente.

### 3.1. INTRODUCCION

Los resultados expuestos en este trabajo sugieren que el acoplamiento de los procesos de disociación y desplegamiento de las enzimas Pfk-2 y Kdgk de *Escherichia coli* pueden ser explicados en función de la conectividad reentrante de la cadena polipeptídica entre dominios y de la topología de plegamiento del homodímero.

Dado que un sinnúmero de evidencia experimental ha demostrado que la topología es el factor determinante del mecanismo de (des)plegamiento de diversas proteínas monoméricas [27, 28, 29, 30, 31, 32, 33, 34, 35] y oligoméricas [46, 47, 48], y que el mapa de contactos nativos es un buen descriptor de la topología del estado nativo [214], incluyendo implícitamente información acerca del orden de contacto de cada proteína de interés [21], se han desarrollado diversos modelos simplificados basados en la estructura nativa [43, 215] para predecir y/o determinar con éxito la termodinámica y cinética del proceso de (des)plegamiento, y los cambios estructurales que ocurren durante estos procesos, para una gran variedad de modelos de estudio con distintos estados de agregación [43, 45, 50, 51, 55]. En base a esto, es esperable que las características predominantes del (des)plegamiento de las enzimas homodiméricas de la familia riboquinasa puedan ser también predichas utilizando estas funciones de energía potencial que incluyen implícitamente un sesgo hacia el estado nativo (o nativo-céntricos).

Sin embargo, el incremento en el orden de contacto y en el largo de la cadena polipeptídica también involucra un enlentecimiento de las constantes cinéticas de plegamiento [21]; es decir, topologías complejas tardan más tiempo en alcanzar el estado nativo. Este fenómeno también es detectable en las simulaciones con modelos basados en la estructura nativa [21]. En efecto, los parámetros cinéticos y termodinámicos escalan de manera consistente en estos modelos, ya que la velocidad de plegamiento depende del número de contactos nativos [24], de manera que la altura de la(s) barrera(s) energética(s) que separa(n) al estado desnaturado del estado nativo

está(n) estrechamente relacionada(s) con la velocidad de plegamiento [216]. Por consiguiente, para proteínas con topologías complejas compuestas por un gran número de residuos, como en el caso de las enzimas homodiméricas de la familia riboquinasa, las transiciones entre el estado desnaturalado y el estado nativo son difíciles de capturar incluso utilizando estos modelos, en los cuales la escala de tiempo es  $10^4$  veces mayor que en simulaciones moleculares con campos de fuerza empíricos [217, 218].

Otro aspecto importante es que nuestros experimentos de HXMS con la enzima Pfk-2 sugieren la presencia de un intermediario penetrado por solvente [148], un evento que no es posible capturar en los modelos nativo-céntricos actuales, ya que carecen de solvente circundante. Por otro lado, la energía de interacción entre pares de átomos o residuos que forman parte de un contacto nativo están descritas mediante un potencial de Lennard-Jones cuyo mínimo energético está centrado en la distancia entre estos átomos o residuos en el estado nativo [43], y si bien varios mínimos energéticos a distintas distancias de interacción en el estado nativo pueden introducirse mediante potenciales basados en múltiples curvas de Gauss [219], se requiere de conocimiento previo de los contactos que aceptan más de una configuración de distancia en el estado nativo, mediante la disponibilidad de varias estructuras cristalográficas para una misma proteína, las cuales describen conformaciones diferentes.

Sin embargo, recientemente se ha descrito que los mecanismos de desplegamiento de proteínas por efecto de la presión y de la baja temperatura son microscópicamente similares, involucrando la formación de una capa de solvatación del grosor de una molécula de agua, la cual se forma alrededor de los residuos no polares en desmedro de los contactos hidrofóbicos directos, a medida que la temperatura desciende o la presión aumenta [220]. Potenciales que admiten como parte del paisaje energético, la formación de configuraciones de interacción directa y mediadas por una molécula de agua entre residuos en contacto en el estado nativo, separadas energéticamente por una barrera de desolvatación, han sido descritas previamente [221], evidenciando la importancia de la hidratación como un "lubricante" que permite que el centro

hidrofóbico encuentre su estado empacado óptimo, previniendo la formación de contactos no-nativos [222]. De este modo, se propone la manipulación directa de la energía de los contactos directos y solvatados de este potencial [221], para favorecer la formación de contactos nativos mediados por agua, como una estrategia para acceder a las conformaciones solvatadas observadas experimentalmente como consecuencia de la aplicación de presión o de la exposición de proteínas a baja temperatura [220].

Por estos motivos, hemos introducido dos modificaciones a los modelos simplificados nativo-céntricos convencionales, las cuales son evaluadas en distintos modelos proteicos. Estas modificaciones corresponden a: i) una estrategia de muestreo acelerado de las transiciones entre el estado plegado y desplegado; ii) un potencial modificado que admite configuraciones de contactos nativos mediados por moléculas de agua [221]. Mientras el primero tiene como objetivo optimizar las estrategias de muestreo y análisis del proceso de plegamiento de proteínas de gran tamaño y alta complejidad topológica, el segundo pretende la incorporación del fenómeno de penetración del solvente al centro hidrofóbico de proteínas.

Nuestros resultados muestran que las características del estado de transición y de la altura de la barrera energética no se ven modificadas por la estrategia de muestreo acelerado, mientras que la introducción de barreras de solvatación permite el alcance de configuraciones hidratadas en proteínas.

## **3.2. MATERIALES Y METODOS.**

### **3.2.1. Modelos de estudio**

Las proteínas escogidas como modelo de estudio para la aplicación de estos métodos de simulación corresponden a: inhibidor de quimotripsina CI2 de *Hordeum vulgare* (PDB ID 1YPA) [223], el dominio SH3 de la quinasa humana de tirosina (PDB ID 1FMK) [224], ubiquitina humana (PDB ID 1UBQ) [225] y el dominio B1 de la

proteína G de *Streptococcus* (PDB ID 2GB1) [226]. Además, realizamos simulaciones de plegamiento sobre la quinasa de piridoxal humana (PDB ID 2F7K) [227] (Figura 22).

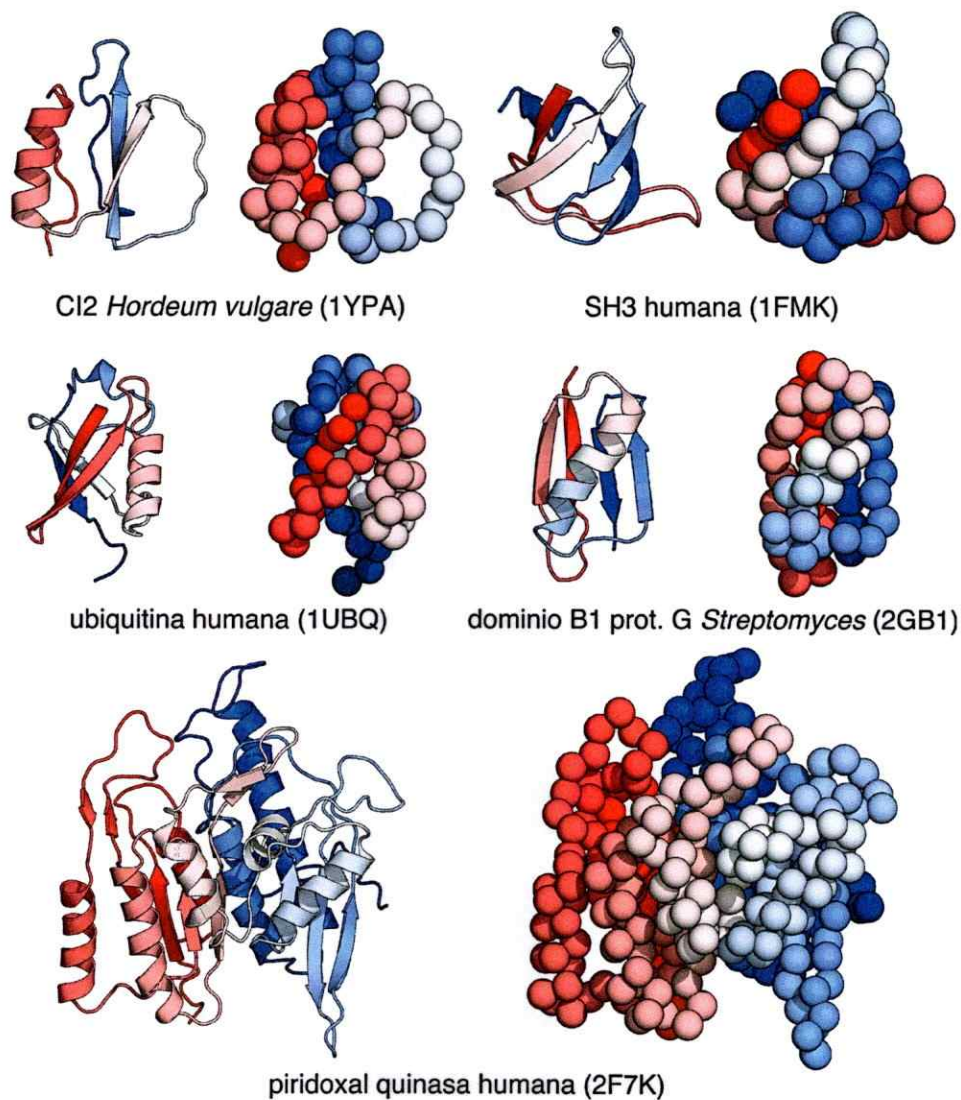
### 3.2.2. Modelo simplificado basado en estructura nativa.

Para las simulaciones de tipo Gō [42], utilizamos un modelo simplificado basado en la estructura nativa equivalente al descrito en el trabajo de Clementi y cols [43]. En este modelo, cada residuo es representado como una esfera de masa unitaria centrada en las coordenadas del carbono alfa del residuo correspondiente ( $C\alpha$ ), y la estructura nativa es definida como el mínimo energético global. Dado que este modelo (en adelante, modelo  $C\alpha$ ) se basa en el principio de la frustración mínima [12], la frustración energética es descartada, y solo las restricciones geométricas impuestas por la topología del estado nativo son consideradas [45].

En este modelo, la distancia de los enlaces y la rotación de los ángulos y ángulos dihedros entre residuos es controlada a través de potenciales harmónicos. Los pares de residuos separados por al menos 3 residuos ( $i > j + 3$ ), que establecen al menos un contacto átomo-átomo en el estado nativo a una distancia límite de 6 Å, son evaluados mediante un potencial de Lennard-Jones tipo 12-10 cuyo mínimo energético corresponde a la distancia calculada a partir de la estructura cristalográfica, mientras que otras interacciones no locales son consideradas como repulsivas. Todos los parámetros de los componentes atractivos del potencial son extraídos a partir del estado nativo. De este modo, la conectividad del polímero es mantenida mediante enlaces y ángulos, mientras que la geometría del estado nativo es preservada a través de ángulos dihedros y potenciales no locales. La forma funcional detallada de este potencial energético se encuentra descrita previamente [43]. El cálculo del mapa de contactos nativos y la generación del modelo  $C\alpha$  fue realizada mediante el servidor SMOG [228].

### 3.2.2. Simulaciones moleculares utilizando modelos simplificados.

Las simulaciones de estas proteínas fueron realizadas utilizando el programa de acceso libre GROMACS [229], utilizando unidades reducidas. El paso de tiempo  $\tau$



**Figura 22. Modelos de estudio para la inclusión de potenciales adicionales en modelos nativo-céntricos.**

Se muestra la estructura secundaria y terciaria de cada una de las proteínas estudiadas en representación de cintas, y la representación equivalente de esferas sobre el carbono alfa de cada residuo, correspondiente al modelo C $\alpha$ . El gradiente de color representa la posición en la secuencia primaria, el cual va desde el amino-terminal, en rojo, al carboxilo-terminal, en azul. El nombre y fuente de cada proteína se indica en la parte inferior de cada estructura, y el código PDB se indica en paréntesis.

utilizado fue de 0.0005, equivalente a  $\sim 2.5$  ps [217, 218]. La temperatura fue controlada mediante dinámica estocástica con una constante de acoplamiento de 1. Para cada proteína se realizaron varias trayectorias de simulación a temperatura constante. En el caso de las simulaciones con muestreo acelerado (ver sección 3.2.3), las simulaciones fueron realizadas a la temperatura de plegamiento. En el caso de las simulaciones que incorporan un potencial de solvatación (ver sección 3.2.4), se realizaron varias simulaciones a distintas temperaturas en paralelo, y configuraciones representativas de cada temperatura fueron intercambiadas entre trayectorias de temperaturas vecinas por el método de intercambio de réplicas (*Replica Exchange*, [230]). El tiempo escogido para realizar el intercambio entre réplicas fue de  $5000 \tau$ .

### 3.2.3. Potencial polarizante externo para muestreo acelerado de transiciones de plegamiento en modelos simplificados.

Para incrementar el muestreo de las transiciones entre el estado plegado, el estado desplegado y los posibles estados intermediarios de una proteína objetivo, se introdujo un potencial externo (*umbrella sampling*, [231]) basado en el número de contactos nativos ( $Q$ ) del sistema. La forma funcional de este potencial es:

$$V_C = \frac{1}{2} k (V_Q - V_{Q^*})^2$$

Que corresponde a un potencial armónico, cuya fuerza es controlada por el parámetro  $k$ , centrado en valores particulares del potencial de conteo de contactos nativos  $V_Q$  (llamado  $V_{Q^*}$ ). El potencial de conteo de contactos nativos ( $V_Q$ ) corresponde a una descripción continua de la función escalonada de conteo de contactos nativos:

$$V_Q = \sum_{ij}^Q \frac{1}{2} (1 - \tanh[\gamma(r_{ij} - 1,2\mu_{ij})])$$

Donde  $r_{ij}$  representa la distancia entre el par de residuos  $i,j$  para cada paso de la simulación,  $\mu_{ij}$  corresponde a la distancia en el estado nativo y  $\gamma$  define la pendiente del potencial de conteo de contactos nativos.

Este potencial armónico sobre  $V_Q$  limita la simulación al muestreo de un rango particular de contactos nativos, dependiendo de los valores de  $V_{Q^*}$  y  $k$ , de tal manera que fomenta la formación de contactos nativos cuando  $V_Q < V_{Q^*}$ , y disuadiendo su formación cuando  $V_Q > V_{Q^*}$ .

Este potencial armónico fue codificado como una versión modificada de GROMACS [229].

#### 3.2.4. Potencial de solvatación de contactos nativos.

En base a la descripción microscópica del desplegamiento a baja temperatura como un evento de solvatación de los residuos que conforman el centro hidrofóbico de una proteína [220], introdujimos un potencial de solvatación basado en el trabajo de Cheung y cols [221, 232]. En este potencial, la energía del contacto nativo directo ( $\epsilon$ ), de la barrera de desolvatación ( $\epsilon_{db}$ ) y del contacto nativo mediado por agua ( $\epsilon_{ssm}$ ) pueden ser manipuladas directamente, pudiendo favorecer un tipo de contacto por sobre otro. El mínimo energético del contacto nativo mediado por agua está centrado a una distancia de  $\mu_{ssm} = \mu_{ij} + 0.3$  nm, donde  $\mu_{ssm} - \mu_{ij} = 0.3$  nm equivale al diámetro aproximado de una molécula de agua y  $\mu_{ij}$  corresponde a la distancia del par de residuos  $i,j$  en contacto en el estado nativo [221, 232]. La barrera de desolvatación está centrada a una distancia  $\mu_{db} = (\mu_{ij} + \mu_{ssm}) / 2$  [221, 232]. Para las simulaciones presentadas en este trabajo, escogimos los siguientes valores para los parámetros energéticos:  $\epsilon = 1$ ;  $\epsilon_{db} = 0.3$ ;  $\epsilon_{ssm} = 0.9$ .

Este potencial de solvatación fue codificado como una versión modificada de GROMACS [229].

#### 3.2.5. Análisis de las simulaciones.

El cálculo de parámetros termodinámicos (capacidad calórica, perfil de energía libre en función de  $Q$ ) fue realizado combinando múltiples trayectorias de simulaciones a distintos valores de  $V_{Q^*}$  a temperatura constante cercana a la temperatura de plegamiento ( $T_F$ ) (para las simulaciones con potenciales de muestreo acelerado) o a



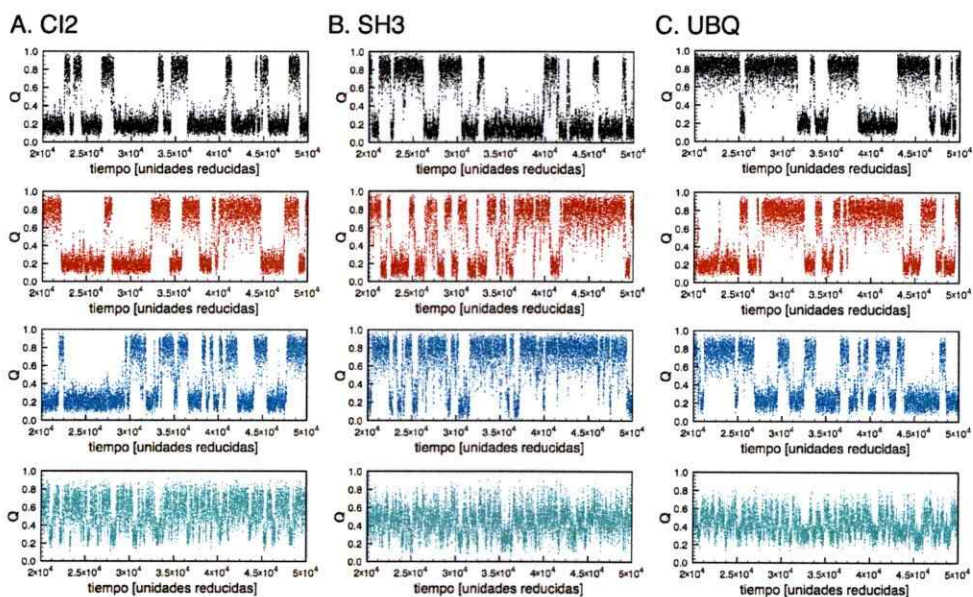
distintas temperaturas (para las simulaciones con potenciales de solvatación), mediante el método de análisis de histogramas ponderados (WHAM, *Weighted Histogram Analysis Method*) [233, 234, 235]. Para obtener la información de la probabilidad de las distintas configuraciones observadas en las simulaciones en ausencia del sesgo impuesto por el potencial armónico sobre  $V_{Q^*}$ , utilizamos el método de análisis reponderado (rWHAM) [236]. Los valores de  $\phi$  [26] fueron calculados para cada par de residuos  $i,j$  que forman un contacto nativo a partir de estas simulaciones como la probabilidad de formación de este contacto en el estado de transición sobre la probabilidad de formación en el estado nativo, de acuerdo a lo descrito en otros trabajos [43, 45].

### 3.3. RESULTADOS.

#### 3.3.1. Aceleración del muestreo de configuraciones mediante la aplicación del potencial armónico sobre los contactos nativos.

El método de muestreo acelerado fue utilizado para analizar las transiciones de plegamiento/desplegamiento de las proteínas modelo CI2, SH3 y ubiquitina, las cuales han sido previamente caracterizadas como representativas de mecanismos de plegamiento de dos estados [43, 237].

Como se observa en la figura 23, para todos los modelos estudiados (CI2, SH3 y ubiquitina), la utilización del potencial armónico sobre los contactos nativos aumenta la frecuencia de la observación de la transición entre el estado plegado y desplegado para valores de  $V_{Q^*}$  cercanos a  $Q$  (fracción del total de contactos nativos)  $\sim 0.55$  para CI2 y SH3, y cercano a  $Q \sim 0.45$  para ubiquitina. A medida que la fuerza del potencial armónico aumenta, es decir, a medida que el valor  $k$  es incrementado, la frecuencia de las transiciones entre el estado plegado y desplegado es mayor. Cuando  $k = 0.005$ , las transiciones entre ambos estados son muy frecuentes, pero la varianza de la distribución de  $Q$  es menor, explicando así la diferencia respecto a las trayectorias cuando  $k < 0.005$  (Figura 23).



**Figura 23. Transiciones entre el estado plegado y desplegado de distintas proteínas modelo con mecanismo de plegamiento de dos estados mediante potenciales armónicos sobre  $Q$ .**

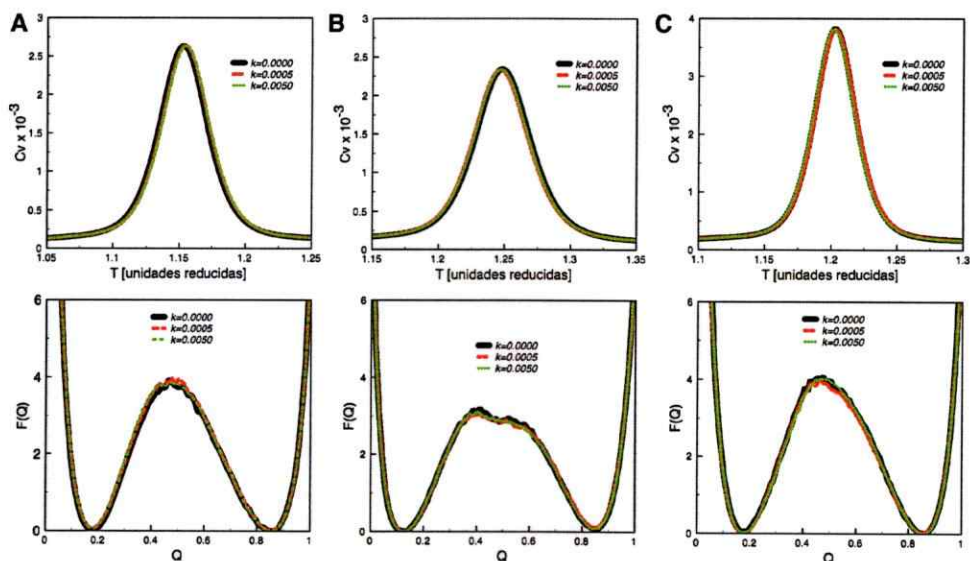
Las transiciones de plegamiento fueron obtenidas mediante simulaciones a temperatura constante  $\sim T_F$ , utilizando un potencial armónico sobre una definición continua de  $Q$  ( $V_Q$ ) para incrementar la frecuencia de estas transiciones. Distintas fuerzas para potencial armónico, determinada por el parámetro  $k$ , fueron utilizadas: 0 (símbolos negros), 0.0005 (símbolos rojos), 0.0010 (símbolos azules) y 0.0050 (símbolos verdes). Las proteínas utilizadas fueron CI2 (A), SH3 (B) y ubiquitina (UBQ, C), y los potenciales se centraron en valores de  $V_{Q^*}$  tales que  $Q \sim 0.55$  para CI2 y SH3, y  $Q \sim 0.45$  para ubiquitina.

A partir de la combinación de múltiples trayectorias utilizando distintos valores de  $VQ^*$  mediante el método de rWHAM, se obtuvo la capacidad calórica en función de la temperatura para el proceso de plegamiento y el perfil de energía libre en función de  $Q$  a la temperatura de plegamiento  $T_F$  (Figura 24). En ambos casos, no se observan cambios en la  $T_F$  (temperatura a la cual se encuentra máximo de la curva de calor específico) o en la barrera energética calculada para el mecanismo de plegamiento. Asimismo, la probabilidad de formación de los contactos nativos en el estado de transición (valor de  $Q$  en donde se encuentra el máximo de la barrera energética que separa al estado plegado del estado desplegado) es la misma, independiente de la fuerza empleada en el potencial armónico sobre  $Q$  (Figura 25).

De este modo, demostramos que esta metodología es adecuada para la aceleración de las transiciones entre los estados que una proteína determinada visita en el proceso de plegamiento termodinámico, estudiado utilizando un modelo de tipo Gō.

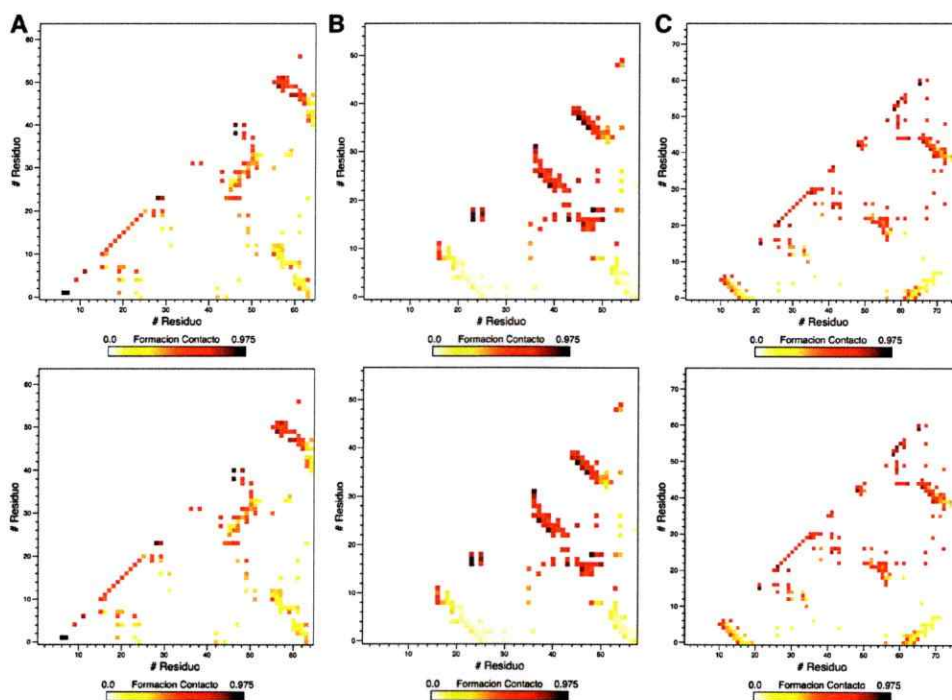
### 3.3.2. El método de muestreo acelerado permite describir el paisaje energético de plegamiento de la quinasa humana de piridoxal.

Para demostrar que estos modelos simplificados basados en la estructura nativa son útiles en el estudio del plegamiento de proteínas con topologías complejas (elevado orden de contactos y elevado número de contactos nativos [21, 24]) y, en particular, para el estudio de enzimas de la superfamilia riboquinasa [88], decidimos utilizar como modelo de prueba a la quinasa humana de piridoxal [227]. Esto, ya que corresponde a uno de los miembros de menor tamaño de la superfamilia, consistiendo solamente de un dominio (común para todos los miembros) [95], y porque el plegamiento de esta enzima [238] y de la quinasa homóloga de oveja [115] han sido caracterizados como mecanismos de dos estados, con una sola transición ( $N \leftrightarrow D$ ) cuyos valores  $m$  asociados son de 2.4 kcal/mol/M y 1.2 kcal/mol/M, respectivamente. Sin embargo, se ha sugerido la presencia de un intermediario en el proceso de desplegamiento de estas quinazas, dado



**Figura 24. Propiedades termodinámicas de las proteínas CI2, SH3 y ubiquitina para diferentes fuerzas del potencial armónico sobre  $Q$ .**

Los gráficos en la parte superior representan el cambio en la capacidad calórica en función de la temperatura, mientras que los gráficos en la parte inferior representan el perfil de energía libre del proceso de plegamiento en función de la fracción de contactos nativos ( $Q$ ). Las propiedades termodinámicas fueron determinadas mediante rWHAM [236], combinando múltiples trayectorias a distintos valores de  $V_{Q^*}$ , utilizando distintas fuerzas sobre el potencial armónico. Los valores de  $k$  utilizados en el potencial armónico corresponden a 0 (línea negra), 0.0005 (línea roja) y 0.005 (línea verde). Las proteínas analizadas corresponden a CI2 (A), SH3 (B) y ubiquitina (C)

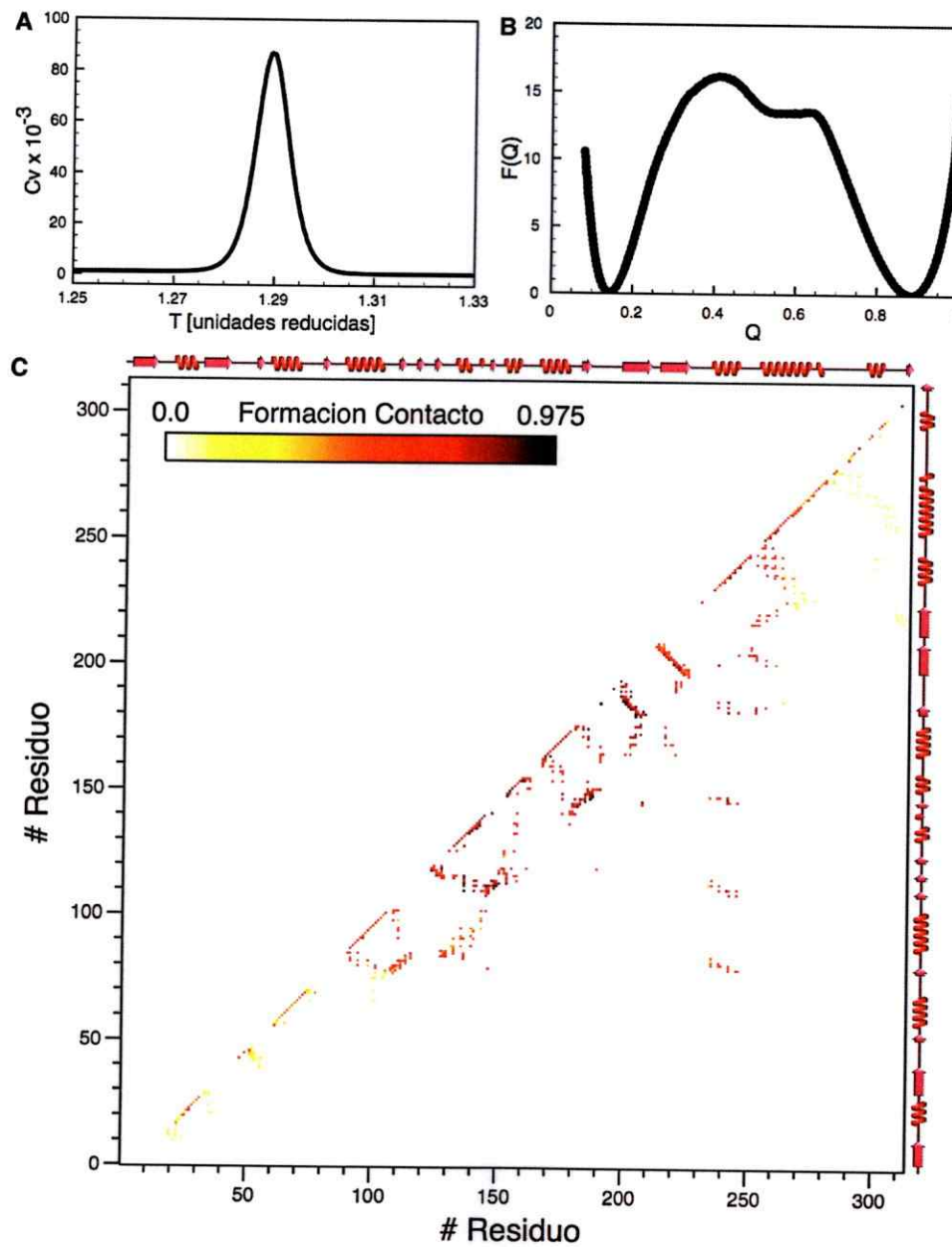


**Figura 25. Probabilidad de formación de los contactos nativos de las proteínas CI2, SH3 y ubiquitina en el estado de transición de plegamiento.**

Los gráficos corresponden a mapas de contactos nativos entre pares de residuos  $i, j$  ( $i$  en el eje  $x$ ,  $j$  en el eje  $y$ ), donde el color representa la probabilidad de formación del contacto nativo en el estado de transición. No existen diferencias entre la probabilidad de formación de contactos nativos en el estado de transición calculada en ausencia de un potencial armónico sobre  $Q$  ( $k = 0$ ; mapas de contacto en la parte superior) y en presencia de un potencial armónico elevado ( $k = 0.005$ ; mapas de contacto en parte inferior). Los mapas de contacto fueron calculados para CI2 (A), SH3 (B) y ubiquitina (C).

que el valor  $m$  esperado para el desplegamiento de la proteína completa es de 8.1 kcal/mol/M [114, 160].

Al utilizar una constante de fuerza para el potencial armónico de 0.001, logramos obtener el paisaje energético de plegamiento de la quinasa humana de piridoxal en función de  $Q$  (Figura 26). Nuestros resultados muestran que la temperatura de plegamiento ( $T_F$ ) corresponde a  $\sim 1.29$  unidades reducidas (Figura 26A). Para esta temperatura, se determinó la barrera energética que separa el estado desplegado del estado nativo, la cual es muy elevada, alcanzando un valor de  $\sim 16 k_B T$  en el máximo a  $Q \sim 0.4$ . Esto es esperable para una proteína de 312 aminoácidos y 988 contactos nativos entre pares de residuos. Además, es evidente la presencia de un intermediario meta-estable a una energía de  $\sim 15 k_B T$  cuando la fracción de contactos nativos es  $\sim 0.6$  (Figura 26B), lo que es consistente respecto a los argumentos basados en la discrepancias del valor  $m$  calculado en base a la estructura terciaria de la proteína y el observado experimentalmente [114, 115, 160, 238]. Además, determinamos la probabilidad de formación de los contactos nativos en el estado de transición de la proteína ( $Q \sim 0.40$ ), encontrando que la mitad de la región correspondiente al dominio Rossmann (residuos 1-100) exhiben una baja probabilidad de formación en el máximo de la barrera energética (Figura 26C). Esto es consistente con lo observado mediante el seguimiento del desplegamiento reversible de la quinasa de piridoxal de oveja utilizando análogos fluorescentes de ATP y piridoxal, en los cuales se demuestra que el sitio de unión a piridoxal (localizado en el extremo amino terminal), se despliega a concentraciones menores de GndHCl que el extremo carboxilo terminal, donde se une ATP [115]. Estos antecedentes sugieren fuertemente que el comportamiento observado para el mecanismo de plegamiento termodinámico de la quinasa humana de piridoxal, utilizando estos modelos simplificados sometidos a potenciales de aceleración armónicos, representa características generales de los cambios estructurales observados experimentalmente durante el desplegamiento reversible inducido por GndHCl de esta enzima [115].



**Figura 26. Análisis termodinámico del proceso de plegamiento de la quinasa humana de piridoxal mediante el uso de modelos nativo-céntricos y potenciales harmónicos sobre  $Q$ .**

A) Cambio en la capacidad calórica en función de la temperatura para el proceso de plegamiento de la quinasa humana de piridoxal (PDB ID 2F7K).

B) Perfil de energía libre del proceso de plegamiento de la quinasa humana de piridoxal en función de la fracción de contactos nativos ( $Q$ ). Se observa el estado de transición a  $Q \sim 0.4$  y un intermediario meta-estable a  $Q \sim 0.6$ .

C) Mapa de contactos nativos entre pares de residuos  $i, j$  ( $i$  en el eje  $x$ ,  $j$  en el eje  $y$ ), que muestra la probabilidad de formación en el estado de transición ( $Q \sim 0.4$ ) de la quinasa humana de piridoxal. Se muestra el contenido de estructura secundaria a lo largo de la secuencia primaria.

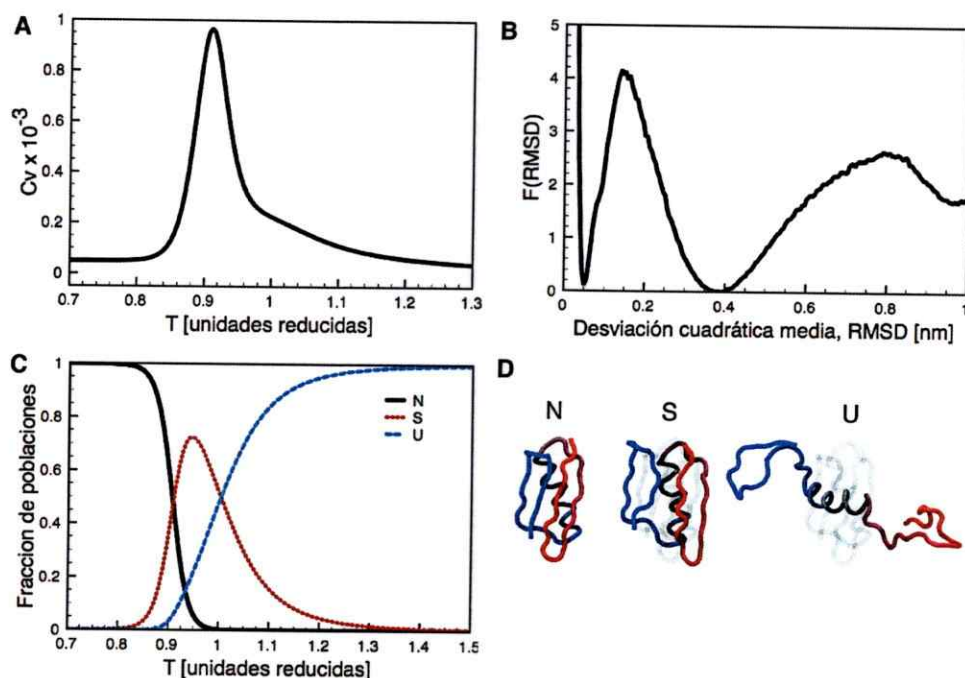
### 3.3.3. El potencial de solvatación permite describir la hidratación del dominio B1 de la proteína G de *Streptomyces*.

Los modelos Ca incluyen implícitamente información acerca del estado nativo de una proteína mediante la utilización del mapa de contactos nativos y la evaluación de la formación o rompimiento de estos contactos a través de un potencial de Lennard-Jones, cuyo único mínimo energético corresponde a la distancia entre pares de residuos en el estado nativo [43], o a través de múltiples potenciales de Gauss, cuyos mínimos energéticos corresponden a distancias conocidas *a priori* [219]. Sin embargo, estos modelos no son capaces de describir los eventos de penetración del solvente en el centro hidrofóbico de proteínas, los cuales son frecuentes cuando la perturbación utilizada para inducir al desplegamiento corresponde a presión o bajas temperaturas [220].

Es por ello que incorporamos el potencial de Cheung y cols [221] para la evaluación de los contactos nativos de la proteína, con el propósito de incluir un segundo mínimo energético correspondiente al establecimiento de un contacto nativo mediado por una molécula de agua. De esta manera, pretendemos explicar eventos de desnaturación a bajas temperaturas, como en el caso de Pfk-2 [148] o de desplegamiento inducido por presión, como en el caso del dominio B1 de la proteína G de *Streptomyces* [226]. El sometimiento de este dominio a altas presiones (2 kbar) causa la penetración de moléculas de agua en el extremo carboxilo-terminal de la única hélice  $\alpha$  presente en la estructura, alcanzando así una estructura compacta alternativa [239]. Escogimos probar nuestro potencial con éste último debido al tamaño del dominio (compuesto por 56 residuos aminoácidos), y utilizamos el método de intercambio de réplicas [230] para evitar la formación de posibles estructuras mal plegadas durante el proceso.

Utilizamos este potencial sobre el dominio B1 de la proteína G de *Streptomyces* de tal manera que, para un par de residuos  $i,j$  en contacto en el estado nativo, los contactos nativos directos y mediados por agua son cuasi-isoenergéticos (es decir,  $\epsilon \approx \epsilon_{ssm}$ ). Como se observa en la figura 27, el paisaje energético del plegamiento de esta





**Figura 27. Intermediarios penetrados por solvente en el desplegamiento termodinámico del dominio B1 de la proteína G de *Streptomyces*.**

A) Cambio en la capacidad calórica en función de la temperatura. La temperatura de plegamiento es  $\sim 0.9 T$ .

B) Perfil de energía libre en función de la desviación cuadrática media (RMSD, *root mean square deviation*) calculada a partir de la estructura nativa del dominio B1 de la proteína G de *Streptomyces*. El estado intermediario solvatado posee un RMSD de 0.4 nm respecto a la estructura nativa.

C) Fracción de las poblaciones nativa (N, *native*), penetradas por solvente (S, *swollen*) y desplegadas (U, *unfolded*), en función de la temperatura. El desplegamiento del intermediario hidratado para llegar al estado desplegado es no cooperativo.

D) Representación visual de la estructura de los estados nativo (N), penetrado por solvente (S) y desplegado (U) del dominio B1 de la proteína G de *Streptomyces*. El gradiente de color representa la posición en la secuencia primaria, el cual va desde el amino-terminal, en rojo, al carboxilo-terminal, en azul.

proteína corresponde a un mecanismo de tres estados, en buen acuerdo con lo esperado para una proteína con un intermediario penetrado por solvente, como en el caso de Pfk-2 [128] y ubiquitina [240], entre otros (Figura 27B). En efecto, al incrementar la temperatura de la simulación, se produce el poblamiento de un estado intermediario penetrado por solvente (S, Figura 27C) el cual tiene un radio de giro de 1.25 nm, mayor al calculado para el monómero compacto (1.02 nm). Este intermediario se despliega de manera no cooperativa a medida que la temperatura continúa incrementado (Figura 27C). De esta manera, el favorecimiento de las interacciones nativas mediadas por agua conduce a la formación de intermediarios penetrados por solvente que alcanzan el estado desnaturalado a través de un proceso de desplegamiento no cooperativo, consistente con teorías propuestas al respecto de este fenómeno con anterioridad [204].

### 3.4. DISCUSION.

La combinación de distintas estrategias de muestreo para acelerar los procesos de plegamiento termodinámico, y la introducción de frustración energética a través de potenciales modificados basados en parámetros nativo-céntricos, son herramientas de gran utilidad para capturar los cambios estructurales generales que ocurren durante este proceso.

En el caso de la quinasa humana de piridoxal, las características del estado de transición observado mediante el uso de modelos nativo-céntricos (Figura 26) son consistentes con la evidencia experimental [115]. De este modo, nuestros datos sugieren fuertemente que las características generales del plegamiento de esta proteína pueden ser explicadas en función del principio de frustración mínima [12], es decir, solo en base a la frustración topológica [43]. Además, considerando la elevada barrera energética que separa los estados nativo y desplegado, la implementación del potencial armónico sobre Q (Figura 23 y 24) se presenta como una herramienta invaluable para la descripción del mecanismo de plegamiento de proteínas complejas, cuyo número de

contactos nativos y orden de contacto hacen que las transiciones entre los diversos estados sean inaccesibles o muy poco frecuentes en tiempos de cómputo razonables.

Por otra parte, la manipulación directa de los mínimos energéticos del potencial de solvatación desarrollado por Cheung y cols [221] permite la descripción de las características estructurales esenciales de los estados penetrados por solvente en proteínas. La importancia de la implementación de este tipo de potenciales radica en su potencial utilización, en combinación con experimentos de difracción de rayos X de ángulo pequeño, para describir los estados intermedarios en solución de proteínas [85, 241, 242, 243, 244]. De esta manera, las características estructurales del estado intermedario monomérico de enzimas como la mutante L93A de Pfk-2 o la mutante W103A de Kdgk (ver Capítulo 2) pueden ser descritas con mayor precisión.

### 3.5. AGRADECIMIENTOS

El método de muestreo acelerado implementado en este capítulo fue utilizado para estudiar una proteína anudada, VirC2, cuyos resultados, en parte, aparecen en *Biochemical Society Transactions*, (2013), Sułkowska, J. I., Noel, J. K., Ramírez-Sarmiento, C. A., Rawdon, E. J., Millett, K. C., Onuchic, J. N. El autor de este trabajo doctoral es co-autor e investigador externo del artículo. El autor de este trabajo doctoral desea agradecer al Dr. Jose N. Onuchic, Rice University, Houston, Texas, por admitirme en el Centro de Física Biológica Teórica de la Universidad de California, San Diego, Estados Unidos; al Dr. Jeffrey K. Noel, Rice University, Houston, Texas, Estados Unidos, por su invaluable contribución en el aprendizaje de los métodos de simulación molecular con modelos simplificados y en el desarrollo del potencial de muestreo acelerado; a la Dra. Joanna I. Sułkowska, Universidad de Varsovia, Polonia, por su cordial invitación a participar del artículo publicado en conjunto; y al Dr. Mauricio Báez, por sus conversaciones fundamentales respecto a los modelos proteicos anudados y por las futuras colaboraciones en torno a este ámbito. Las simulaciones moleculares fueron realizadas en el Centro de Física Biológica Teórica, Universidad de California,

San Diego, Estados Unidos, y en el Laboratorio Nacional para Computación de Alto Rendimiento del Centro de Modelamiento Matemático, Universidad de Chile, Santiago, Chile (PIA ECM-02), además de recursos de cómputo locales adquiridos mediante financiamiento del Fondo Nacional de Desarrollo Científico y Tecnológico (FONDECYT) Nos. 1090336 y 1110137. CAR-S fue beneficiario de una beca de estudios doctorales de la Comisión Nacional de Investigación Científica y Tecnológica (CONICYT), de una beca de apoyo a la investigación en el extranjero del organismo de Mejoramiento de la Calidad y Equidad de la Educación Superior (MECESUP) en conjunto con el programa de Ciencias, Tecnología e Innovación para las Américas (STIPAS, *Sciences, Technology and Innovation Program for the Americas*) No. UCH0717 y de una beca de apoyo a la investigación en el extranjero de MECESUP No. UCH0713.

## DISCUSION GENERAL

Las proteínas homodiméricas exhiben diversos mecanismos de (des)plegamiento, siendo los más habituales los mecanismos que siguen un modelo de dos estados y los que siguen un modelo de tres estados, con un intermediario que puede estar compuesto por una o dos cadenas polipeptídicas [49]. En proteínas homo/heterodiméricas pequeñas (< 100 residuos), como el represor Arc [46], represor cro del fago lambda [47] y represor de triptófano [48], el mecanismo de (des)plegamiento puede ser sugerido *a priori* al examinar la estructura nativa del oligómero, particularmente la proporción de contactos nativos intermoleculares sobre los contactos nativos intramoleculares y la hidrofobicidad de la interfaz de asociación entre subunidades [50].

Las proteínas que presentan más de 100 residuos en su cadena polipeptídica generalmente se componen de más de un dominio, lo que les permite alcanzar el estado nativo mediante vías de plegamiento paralelas e independientes para cada uno de ellos [69, 71]. Sin embargo, en el caso de topologías discontinuas entre dominios [74], la conectividad obligada a través de los cruces de la cadena polipeptídica entre éstos promueve la cooperatividad e impone el acoplamiento entre los procesos de desplegamiento de cada dominio [37]. Este es el caso para la enzima Pfk-2 de *Escherichia coli*, en la cual tanto el desplegamiento inducido químicamente como la desnaturación inducida por bajas temperaturas llevan a la acumulación de una especie monomérica expandida cuya accesibilidad al solvente, de acuerdo a nuestros experimentos de HXMS, es mayor que en el caso del dímero nativo (Capítulo 1 y Capítulo 2, Figuras 13 y 21) [128, 148]. Esto es consecuencia de la topología reentrante, que conecta el dominio mayor y menor de la enzima, y de la topología de la interfaz de asociación de la cual éste último forma parte, los cuales restringen estructuralmente al dímero nativo a conformar una sola unidad cooperativa [128].

La observación del estado monomérico penetrado por solvente como consecuencia de un proceso de desnaturación en frío, es consistente con observaciones

previas en otros modelos de estudio, como en la ribonucleasa A bovina pancreática [246], frataxina de *Saccharomyces cerevisiae* [133], ubiquitina humana [138], la mutante I98A del dominio C-terminal de la proteína ribosomal L9 [139], el inhibidor de subtilisina de *Streptomyces* [163] y el homodímero CylR2 de *Enterococcus faecalis* [247]. En el caso de la proteína L9, el estado desnaturalado en frío ha sido referido como compacto, pero expandido a bajas temperaturas [136]. El estado desnaturalado en frío del inhibidor de subtilisina ha sido caracterizado mediante difracción de rayos X de ángulo pequeño como un estado más compacto que el observado para la proteína desnaturalada en urea [163]. Para CylR2, se ha resuelto la estructura en solución por RMN de diversos estados conformacionales a distintas temperaturas bajo 0 °C, revelando la disociación de la proteína hacia un intermediario compacto similar al estado nativo, el cual alcanza un estado parcialmente desplegado altamente flexible a medida que la temperatura es menor [247].

Tanto para el dominio C-terminal de la proteína L9 como para frataxina, el proceso de desnaturalación en frío es cooperativo [133, 139], lo cual es similar para el caso de Pfk-2 [128], pero difiere de otros modelos de estudio, en los cuales este proceso ha sido descrito como no-cooperativo [138, 246]. Hemos argumentado que esta inconsistencia se debe al uso de perturbaciones adicionales para acceder al proceso de desnaturalación en frío [138, 246] y a la inaccesibilidad cinética de algunos modelos de estudio para alcanzar el estado desnaturalado en frío en condiciones de equilibrio, como es el caso de tioredoxina de *Escherichia coli* [141]. En efecto, los tres modelos de estudio cuyo desplegamiento en frío ha sido descrito como cooperativo (Pfk-2 [128], frataxina [133] y el dominio C-terminal de la proteína L9 [139]) han sido estudiados en ausencia de perturbaciones adicionales, lo que corrobora parcialmente esta afirmación.

Los estados penetrados por solvente también han sido detectados cuando las proteínas se someten a altas presiones, como en el caso de ribonucleasa A bovina pancreática [246], el homodímero del represor Arc [248], el homodímero de  $\beta$ -lactoglobulina A [249] y la ribonucleasa HI de *Escherichia coli* [250]. Esto se debe a

que el incremento de presión distorsiona la estructura nativa de la proteína, aumentando la flexibilidad conformacional e incrementando las fluctuaciones en el volumen proteico, lo que provoca la formación de cavidades en el centro hidrofóbico que pueden ser penetradas por el solvente circundante [246]. En efecto, la presencia de cavidades en el estado plegado y la reestructuración y consecuente penetración del solvente en estas cavidades producto de la alta presión han sido sugeridos como un mecanismo general del desplegamiento de proteínas mediante esta perturbación [251]. En el caso de ribonucleasa A bovina pancreática, el estado desnaturado por presión ha sido descrito como un glóbulo fundido [252], correspondiente a un estado intermediario compacto con estructura secundaria nativa remanente y un interior "líquido" [246]. En este estado inducido por presión, la mayoría de las amidas que se encuentran al interior de la proteína intercambian los átomos de hidrógeno con el solvente circundante más rápido que en el caso del estado nativo, pero más lento que en el caso de una cadena polipeptídica completamente desplegada [246]. Asimismo, en el caso de la ribonucleasa HI de *Escherichia coli*, la conformación inducida por presión se ha descrito como un estado novel similar a un glóbulo fundido, el cual es compacto pero exhibe una alta entropía conformacional para las cadenas laterales y la cadena principal [250]. Todas estas descripciones están en buen acuerdo con lo observado para el caso de Pfk-2 (Capítulo 1) [148] y son consistentes con la visión microscópica unificadora del mecanismo de desnaturación inducido por bajas temperaturas y por presión [220]. Además, dado que las características del intermediario inducido por incubación de la enzima Pfk-2 a 3° C (Capítulo 1) [128, 148] y del intermediario inducido mediante el uso de agentes caotrópicos (Capítulo 2, Figura 13 y 21) [114] son similares, la estructura del intermediario inducido a 0.85 M de GndHCl puede ser concebida como un estado expandido que presenta cavidades hidrofóbicas remanentes pero es altamente hidratado, como fue propuesto inicialmente [114].

Dado que el mecanismo microscópico de desplegamiento de proteínas inducido por presión y por exposición a baja temperatura son similares, en cuanto a la penetración

del agua en el centro hidrofóbico de proteínas y la subsecuente formación de contactos mediados por agua [220], estos eventos pueden ser caracterizados mediante la utilización de modelos simplificados [43, 45] que incorporen potenciales de solvatación de contactos nativos y barreras de desolvatación para alcanzar el estado plegado [221]. La incorporación de estos potenciales para describir el desplegamiento termodinámico inducido por presión de una proteína modelo (Capítulo 3, Figura 27) confirman que estos potenciales permiten capturar las características generales de la hidratación del centro hidrofóbico de proteínas. Esto, acompañado de la aceleración de las transiciones de plegamiento mediante la aplicación de potenciales armónicos, utilizando a los contactos nativos como variable colectiva (Capítulo 3, Figura 23), permitirán el estudio de la penetración del solvente en proteínas cuya complejidad topológica hace que estos procesos sean generalmente inaccesibles desde el punto de vista computacional, sin inducir a la perturbación de los estados de transición observados (Capítulo 3, Figura 25).

La disociación de homodímeros también puede ser inducida mediante la utilización de perturbaciones locales, a través de la introducción de mutaciones sitio-dirigidas en la interfaz de asociación proteína-proteína. La aplicación de este tipo de perturbaciones sobre Pfk-2 conduce al alcance de un estado monomérico (Capítulo 2, Figura 11) con mayor grado de compactación que las especies observadas mediante perturbaciones globales (cambios de temperatura o de composición del solvente; Capítulo 2, Tabla 2), pero con propiedades hidrodinámicas y características estructurales locales propias de una especie con mayor grado de accesibilidad al solvente y menor estabilidad conformacional que el dímero nativo (Capítulo 2, Figuras 14 y 21). Estos rearrreglos estructurales no ocurren de manera homogénea sobre la estructura de la proteína, sino que se encuentran localizados en el dominio Rossmann (Capítulo 2, Figura 15), producto de la topología reentrante de la cadena polipeptídica entre el dominio menor y la región amino-terminal del dominio mayor [94]. Las propiedades hidrodinámicas de los monómeros aislados producto de perturbaciones locales sobre la interfaz de asociación de homodímeros ha sido previamente caracterizada en otros



modelos de estudio, como el dominio Apple 4 (A4) del factor de coagulación XI [164] y la quinasa dimérica de creatina de músculo de conejo (MM-CK) [169]. En ambos casos, la perturbación de la interfaz de asociación lleva a la formación de una especie monomérica, cuyo grado de compactación es menor al de la subunidad aislada cuando la disociación es inducida por agentes caotrópicos. En el caso de MM-CK, la unión de ANS para la mutante monomérica es mayor que en el caso del dímero en condiciones nativas, sugiriendo que esta estructura corresponde también a un glóbulo fundido [169]. En efecto, el cambio de estructura secundaria producto de la adición e incremento de concentración de agentes caotrópicos para esta mutante monomérica exhibe un comportamiento no cooperativo, característico de especies hidratadas observadas en procesos de desnaturación a altas presiones [246] y consistente con otras propuestas teóricas al respecto [204]. Estos aspectos son consistentes con lo observado para Pfk-2, considerando que el mayor cambio de estabilidad de la enzima corresponde a la disociación de las subunidades [114], mientras que el intermediario monomérico inducido mediante GndHCl o cambios de temperatura es marginalmente estable [128]. En su conjunto, estos argumentos sugieren que la estabilidad del monómero depende de la formación del dímero. La combinación de los modelos simplificados de simulación molecular propuestos que incorporan configuraciones separadas por solvente (Capítulo 3, Figura 27) con estrategias experimentales como difracción de rayos X de ángulo pequeño, la cual permite determinar características estructurales elementales acerca de la forma y las propiedades hidrodinámicas de proteínas (Capítulo 3, Figura 20), nos permitirán resolver la estructura en solución de estos monómeros aislados hidratados, de manera similar a lo que ha sido realizado en otros modelos proteicos [85, 241, 242, 243, 244].

Nuestros resultados también sugieren que la observación de especies monoméricas flexibles y penetradas por solvente no es una característica exclusiva de la enzima Pfk-2 de *Escherichia coli*, sino que también pueden ser observadas como consecuencia de la disociación de otros homodímeros de la familia riboquinasa, como lo

expuesto para la enzima Kdgk de *Escherichia coli*. La disociación de las subunidades de esta quinasa por medio de perturbaciones locales (Capítulo 2, Figura 11) y globales (Capítulo 2, Figura 12) conducen a cambios en las propiedades hidrodinámicas y en el contenido de estructura secundaria que son comparables y/o similares a los que ocurren en el caso de Pfk-2 (Capítulo 2, Tabla 2 y Figura 17). Pese a que, a diferencia de la enzima Pfk-2, no pudimos determinar la flexibilidad local del monómero aislado de la enzima Kdgk mediante HXMS, los resultados de la difracción de rayos X de ángulo pequeño del dímero nativo y de la mutante monomérica W103A (Capítulo 2, Figura 18 y 19) nos permitieron confirmar que la subunidad aislada es relativamente compacta, pero más expandida que lo esperado para el monómero en el contexto dímérico (Capítulo 2, Figura 20). Nuestros resultados sugieren fuertemente que la dependencia de la estabilidad monomérica en la formación del dímero nativo es una propiedad general de los homodímeros de la familia riboquinasa.

Esta dependencia de la estabilidad puede ser explicada en base al principio de la frustración mínima [12]. Mediante estudios *in silico*, determinamos que los residuos que estabilizan en mayor medida la interacción entre subunidades se encuentran localizados en las regiones similares del dominio menor para los distintos homodímeros, correspondiente a la última hebra  $\beta$  carboxilo-terminal de este dominio (Capítulo 2, Figura 9). Esto ocurre pese a que las características del dominio menor, la rotación de los monómeros en el contexto dímérico y la forma del dímero son diferentes entre los distintos miembros (Introducción, Figura 2). Sin embargo, la manera en que los residuos en estas posiciones interactúan en el contexto dímérico difieren para los distintos modelos de estudio: Para Pfk-2, esta hebra  $\beta$  establece interacciones con la hebra proveniente de la subunidad adyacente [94], mientras que en el caso de las enzimas Kdgk, esta posición es generalmente ocupada por un residuo aromático que se encuentra en el medio de la sábana  $\beta$  curvada constituida por la asociación entre subunidades, y que establece interacciones con el residuo aromático equivalente de la subunidad

adyacente [110, 112]. Además, existe un segundo hot-spot en la hélice 3<sub>10</sub> del dominio menor (Capítulo 2, Figura 9). Considerando que la asociación entre subunidades provoca la disminución de la frustración mutacional y configuracional en estas regiones de la proteína (Capítulo 2, Figura 10), nuestros antecedentes sugieren que tanto la identidad de estos residuos como la posición que ocupan en el dominio menor son una consecuencia de la topología de la interfaz de asociación proteína-proteína.

## CONCLUSIONES

- La exposición a baja temperatura de la enzima homodimérica Pfk-2 de *Escherichia coli* induce un proceso de desplegamiento reversible altamente cooperativo. Mediante experimentos de ultracentrifugación analítica, dicroísmo circular y fluorescencia, determinamos que el estado final de este proceso de desplegamiento corresponde a un intermediario monomérico que exhibe propiedades estructurales globales (contenido de estructura secundaria, presencia de cavidades hidrofóbicas) similares a las del intermediario inducido por concentraciones bajas de GndHCl.
- El análisis de los cambios estructurales locales de la enzima Pfk-2 durante el proceso de desnaturación a baja temperatura mediante experimentos de HXMS, demostraron que este proceso ocurre en la escala de tiempo de horas y que el proceso de disociación y desplegamiento son concertados, es decir, la cinética de disociación medida por fluorescencia intrínseca y la cinética de incorporación de deuterio para las distintas regiones de la proteína, son similares dentro del error experimental ( $k \sim 1.00 \times 10^{-4} \pm 0.25 \times 10^{-4} \text{ s}^{-1}$ ).
- El análisis *in silico* de la interfaz de asociación entre subunidades demostró que los residuos responsables de la estabilidad de la interacción proteína-proteína en las enzimas homodiméricas de la familia riboquinasa se localizan principalmente en la última hebra  $\beta$  en posición carboxilo-terminal del dominio menor de estas quinasas. La identidad y posición de estos residuos permite la minimización de la frustración energética luego de la asociación entre subunidades.
- La perturbación local de la interfaz de asociación de la enzima Pfk-2, mediante la mutación sitio-dirigida del residuo L93 por alanina, permite la obtención del monómero aislado de esta enzima a concentraciones de proteína de 20  $\mu\text{M}$ . Mediante experimentos de ultracentrifugación analítica y desplegamiento

inducido por GndHCl, demostramos que este monómero es más compacto ( $R_h = 2.9 \pm 0.1$  nm) que el intermediario monomérico observado mediante perturbaciones globales ( $R_h = 3.8$  nm) y que la adición de agentes caotrópicos causa el desplegamiento no cooperativo de esta mutante. Experimentos de HXMS sobre esta mutante permitieron determinar que los cambios estructurales ocurren principalmente sobre el dominio Rossmann, el cual está conectado al dominio menor a través de la topología reentrante de la cadena polipeptídica, sugiriendo que el estado configuracional de esta mutante representa los primeros arreglos estructurales producto de la disociación del homodímero.

- La mutación sitio-dirigida de los residuos W103 y Y101 por alanina de la enzima homodimérica Kdgk de *Escherichia coli* conducen a la obtención de la subunidad aislada y de una especie dimérica cuya interfaz de asociación es inestable en comparación a la enzima silvestre, respectivamente. Mediante experimentos de ultracentrifugación analítica y difracción de rayos X de ángulo pequeño, demostramos que las subunidades aisladas poseen un radio hidrodinámico y un radio de giro mayor ( $R_h \sim 3.1$  nm,  $R_g \sim 2.6$ ) que el esperado para el monómero compacto ( $R_h = 2.6$  nm,  $R_g \sim 2.0$ ), similar a lo observado para la enzima Pfk-2.
- El desplegamiento de la enzima Kdgk silvestre y de la mutante Y101A mediante la adición de concentraciones bajas de GndHCl conduce a la formación de un intermediario monomérico con un menor contenido de estructura secundaria que la enzima dimérica nativa. Este intermediario exhibe un radio hidrodinámico 14% mayor que el de la mutante W103A, pese a que el contenido de estructura secundaria entre ellos es similar.
- Las propiedades hidrodinámicas y estructurales de la subunidad aislada de las enzimas homodiméricas Pfk-2 y Kdgk de *Escherichia coli* sugieren que la dependencia de la estabilidad del monómero en la formación del oligómero es una característica común de los homodímeros de la familia riboquinasa.

- Implementamos un método de aceleración de las transiciones de plegamiento en modelos computacionales nativo-céntricos mediante potenciales armónicos que utiliza a los contactos nativos como variable colectiva, y un potencial de solvatación de contactos nativos. El potencial armónico permitió estudiar el desplegamiento termodinámico de la quinasa humana de piridoxal, demostrando que sigue un mecanismo de desplegamiento de dos estados cuyo estado de transición es consistente con la evidencia experimental, a la vez que permitió demostrar la utilidad del potencial armónico para describir el desplegamiento termodinámico de proteínas topológicamente complejas. El potencial de solvatación fue utilizado para describir el desplegamiento termodinámico por solvatación del dominio B1 de la proteína G de *Streptomyces*, en el cual se observa la acumulación cooperativa de un intermediario penetrado por solvente, lo que está en buen acuerdo con la visión microscópica del mecanismo de desplegamiento inducido por presión y por baja temperatura.

## REFERENCIAS

- [1] Strandberg, B., Dickerson, R. E. Rossmann, M. G. (2009) 50 Years of Protein Structure Analysis. *J Mol Biol* **392**, 2-32.
- [2] Zwanzig, R., Szabo, A., Bagchi, B. (1992) Levinthal's paradox. *Proc Natl Acad Sci USA* **89**, 20-22.
- [3] Anfinsen, C. B. (1973) Principles that govern the folding of protein chains. *Science* **181**, 223-230.
- [4] Levinthal, C. (1968). Are there pathways for protein folding? *J Chim Phys* **85**, 44-45.
- [5] Dill, K. A., Chan, H. S. (1997) From Levinthal to pathways to funnels. *Nat Struct Biol* **4**, 10-19.
- [6] Barrick, D. (2009) What have we learned from the studies of two-state folders, and what are the unanswered questions about two-state protein folding? *Phys Biol* **6**, 015001.
- [7] Frauenfelder, H, Sligar, S. G., Wolynes, P. G. (1991). The energy landscapes and motions of proteins. *Science* **254**, 1598-1603.
- [8] Onuchic, J. N., Luthey-Schulten, Z., Wolynes, P. G. (1997) Theory of protein folding: the energy landscape perspective. *Annu Rev Phys Chem* **48**, 545-600.
- [9] Rose, G. D., Fleming, P. J., Banavar, J. R., Maritan, A. (2006) A backbone-based theory of protein folding. *Proc Natl Acad Sci USA* **103**, 16623-16633.
- [10] Wolynes, PG. (2008). The protein folding energy landscape: a primer. *Protein Folding, Misfolding and Aggregation: Classical themes and novel approaches*; Nedie, S., Muñoz, V., Ed.; RSC Publishing: London, UK; 49-69.
- [11] Onuchic, J. N., Wolynes, P. G. (2004) Theory of protein folding. *Curr Opin Struct Biol* **14**, 70-75.
- [12] Bryngelson, J. D., Wolynes, P. G. (1987) Spin glasses and the statistical mechanics of protein folding. *Proc Natl Acad Sci USA* **84**, 7524-7528.

- [13] 1. Noé, F., Schütte, C., Vanden-Eijnden, E., Reich, L., Weikl, T. R. (2009) Constructing the equilibrium ensemble of folding pathways from short off-equilibrium simulations. *Proc Natl Acad Sci USA* **106**, 19011–19016.
- [14] Bryngelson, J. D., Onuchic, J. N., Socci, N. D., Wolynes, P. G. (1995) Funnels, pathways, and the energy landscape of protein folding: a synthesis. *Proteins* **21**, 167–195.
- [15] Onuchic, J. N., Nymeyer, H., García, A. E., Chahine, J., Socci, N. D. (2000) The energy landscape theory of protein folding: insights into folding mechanisms and scenarios. *Adv. Protein Chem.* **53**, 87–152
- [16] Cuff, A. L., Sillitoe, I., Lewis, T., Redfern, O. C., Garratt, R., Thornton, J., Orengo, C. A. (2009) The CATH classification revisited—architectures reviewed and new ways to characterize structural divergence in superfamilies. *Nucleic Acids Res* **37**, D310–4.
- [17] Plaxco, K. W., Simons, K. T., Baker, D. (1998) Contact order, transition state placement and the refolding rates of single domain proteins. *J Mol Biol* **277**, 985–994.
- [18] Ivankov, D. N., Garbuzynskiy, S. O., Alm, E., Plaxco, K. W., Baker, D., Finkelstein, A. V. (2003) Contact order revisited: influence of protein size on the folding rate. *Protein Sci.* **12**, 2057–2062.
- [19] Gong, H., Isom, D. G., Srinivasan, R., Rose, G. D. (2003) Local secondary structure content predicts folding rates for simple, two-state proteins. *J Mol Biol* **327**, 1149–1154.
- [20] Plaxco, K. W., Simons, K. T., Ruczinski, I., Baker, D. (2000) Topology, stability, sequence, and length: defining the determinants of two-state protein folding kinetics. *Biochemistry* **39**, 11177–11183.
- [21] Koga, N., Takada, S. (2001) Roles of native topology and chain-length scaling in protein folding: a simulation study with a Go-like model. *J Mol Biol* **313**, 171–180.



- [22] Ivankov, D. N., Bogatyreva, N. S., Lobanov, M. Y., and Galzitskaya, O. V. (2009) Coupling between properties of the protein shape and the rate of protein folding. *PLoS ONE* 4, e6476.
- [23] Rustad, M., and Ghosh, K. (2012) Why and how does native topology dictate the folding speed of a protein? *J Chem Phys* 137, 205104.
- [24] Makarov, D. E., Keller, C. A., Plaxco, K. W., Metiu, H. (2002) How the folding rate constant of simple, single-domain proteins depends on the number of native contacts. *Proc Natl Acad Sci USA* 99, 3535–3539.
- [25] Fersht, A. R., Matouschek, A., Serrano, L. (1992) The folding of an enzyme. I. Theory of protein engineering analysis of stability and pathway of protein folding. *J Mol Biol* 224, 771–782.
- [26] Fersht, A. R., Sato, S. (2004) Phi-value analysis and the nature of protein-folding transition states. *Proc Natl Acad Sci USA* 101, 7976–7981.
- [27] Serrano, L., Matouschek, A., Fersht, A. R. (1992) The folding of an enzyme. III. Structure of the transition state for unfolding of barnase analysed by a protein engineering procedure. *J Mol Biol* 224, 805–818.
- [28] Itzhaki, L. S., Otzen, D. E., Fersht, A. R. (1995) The structure of the transition state for folding of chymotrypsin inhibitor 2 analysed by protein engineering methods: evidence for a nucleation-condensation mechanism for protein folding. *J Mol Biol* 254, 260–288.
- [29] Grantcharova, V. P., Riddle, D. S., Santiago, J. V., Baker, D. (1998) Important role of hydrogen bonds in the structurally polarized transition state for folding of the src SH3 domain. *Nat. Struct. Biol.* 5, 714–720.
- [30] Riddle, D. S., Grantcharova, V. P., Santiago, J. V., Alm, E., Ruczinski, I., Baker, D. (1999) Experiment and theory highlight role of native state topology in SH3 folding. *Nat. Struct. Biol.* 6, 1016–1024.

- [31] Sosnick, T. R., Jackson, S., Wilk, R. R., Englander, S. W., DeGrado, W. F. (1996) The role of helix formation in the folding of a fully alpha-helical coiled coil. *Proteins* **24**, 427–432.
- [32] López-Hernández, E., Serrano, L. (1996) Structure of the transition state for folding of the 129 aa protein CheY resembles that of a smaller protein, CI-2. *Fold Des.* **1**, 43–55.
- [33] Villegas, V., Martínez, J. C., Avilés, F. X., Serrano, L. (1998) Structure of the transition state in the folding process of human procarboxypeptidase A2 activation domain. *J Mol Biol* **283**, 1027–1036.
- [34] Chiti, F., Taddei, N., White, P. M., Bucciantini, M., Magherini, F., Stefani, M., Dobson, C. M. (1999) Mutational analysis of acylphosphatase suggests the importance of topology and contact order in protein folding. *Nat. Struct. Biol.* **6**, 1005–1009.
- [35] Bueno, M., Ayuso-Tejedor, S., Sancho, J. (2006) Do proteins with similar folds have similar transition state structures? A diffuse transition state of the 169 residue apoflavodoxin. *J Mol Biol* **359**, 813–824.
- [36] Baker, D. (2000) A surprising simplicity to protein folding. *Nature* **405**, 39–42.
- [37] Cellitti, J., Llinas, M., Echols, N., Shank, E. A., Gillespie, B., Kwon, E., Crowder, S. M., Dahlquist, F. W., Alber, T., Marqusee, S. (2007) Exploring subdomain cooperativity in T4 lysozyme I: structural and energetic studies of a circular permutant and protein fragment. *Protein Sci.* **16**, 842–851.
- [38] Otzen, D. E., Fersht, A. R. (1998) Folding of circular and permuted chymotrypsin inhibitor 2: retention of the folding nucleus. *Biochemistry* **37**, 8139–8146.
- [39] Viguera, A. R., Serrano, L., Wilmanns, M. (1996) Different folding transition states may result in the same native structure. *Nat. Struct. Biol.* **3**, 874–880.
- [40] Lindberg, M., Tångrot, J., Oliveberg, M. (2002) Complete change of the protein folding transition state upon circular permutation. *Nat. Struct. Biol.* **9**, 818–822.

- [41] Lindberg, M. O., Tångrot, J., Otzen, D. E., Dolgikh, D. A., Finkelstein, A. V., Oliveberg, M. (2001) Folding of circular permutants with decreased contact order: general trend balanced by protein stability. *J Mol Biol* **314**, 891–900.
- [42] Ueda, Y., Taketomi, H., Gō, N. (1978) Studies on protein folding, unfolding, and fluctuations by computer simulation. II. A. Three-dimensional lattice model of lysozyme. *Biopolymers* **17**, 1531–1548.
- [43] Clementi, C., Nymeyer, H., Onuchic, J. N. (2000) Topological and energetic factors: what determines the structural details of the transition state ensemble and "en-route" intermediates for protein folding? an investigation for small globular proteins. *J Mol Biol* **298**, 937–953.
- [44] Tozzini, V. (2005) Coarse-grained models for proteins. *Curr. Opin. Struct. Biol.* **15**, 144–150.
- [45] Clementi, C., Jennings, P. A., Onuchic, J. N. (2000) How native-state topology affects the folding of dihydrofolate reductase and interleukin-1beta. *Proc Natl Acad Sci USA* **97**, 5871–5876.
- [46] Milla, M. E., Brown, B. M., Waldburger, C. D., Sauer, R. T. (1995) P22 Arc Repressor: Transition State Properties Inferred from Mutational Effects on the Rates of Protein Unfolding and Refolding. *Biochemistry* **34**, 13914–13919.
- [47] Filimonov, V. V., Rogov, V. V. (1996) Reversible association of the equilibrium unfolding intermediate of lambda Cro repressor. *J Mol Biol* **255**, 767–777.
- [48] Mann, C. J., Matthews, C. R. (1993) Structure and stability of an early folding intermediate of Escherichia coli trp aporepressor measured by far-UV stopped-flow circular dichroism and 8-anilino-1-naphthalene sulfonate binding. *Biochemistry* **32**, 5282–5290.
- [49] Rumfeldt, J. A. O., Galvagnion, C., Vassall, K. A., Meiering, E. M. (2008) Conformational stability and folding mechanisms of dimeric proteins. *Prog. Biophys. Mol. Biol.* **98**, 61–84.

- [50] Levy, Y., Wolynes, P. G., Onuchic, J. N. (2004) Protein topology determines binding mechanism. *Proc Natl Acad Sci USA* **101**, 511–516.
- [51] Levy, Y., Cho, S. S., Onuchic, J. N., Wolynes, P. G. (2005) A Survey of Flexible Protein Binding Mechanisms and their Transition States Using Native Topology Based Energy Landscapes. *J Mol Biol* **346**, 1121–1145.
- [52] Bennett, M. J., Schlunegger, M. P., Eisenberg, D. (1995) 3D domain swapping: a mechanism for oligomer assembly. *Protein Sci.* **4**, 2455–2468.
- [53] Bennett, M. J., Choe, S., Eisenberg, D. (1994) Domain swapping: entangling alliances between proteins. *Proc Natl Acad Sci USA* **91**, 3127–3131.
- [54] Rousseau, F., Schymkowitz, J. W. H., Itzhaki, L. S. (2003) The unfolding story of three-dimensional domain swapping. *Structure* **11**, 243–251.
- [55] Yang, S., Cho, S. S., Levy, Y., Cheung, M. S., Levine, H., Wolynes, P. G., Onuchic, J. N. (2004) Domain swapping is a consequence of minimal frustration. *Proc Natl Acad Sci USA* **101**, 13786–13791.
- [56] Yang, S., Levine, H., Onuchic, J. N. (2005) Protein oligomerization through domain swapping: role of inter-molecular interactions and protein concentration. *J Mol Biol* **352**, 202–211.
- [57] Tovchigrechko, A., Vakser, I. A. (2001) How common is the funnel-like energy landscape in protein-protein interactions? *Protein Sci.* **10**, 1572–1583.
- [58] Shen, Y., Paschalidis, I. C., Vakili, P., Vajda, S. (2008) Protein docking by the underestimation of free energy funnels in the space of encounter complexes. *PLoS Comput Biol* **4**, e1000191.
- [59] Tsai, C. J., Kumar, S., Ma, B., Nussinov, R. (1999) Folding funnels, binding funnels, and protein function. *Protein Sci.* **8**, 1181–1190.
- [60] Luque, I., Freire, E. (2000) Structural stability of binding sites: consequences for binding affinity and allosteric effects. *Proteins* **4**, 63–71.

- [61] Shoemaker, B. A., Portman, J. J., Wolynes, P. G. (2000) Speeding molecular recognition by using the folding funnel: the fly-casting mechanism. *Proc Natl Acad Sci USA* **97**, 8868–8873.
- [62] Forman-Kay, J. D. (1999) The “dynamics” in the thermodynamics of binding. *Nat. Struct. Biol.* **6**, 1086–1087.
- [63] Vakser, I. A. (1996) Main-chain complementarity in protein-protein recognition. *Protein Eng.* **9**, 741–744.
- [64] Vakser, I. A., Matar, O. G., Lam, C. F. (1999) A systematic study of low-resolution recognition in protein–protein complexes. *Proc Natl Acad Sci USA* **96**, 8477–8482.
- [65] Tsai, C. J., Ma, B., Nussinov, R. (1999) Folding and binding cascades: shifts in energy landscapes. *Proc Natl Acad Sci USA* **96**, 9970–9972.
- [66] Kumar, S., Ma, B., Tsai, C. J., Sinha, N., Nussinov, R. (2000) Folding and binding cascades: dynamic landscapes and population shifts. *Protein Sci.* **9**, 10–19.
- [67] Jackson, S. E. (1998) How do small single-domain proteins fold? *Fold Des* **3**, R81–91.
- [68] Brockwell, D. J., Radford, S. E. (2007) Intermediates: ubiquitous species on folding energy landscapes? *Curr. Opin. Struct. Biol.* **17**, 30–37.
- [69] Jahn, T. R., Radford, S. E. (2005) The Yin and Yang of protein folding. *FEBS J.* **272**, 5962–5970.
- [70] Vabulas, R. M., Raychaudhuri, S., Hayer-Hartl, M., Hartl, F. U. (2010) Protein folding in the cytoplasm and the heat shock response. *Cold Spring Harb Perspect Biol* **2**, a004390.
- [71] Radford, S. E., Dobson, C. M., Evans, P. A. (1992) The folding of hen lysozyme involves partially structured intermediates and multiple pathways. *Nature* **358**, 302–307.
- [72] Apic, G., Gough, J., Teichmann, S. A. (2001) Domain combinations in archaeal, eubacterial and eukaryotic proteomes. *J Mol Biol* **310**, 311–325.

- [73] Netzer, W. J., Hartl, F. U. (1997) Recombination of protein domains facilitated by co-translational folding in eukaryotes. *Nature* **388**, 343–349.
- [74] Wetlaufer, D. B. (1973) Nucleation, rapid folding, and globular intrachain regions in proteins. *Proc Natl Acad Sci USA* **70**, 697–701.
- [75] Thornton, J. M., Sibanda, B. L. (1983) Amino and carboxy-terminal regions in globular proteins. *J Mol Biol* **167**, 443–460.
- [76] Weaver, L. H., Matthews, B. W. (1987) Structure of bacteriophage T4 lysozyme refined at 1.7 Å resolution. *J Mol Biol* **193**, 189–199.
- [77] Shank, E. A., Cecconi, C., Dill, J. W., Marqusee, S., Bustamante, C. (2010) The folding cooperativity of a protein is controlled by its chain topology. *Nature* **465**, 637–640.
- [78] Blattner, F. R., Plunkett, G., Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B., Shao, Y. (1997) The complete genome sequence of *Escherichia coli* K-12. *Science* **277**, 1453–1462.
- [79] Ouzounis, C. A., Karp, P. D. (2000) Global properties of the metabolic map of *Escherichia coli*. *Genome Res.* **10**, 568–576.
- [80] Bennett, B. D., Kimball, E. H., Gao, M., Osterhout, R., Van Dien, S. J., Rabinowitz, J. D. (2009) Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*. *Nat. Chem. Biol.* **5**, 593–599.
- [81] Jardine, O., Gough, J., Chothia, C., Teichmann, S. A. (2002) Comparison of the small molecule metabolic enzymes of *Escherichia coli* and *Saccharomyces cerevisiae*. *Genome Res.* **12**, 916–929.
- [82] Whitford, P. C., Miyashita, O., Levy, Y., Onuchic, J. N. (2007) Conformational Transitions of Adenylate Kinase: Switching by Cracking. *J Mol Biol* **366**, 1661–1671.

- [83] Whitford, P. C., Gosavi, S., and Onuchic, J. N. (2007) Conformational Transitions in Adenylate Kinase: ALLOSTERIC COMMUNICATION REDUCES MISLIGATION. *J. Biol. Chem.* **283**, 2042–2048.
- [84] Hyeon, C., Jennings, P. A., Adams, J. A., Onuchic, J. N. (2009) Ligand-induced global transitions in the catalytic domain of protein kinase A. *Proc Natl Acad Sci USA* **106**, 3023–3028.
- [85] Jamros, M. A., Oliveira, L. C., Whitford, P. C., Onuchic, J. N., Adams, J. A., Jennings, P. A. (2012) Substrate-Specific Reorganization of the Conformational Ensemble of CSK Implicates Novel Modes of Kinase Function. *PLoS Comput Biol* **8**, e1002695.
- [86] Radou, G., Enciso, M., Krivov, S., Paci, E. (2013) Modulation of a protein free-energy landscape by circular permutation. *J Phys Chem B* **117**, 13743–13747.
- [87] Ferreira, D. U., Hegler, J. A., Komives, E. A., Wolynes, P. G. (2011) On the role of frustration in the energy landscapes of allosteric proteins. *Proc Natl Acad Sci USA* **108**, 3499–3503.
- [88] 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**, 183–193.
- [89] Wu, L. F., Reizer, A., Reizer, J., Cai, B., Tomich, J. M., Saier, M. H. (1991) Nucleotide sequence of the Rhodobacter capsulatus fruK gene, which encodes fructose-1-phosphate kinase: evidence for a kinase superfamily including both phosphofructokinases of Escherichia coli. *J. Bacteriol.* **173**, 3117–3127.
- [90] Bork, P., Sander, C., Valencia, A. (1993) Convergent evolution of similar enzymatic function on different protein folds: the hexokinase, ribokinase, and galactokinase families of sugar kinases. *Protein Sci.* **2**, 31–40.

- [91] Guixé, V., Babul, J. (1985) Effect of ATP on phosphofructokinase-2 from *Escherichia coli*. A mutant enzyme altered in the allosteric site for MgATP. *J Biol Chem* **260**, 11001–11005.
- [92] Ito, S., Fushinobu, S., Yoshioka, I., Koga, S., Matsuzawa, H., and Wakagi, T. (2001) Structural basis for the ADP-specificity of a novel glucokinase from a hyperthermophilic archaeon. *Structure* **9**, 205–214.
- [93] Rossmann, M. G., Moras, D., Olsen, K. W. (1974) Chemical and biological evolution of a nucleotide-binding protein. *Nature* **250**, 194–199.
- [94] Cabrera, R., Ambrosio, A. L. B., Garratt, R. C., Guixé, V., Babul, J. (2008) Crystallographic Structure of Phosphofructokinase-2 from *Escherichia coli* in Complex with Two ATP Molecules. Implications for Substrate Inhibition. *J Mol Biol* **383**, 588–602.
- [95] Zhang, Y., Dougherty, M., Downs, D. M., Ealick, S. E. (2004) Crystal structure of an aminoimidazole riboside kinase from *Salmonella enterica*: implications for the evolution of the ribokinase superfamily. *Structure* **12**, 1809–1821
- [96] Sigrell, J. A., Cameron, A. D., Mowbray, S. L. (1999) Induced fit on sugar binding activates ribokinase. *J Mol Biol* **290**, 1009–1018.
- [97] Cabrera, R., Babul, J., Guixé, V. (2010) Ribokinase family evolution and the role of conserved residues at the active site of the PfkB subfamily representative, Pfk-2 from *Escherichia coli*. *Arch Biochem Biophys* **502**, 23–30.
- [98] Miallau, L., Hunter, W. N., McSweeney, S. M., Leonard, G. A. (2007) Structures of *Staphylococcus aureus* D-tagatose-6-phosphate kinase implicate domain motions in specificity and mechanism. *J Biol Chem* **282**, 19948–19957.
- [99] Rivas-Pardo, J. A., Herrera-Morande, A., Castro-Fernandez, V., Fernandez, F. J., Vega, M. C., Guixé, V. (2013) Crystal structure, SAXS and kinetic mechanism of hyperthermophilic ADP-dependent glucokinase from *Thermococcus litoralis* reveal a conserved mechanism for catalysis. *PLoS ONE* **8**, e66687.



- [100] Campobasso, N., Mathews, I. I., Begley, T. P., Ealick, S. E. (2000) Crystal structure of 4-methyl-5-beta-hydroxyethylthiazole kinase from *Bacillus subtilis* at 1.5 Å resolution. *Biochemistry* **39**, 7868–7877.
- [101] Mathews, I. I., Erion, M. D., Ealick, S. E. (1998) Structure of Human Adenosine Kinase at 1.5 Å Resolution. *Biochemistry* **37**, 15607–15620.
- [102] Schumacher, M. A., Scott, D. M., Mathews, I. I., Ealick, S. E., Roos, D. S., Ullman, B., Brennan, R. G. (2000) Crystal structures of *Toxoplasma gondii* adenosine kinase reveal a novel catalytic mechanism and prodrug binding. *J Mol Biol* **296**, 549–567.
- [103] Currie, M. A., Merino, F., Skarina, T., Wong, A. H. Y., Singer, A., Brown, G., Savchenko, A., Caniuguir, A., Guixé, V., Yakunin, A. F., Jia, Z. (2009) ADP-dependent 6-phosphofructokinase from *Pyrococcus horikoshii* OT3: structure determination and biochemical characterization of PH1645. *J. Biol. Chem.* **284**, 22664–22671.
- [104] Cabrera, R., Baez, M., Pereira, H. M., Caniuguir, A., Garratt, R. C., Babul, J. (2011) The crystal complex of phosphofructokinase-2 of *Escherichia coli* with fructose-6-phosphate: kinetic and structural analysis of the allosteric ATP inhibition. *J. Biol. Chem.* **286**, 5774–5783.
- [105] Cheng, G., Bennett, E. M., Begley, T. P., Ealick, S. E. (2002) Crystal structure of 4-amino-5-hydroxymethyl-2-methylpyrimidine phosphate kinase from *Salmonella typhimurium* at 2.3 Å resolution. *Structure* **10**, 225–235.
- [106] Li, M.-H., Kwok, F., Chang, W.-R., Lau, C.-K., Zhang, J.-P., Lo, S. C. L., Jiang, T., Liang, D.-C. (2002) Crystal structure of brain pyridoxal kinase, a novel member of the ribokinase superfamily. *J Biol Chem* **277**, 46385–46390.
- [107] Gandhi, A. K., Desai, J. V., Ghatge, M. S., di Salvo, M. L., Di Biase, S., Danso-Danquah, R., Musayev, F. N., Contestabile, R., Schirch, V., Safo, M. K. (2012) Crystal structures of human pyridoxal kinase in complex with the neurotoxins,

ginkgotoxin and theophylline: insights into pyridoxal kinase inhibition. *PLoS ONE* **7**, e40954.

[108] Jeyakanthan, J., Thamotharan, S., Velmurugan, D., Rao, V. S. N., Nagarajan, S., Shinkai, A., Kuramitsu, S., Yokoyama, S. (2009) New structural insights and molecular-modelling studies of 4-methyl-5-beta-hydroxyethylthiazole kinase from *Pyrococcus horikoshii* OT3 (PhThiK). *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **65**, 978–986.

[109] Zhang, R.-G., Grembecka, J., Vinokour, E., Collart, F., Dementieva, I., Minor, W., Joachimiak, A. (2002) Structure of *Bacillus subtilis* YXKO—a member of the UPF0031 family and a putative kinase. *J. Struct. Biol.* **139**, 161–170.

[110] Ohshima, N., Inagaki, E., Yasuike, K., Takio, K., Tahirov, T. H. (2004) Structure of *Thermus thermophilus* 2-Keto-3-deoxygluconate kinase: evidence for recognition of an open chain substrate. *J Mol Biol* **340**, 477–489.

[111] Richardson, J. S., Getzoff, E. D., Richardson, D. C. (1978) The beta bulge: a common small unit of nonrepetitive protein structure. *Proc Natl Acad Sci USA* **75**, 2574–2578.

[112] Mathews, I. I., McMullan, D., Miller, M. D., Canaves, J. M., Elsliger, M.-A., Floyd, R., Grzechnik, S. K., Jaroszewski, L., Klock, H. E., Koesema, E., Kovarik, J. S., Kreuzsch, A., Kuhn, P., McPhillips, T. M., Morse, A. T., Quijano, K., Rife, C. L., Schwarzenbacher, R., Spraggon, G., Stevens, R. C., van den Bedem, H., Weekes, D., Wolf, G., Hodgson, K. O., Wooley, J., Deacon, A. M., Godzik, A., Lesley, S. A., Wilson, I. A. (2008) Crystal structure of 2-keto-3-deoxygluconate kinase (TM0067) from *Thermotoga maritima* at 2.05 Å resolution. *Proteins* **70**, 603–608.

[113] Potter, J. A., Kerou, M., Lambie, H. J., Bull, S. D., Hough, D. W., Danson, M. J., Taylor, G. L. (2008) The structure of *Sulfolobus solfataricus* 2-keto-3-deoxygluconate kinase. *Acta Crystallogr D Biol Crystallogr* **64**, 1283–1287.

[114] Baez, M., Babul, J. (2009) Reversible unfolding of dimeric phosphofructokinase-2 from *Escherichia coli* reveals a dominant role of inter-subunit contacts for stability. *FEBS Lett.* **583**, 2054–2060.

[115] Pineda, T., Churchich, J. E. (1993) Reversible unfolding of pyridoxal kinase. *J Biol Chem* **268**, 20218–20222.

[116] Kotlarz, D., Buc, H. (1977) Two *Escherichia coli* fructose-6-phosphate kinases. Preparative purification, oligomeric structure and immunological studies. *Biochim. Biophys. Acta* **484**, 35–48.

[117] Sigrell, J. A., Cameron, A. D., Jones, T. A., Mowbray, S. L. (1997) Purification, characterization, and crystallization of *Escherichia coli* ribokinase. *Protein Sci.* **6**, 2474–2476.

[118] Babul, J. (1978) Phosphofructokinases from *Escherichia coli*. Purification and characterization of the nonallosteric isozyme. *J Biol Chem* **253**, 4350–4355.

[119] Kotlarz, D., Buc, H. (1981) Regulatory properties of phosphofructokinase 2 from *Escherichia coli*. *Eur. J. Biochem.* **117**, 569–574.

[120] Guixé, V., Rodríguez, P. H., Babul, J. (1998) Ligand-Induced Conformational Transitions in *Escherichia coli* Phosphofructokinase 2: Evidence for an Allosteric Site for MgATP<sup>2-</sup>. *Biochemistry* **37**, 13269–13275.

[121] Guixé, V., Babul, J. (1988) Influence of ligands on the aggregation of the normal and mutant forms of phosphofructokinase 2 of *Escherichia coli*. *Arch Biochem Biophys* **264**, 519–524.

[122] Parducci, R. E., Cabrera, R., Baez, M., Guixé, V. (2006) Evidence for a catalytic Mg<sup>2+</sup> ion and effect of phosphate on the activity of *Escherichia coli* phosphofructokinase-2: regulatory properties of a ribokinase family member. *Biochemistry* **45**, 9291–9299.

- [123] Kotlarz, D., Garreau, H., Buc, H. (1975) Regulation of the amount and of the activity of phosphofructokinases and pyruvate kinases in *Escherichia coli*. *Biochim. Biophys. Acta* **381**, 257–268.
- [124] Caniuguir, A., Cabrera, R., Baez, M., Vásquez, C. C., Babul, J., Guixé, V. (2005) Role of Cys-295 on subunit interactions and allosteric regulation of phosphofructokinase-2 from *Escherichia coli*. *FEBS Lett.* **579**, 2313–2318.
- [125] Baez, M., Cabrera, R., Guixé, V., Babul, J. (2007) Unfolding pathway of the dimeric and tetrameric forms of phosphofructokinase-2 from *Escherichia coli*. *Biochemistry* **46**, 6141–6148.
- [126] Baez, M., Wilson, C. A. M., Babul, J. (2011) Folding kinetic pathway of phosphofructokinase-2 from *Escherichia coli*: A homodimeric enzyme with a complex domain organization. *FEBS Lett.* **585**, 2158–2164.
- [127] Tsai, C. J., Nussinov, R. (1997) Hydrophobic folding units at protein-protein interfaces: implications to protein folding and to protein-protein association. *Protein Sci.* **6**, 1426–1437.
- [128] Baez, M., Wilson, C. A. M., Ramírez-Sarmiento, C. A., Guixé, V., Babul, J. (2012) Expanded Monomeric Intermediate upon Cold and Heat Unfolding of Phosphofructokinase-2 from *Escherichia coli*. *Biophys J* **103**, 2187–2194.
- [129] Privalov, P. L. (1990) Cold Denaturation of Protein. *Crit Rev Biochem Mol Biol* **25**, 281–306.
- [130] Kunugi, S., Tanaka, N. (2002) Cold denaturation of proteins under high pressure. *Biochim Biophys Acta* **1595**, 329–344.
- [131] Tsai, C. J., Maizel, J. V., Nussinov, R. (2002) The hydrophobic effect: a new insight from cold denaturation and a two-state water structure. *Crit Rev Biochem Mol Biol* **37**, 55–69.
- [132] Dias, C. L., Ala-Nissila, T., Wong-ekkabut, J., Vattulainen, I., Grant, M., Karttunen, M. (2010) The hydrophobic effect and its role in cold denaturation. *Cryobiology* **60**, 91–99.

- [133] Adrover, M., Esposito, V., Martorell, G., Pastore, A., Temussi, P. A. (2010) Understanding Cold Denaturation: The Case Study of Yfh1. *J Am Chem Soc* **132**, 16240–16246.
- [134] Buchner, G. S., Shih, N., Reece, A. E., Niebling, S., Kubelka, J. (2012) Unusual cold denaturation of a small protein domain. *Biochemistry* **51**, 6496–6498.
- [135] Privalov, P. L., Griko YuV, Venyaminov SYu, Kutysenko, V. P. (1986) Cold denaturation of myoglobin. *J Mol Biol* **190**, 487–498.
- [136] Li, Y., Shan, B., Raleigh, D. P. (2007) The Cold Denatured State Is Compact but Expands at Low Temperatures: Hydrodynamic Properties of the Cold Denatured State of the C-terminal Domain of L9. *J Mol Biol* **368**, 256–262.
- [137] Richardson, J. M., Lemaire, S. D., Jacquot, J.-P., Makhatadze, G. I. (2000) Difference in the Mechanisms of the Cold and Heat Induced Unfolding of Thioredoxin h from *Chlamydomonas reinhardtii*: Spectroscopic and Calorimetric Studies. *J Am Chem Soc* **39**, 11154–11162.
- [138] Pometun, M. S., Peterson, R. W., Babu, C. R., Wand, A. J. (2006) Cold denaturation of encapsulated ubiquitin. *J Am Chem Soc* **128**, 10652–10653.
- [139] Luan, B., Shan, B., Baiz, C., Tokmakoff, A., Raleigh, D. P. (2013) Cooperative cold denaturation: the case of the C-terminal domain of ribosomal protein L9. *Biochemistry* **52**, 2402–2409.
- [140] Pastore, A., Martin, S. R., Politou, A., Kondapalli, K. C., Stemmler, T., Temussi, P. A. (2007) Unbiased Cold Denaturation: Low- and High-Temperature Unfolding of Yeast Frataxin under Physiological Conditions. *J Am Chem Soc* **129**, 5374–5375.
- [141] Romero-Romero, M. L., Inglés-Prieto, A., Ibarra-Molero, B., Sanchez-Ruiz, J. M. (2011) Highly anomalous energetics of protein cold denaturation linked to folding-unfolding kinetics. *PLoS ONE* **6**, e23050.

- [142] Damaschun, G., Damaschun, H., Gast, K., Misselwitz, R., Mueller, J. J., Pfeil, W., Zirwer, D. (1993) Cold denaturation-induced conformational changes in phosphoglycerate kinase from yeast. *Biochemistry* **32**, 7739–7746.
- [143] Lopez, C. F., Darst, R. K., Rossky, P. J. (2008) Mechanistic Elements of Protein Cold Denaturation. *J Am Chem Soc* **112**, 5961–5967.
- [144] Mandell, J. G., Falick, A. M., Komives, E. A. (1998) Identification of protein-protein interfaces by decreased amide proton solvent accessibility. *Proc Natl Acad Sci USA* **95**, 14705–14710.
- [145] Mandell, J. G., Baerga-Ortiz, A., Akashi, S., Takio, K., Komives, E. A. (2001) Solvent accessibility of the thrombin-thrombomodulin interface. *J Mol Biol* **306**, 575–589.
- [146] Truhlar, S. M. E., Torpey, J. W., Komives, E. A. (2006) Regions of IkappaBalpha that are critical for its inhibition of NF-kappaB•DNA interaction fold upon binding to NF-kappaB. *Proc Natl Acad Sci U S A* **103**, 18951–18956.
- [147] Truhlar, S. M. E., Croy, C. H., Torpey, J. W., Koeppe, J. R., Komives, E. A. (2006) Solvent accessibility of protein surfaces by amide H/2H exchange MALDI-TOF mass spectrometry. *J. Am. Soc. Mass Spectrom.* **17**, 1490–1497.
- [148] Ramírez-Sarmiento, C. A., Baez, M., Wilson, C. A. M., Babul, J., Komives, E. A., Guixé, V. (2013) Observation of Solvent Penetration during Cold Denaturation of *E. coli* Phosphofructokinase-2. *Biophys J* **104**, 2254–2263.
- [149] Dias, C. L., Ala-Nissila, T., Karttunen, M., Vattulainen, I., Grant, M. (2008) Microscopic mechanism for cold denaturation. *Phys. Rev. Lett.* **100**, 118101.
- [150] Schellman, J. A. (2002) Fifty years of solvent denaturation. *Biophys. Chem.* **96**, 91–101.
- [151] Kishore, D., Kundu, S., Kayastha, A. M. (2012) Thermal, chemical and pH induced denaturation of a multimeric  $\beta$ -galactosidase reveals multiple unfolding pathways. *PLoS ONE* **7**, e50380.

- [152] Dubey, V. K., Jagannadham, M. V. (2003) Differences in the unfolding of procerain induced by pH, guanidine hydrochloride, urea, and temperature. *Biochemistry* **42**, 12287–12297.
- [153] Yadav, S. C., Jagannadham, M. V. (2009) Complete conformational stability of kinetically stable dimeric serine protease milin against pH, temperature, urea, and proteolysis. *Eur. Biophys. J.* **38**, 981–991.
- [154] Oliveberg, M., Vuilleumier, S., Fersht, A. R. (1994) Thermodynamic study of the acid denaturation of barnase and its dependence on ionic strength: evidence for residual electrostatic interactions in the acid/thermally denatured state. *Biochemistry* **33**, 8826–8832.
- [155] Tanford, C. (1968) Protein denaturation. *Adv. Protein Chem.* **23**, 121–282.
- [156] Aune, K. C., Salahuddin, A., Zarlengo, M. H., Tanford, C. (1967) Evidence for residual structure in acid- and heat-denatured proteins. *J Biol Chem* **242**, 4486–4489.
- [157] Greene, R. F., Pace, C. N. (1974) Urea and guanidine hydrochloride denaturation of ribonuclease, lysozyme, alpha-chymotrypsin, and beta-lactoglobulin. *J Biol Chem* **249**, 5388–5393.
- [158] Pace, C. N., Laurents, D. V., Thomson, J. A. (1990) pH dependence of the urea and guanidine hydrochloride denaturation of ribonuclease A and ribonuclease T1. *Biochemistry* **29**, 2564–2572.
- [159] Herrmann, L. M., Bowler, B. E. (1997) Thermal denaturation of iso-1-cytochrome c variants: comparison with solvent denaturation. *Protein Sci.* **6**, 657–665.
- [160] Myers, J. K., Pace, C. N., Scholtz, J. M. (1995) Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Sci.* **4**, 2138–2148.

- [161] Tamura, Y., Gekko, K. (1995) Compactness of thermally and chemically denatured ribonuclease A as revealed by volume and compressibility. *Biochemistry* **34**, 1878–1884.
- [162] Kataoka, M., Nishii, I., Fujisawa, T., Ueki, T., Tokunaga, F., Goto, Y. (1995) Structural characterization of the molten globule and native states of apomyoglobin by solution X-ray scattering. *J Mol Biol* **249**, 215–228.
- [163] Konno, T., Kataoka, M., Kamatari, Y., Kanaori, K., Nosaka, A., Akasaka, K. (1995) Solution X-ray scattering analysis of cold- heat-, and urea-denatured states in a protein, Streptomyces subtilisin inhibitor. *J Mol Biol* **251**, 95–103.
- [164] Papagrigoriou, E., McEwan, P. A., Walsh, P. N., Emsley, J. (2006) Crystal structure of the factor XI zymogen reveals a pathway for transactivation. *Nat Struct Mol Biol* **13**, 557–558.
- [165] Bouma, B. N., Griffin, J. H. (1977) Human blood coagulation factor XI. Purification, properties, and mechanism of activation by activated factor XII. *J Biol Chem* **252**, 6432–6437.
- [166] Riley, P. W., Cheng, H., Samuel, D., Roder, H., Walsh, P. N. (2007) Dimer dissociation and unfolding mechanism of coagulation factor XI apple 4 domain: spectroscopic and mutational analysis. *J Mol Biol* **367**, 558–573.
- [167] Samuel, D., Cheng, H., Riley, P. W., Canutescu, A. A., Nagaswami, C., Weisel, J. W., Bu, Z., Walsh, P. N., Roder, H. (2007) Solution structure of the A4 domain of factor XI sheds light on the mechanism of zymogen activation. *Proc Natl Acad Sci U S A* **104**, 15693–15698.
- [168] Mann, C. J., Royer, C. A., Matthews, C. R. (1993) Tryptophan replacements in the trp aporepressor from Escherichia coli: probing the equilibrium and kinetic folding models. *Protein Sci.* **2**, 1853–1861
- [169] Perraut, C., Clottes, E., Leydier, C., Vial, C., Marcillat, O. (1998) Role of quaternary structure in muscle creatine kinase stability: tryptophan 210 is important for dimer cohesion. *Proteins* **32**, 43–51.



- [170] Scapin, G. (2013) Molecular replacement then and now. *Acta Crystallogr D Biol Crystallogr* **69**, 2266–2275.
- [171] Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**, 248–254.
- [172] Shuman, S. (1994) Novel approach to molecular cloning and polynucleotide synthesis using vaccinia DNA topoisomerase. *J Biol Chem* **269**, 32678–32684.
- [173] Mikol, V., Rodeau, J. L., Giegé, R. (1990) Experimental determination of water equilibration rates in the hanging drop method of protein crystallization. *Anal. Biochem.* **186**, 332–339.
- [174] Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R., Leslie, A. G. W. (2011) iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr D Biol Crystallogr* **67**, 271–281.
- [175] Evans, P. (2006) Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr* **62**, 72–82.
- [176] Evans, P. R. (2011) An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr D Biol Crystallogr* **67**, 282–292.
- [177] Potterton, E., Briggs, P., Turkenburg, M., Dodson, E. (2003) A graphical user interface to the CCP4 program suite. *Acta Crystallogr D Biol Crystallogr* **59**, 1131–1137.
- [178] Long, F., Vagin, A. A., Young, P., Murshudov, G. N. (2008) BALBES: a molecular-replacement pipeline. *Acta Crystallogr D Biol Crystallogr* **64**, 125–132.
- [179] Emsley, P., Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126–2132.
- [180] Debreczeni, J. É., Emsley, P. (2012) Handling ligands with Coot. *Acta Crystallogr D Biol Crystallogr* **68**, 425–430.

- [181] Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., Vagin, A. A. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* **67**, 355–367.
- [182] Kortemme, T., Baker, D. (2002) A simple physical model for binding energy hot spots in protein-protein complexes. *Proc Natl Acad Sci U S A* **99**, 14116–14121.
- [183] Kortemme, T., Kim, D. E., Baker, D. (2004) Computational alanine scanning of protein-protein interfaces. *Sci. STKE* **2004**, pl2.
- [184] Pei, J., Kim, B.-H., Grishin, N. V. (2008) PROMALS3D: a tool for multiple protein sequence and structure alignments. *Nucleic Acids Res.* **36**, 2295–2300.
- [185] Jenik, M., Parra, R. G., Radusky, L. G., Turjanski, A., Wolynes, P. G., Ferreiro, D. U. (2012) Protein frustratometer: a tool to localize energetic frustration in protein molecules. *Nucleic Acids Research* **40**, 348–351.
- [186] Schuck, P. (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys J* **78**, 1606–1619.
- [187] Houtman, J. C. D., Brown, P. H., Bowden, B., Yamaguchi, H., Appella, E., Samelson, L. E., Schuck, P. (2007) Studying multisite binary and ternary protein interactions by global analysis of isothermal titration calorimetry data in SEDPHAT: application to adaptor protein complexes in cell signaling. *Protein Sci.* **16**, 30–42.
- [188] Hughes, C. A., Mandell, J. G., Anand, G. S., Stock, A. M., Komives, E. A. (2001) Phosphorylation causes subtle changes in solvent accessibility at the interdomain interface of methyltransferase CheB. *J Mol Biol* **307**, 967–976.
- [189] Chalkley, R. J., Baker, P. R., Huang, L., Hansen, K. C., Allen, N. P., Rexach, M., Burlingame, A. L. (2005) Comprehensive analysis of a multidimensional liquid chromatography mass spectrometry dataset acquired on a quadrupole selecting, quadrupole collision cell, time-of-flight mass spectrometer: II. New developments in

Protein Prospector allow for reliable and comprehensive automatic analysis of large datasets. *Mol. Cell Proteomics* 4, 1194–1204.

[190] Mandell, J. G., Falick, A. M., Komives, E. A. (1998) Measurement of Amide Hydrogen Exchange by MALDI-TOF Mass Spectrometry. *Anal. Chem.* 70, 3987–3995.

[191] Kellermann, G., Vicentin, F., Tamura, E., Rocha, M., Tolentino, H., Barbosa, A., Craievich, A., Torriani, I. (1997) The Small-Angle X-ray Scattering Beamline of the Brazilian Synchrotron Light Laboratory. *J Appl Crystallogr* 30, 880–883.

[192] Guinier A. (1939) La diffraction des rayons X aux tres petits angles: applications a l'etude de phenomenes ultramicroscopiques. *Ann Phys* 12, 161–237.

[193] Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J., Svergun, D. I. (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J Appl Crystallogr* 36, 1277–1282.

[194] Svergun, D. I. (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J Appl Crystallogr* 25, 495–503.

[195] Svergun, D. I. (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys J* 76, 2879–2886.

[196] Svergun, D. I., Petoukhov, M. V., Koch, M. H. (2001) Determination of domain structure of proteins from X-ray solution scattering. *Biophysical Journal* 80, 2946–2953.

[197] Petoukhov, M. V., Franke, D., Shkumatov, A. V., Tria, G., Kikhney, A. G., Gajda, M., Gorba, C., Mertens, H. D. T., Konarev, P. V., Svergun, D. I. (2012) New developments in the ATSAS program package for small-angle scattering data analysis. *J Appl Crystallogr* 45, 342–350.

- [198] Svergun, D., Barberato, C., Koch, M. H. J. (1995) CRYSOLOG- a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *J Appl Crystallogr* **28**, 768–773.
- [199] Andersson, C. E., Mowbray, S. L. (2002) Activation of ribokinase by monovalent cations. *J Mol Biol* **315**, 409–419.
- [200] Ferreiro, D. U., Hegler, J. A., Komives, E. A., Wolynes, P. G. (2007) Localizing frustration in native proteins and protein assemblies. *Proc Natl Acad Sci USA* **104**, 19819–19824.
- [201] Sutto, L., Lätzer, J., Hegler, J. A., Ferreiro, D. U., Wolynes, P. G. (2007) Consequences of localized frustration for the folding mechanism of the IM7 protein. *Proc Natl Acad Sci USA* **104**, 19825–19830.
- [202] Hegler, J. A., Weinkam, P., Wolynes, P. G. (2008) The spectrum of biomolecular states and motions. *HFSP J* **2**, 307–313.
- [203] Ortega, A., Amorós, D., García de la Torre, J. (2011) Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models. *Biophys J* **101**, 892–898.
- [204] Perezzan, R., Rey, A. (2012) Simulating protein unfolding under pressure with a coarse-grained model. *J Chem Phys* **137**, 185102.
- [205] Rambo, R. P., Tainer, J. A. (2011) Characterizing flexible and intrinsically unstructured biological macromolecules by SAS using the Porod-Debye law. *Biopolymers* **95**, 559–571.
- [206] Kozin, M. B., Svergun, D. I. (2001) Automated matching of high- and low-resolution structural models. *J Appl Crystallogr* **34**, 33–41.
- [207] Bernadó, P., Mylonas, E., Petoukhov, M. V., Blackledge, M., Svergun, D. I. (2007) Structural characterization of flexible proteins using small-angle X-ray scattering. *J Am Chem Soc* **129**, 5656–5664.

- [208] Minton, A. P., Wilf, J. (1981) Effect of macromolecular crowding upon the structure and function of an enzyme: glyceraldehyde-3-phosphate dehydrogenase. *Biochemistry* **20**, 4821–4826.
- [209] Engel, R., Westphal, A. H., Huberts, D. H. E. W., Nabuurs, S. M., Lindhoud, S., Visser, A. J. W. G., van Mierlo, C. P. M. (2008) Macromolecular crowding compacts unfolded apoflavodoxin and causes severe aggregation of the off-pathway intermediate during apoflavodoxin folding. *J Biol Chem* **283**, 27383–27394.
- [210] Mikaelsson, T., Adén, J., Johansson, L. B.-Å., and Wittung-Stafshede, P. (2013) Direct observation of protein unfolded state compaction in the presence of macromolecular crowding. *Biophysical Journal* **104**, 694–704.
- [211] Das, P. P., and Matysiak, S. S. (2012) Direct Characterization of Hydrophobic Hydration during Cold and Pressure Denaturation. *J Phys Chem B* **116**, 5342–5348.
- [212] Scott, J. N., Nucci, N. V., Vanderkooi, J. M. (2008) Changes in water structure induced by the guanidinium cation and implications for protein denaturation. *J Phys Chem A* **112**, 10939–10948.
- [213] Mason, P. E., Neilson, G. W., Dempsey, C. E., Barnes, A. C., Cruickshank, J. M. (2003) The hydration structure of guanidinium and thiocyanate ions: implications for protein stability in aqueous solution. *Proc Natl Acad Sci U S A* **100**, 4557–4561.
- [214] Onuchic, J. N., Wolynes, P. G., Luthey-Schulten, Z., Socci, N. D. (1995) Toward an outline of the topography of a realistic protein-folding funnel. *Proc Natl Acad Sci U S A* **92**, 3626–3630.
- [215] Whitford, P. C., Noel, J. K., Gosavi, S., Schug, A., Sanbonmatsu, K. Y., Onuchic, J. N. (2009) An all-atom structure-based potential for proteins: bridging minimal models with all-atom empirical forcefields. *Proteins* **75**, 430–441.

- [216] Chavez, L. L., Onuchic, J. N., Clementi, C. (2004) Quantifying the roughness on the free energy landscape: entropic bottlenecks and protein folding rates. *J Am Chem Soc* **126**, 8426–8432.
- [217] Kouza, M., Li, M. S., O'brien, E. P., Hu, C.-K., Thirumalai, D. (2006) Effect of finite size on cooperativity and rates of protein folding. *J Phys Chem A* **110**, 671–676.
- [218] Whitford, P. C., Geggier, P., Altman, R. B., Blanchard, S. C., Onuchic, J. N., Sanbonmatsu, K. Y. (2010) Accommodation of aminoacyl-tRNA into the ribosome involves reversible excursions along multiple pathways. *RNA* **16**, 1196–1204.
- [219] Lammert, H., Schug, A., Onuchic, J. N. (2009) Robustness and generalization of structure-based models for protein folding and function. *Proteins* **77**, 881–891.
- [220] Dias, C. (2012) Unifying Microscopic Mechanism for Pressure and Cold Denaturations of Proteins. *Phys. Rev. Lett.* **109**, 048104.
- [221] Cheung, M. S., Garcia, A. E., Onuchic, J. N. (2002) Protein folding mediated by solvation: water expulsion and formation of the hydrophobic core occur after the structural collapse. *Proc Natl Acad Sci U S A* **99**, 685–690.
- [222] Levy, Y., Onuchic, J. N. (2006) Water mediation in protein folding and molecular recognition. *Annu Rev Biophys Biomol Struct* **35**, 389–415.
- [223] Harpaz, Y., Elmasry, N., Fersht, A. R., Henrick, K. (1994) Direct observation of better hydration at the N terminus of an alpha-helix with glycine rather than alanine as the N-cap residue. *Proc Natl Acad Sci U S A* **91**, 311–315.
- [224] Xu, W., Harrison, S. C., Eck, M. J. (1997) Three-dimensional structure of the tyrosine kinase c-Src. *Nature* **385**, 595–602.
- [225] Vijay-Kumar, S., Bugg, C. E., Cook, W. J. (1987) Structure of ubiquitin refined at 1.8 Å resolution. *J Mol Biol* **194**, 531–544.

- [226] Gronenborn, A. M., Filpula, D. R., Essig, N. Z., Achari, A., Whitlow, M., Wingfield, P. T., Clore, G. M. (1991) A novel, highly stable fold of the immunoglobulin binding domain of streptococcal protein G. *Science* **253**, 657–661.
- [227] Cao, P., Gong, Y., Tang, L., Leung, Y.-C., Jiang, T. (2006) Crystal structure of human pyridoxal kinase. *J. Struct. Biol.* **154**, 327–332.
- [228] Noel, J. K., Whitford, P. C., Sanbonmatsu, K. Y., Onuchic, J. N. (2010) SMOG@ctbp: simplified deployment of structure-based models in GROMACS. *Nucleic Acids Research* **38**, W657–W661.
- [229] Pronk, S., Páll, S., Schulz, R., Larsson, P., Bjelkmar, P., Apostolov, R., Shirts, M. R., Smith, J. C., Kasson, P. M., van der Spoel, D., Hess, B., Lindahl, E. (2013) GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **29**, 845–854.
- [230] Sugita, Y., Okamoto, Y. (1999) Replica-exchange molecular dynamics method for protein folding. *Chem Phys Lett* **314**, 141–151
- [231] Torrie, G. M., Valleau, J. P. (1977) Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *J Comput Phys* **23**, 187–199.
- [232] Liu, Z., Chan, H. S. (2005) Solvation and desolvation effects in protein folding: native flexibility, kinetic cooperativity and enthalpic barriers under isostability conditions. *Phys Biol* **2**, S75–S85.
- [233] Ferrenberg, A., Swendsen, R. (1988) New Monte Carlo technique for studying phase transitions. *Phys. Rev. Lett.* **61**, 2635–2638.
- [234] Ferrenberg, A., Swendsen, R. (1989) Optimized Monte Carlo data analysis. *Phys. Rev. Lett.* **63**, 1195–1198.
- [235] Kumar, S., Rosenberg, J. M., Bouzida, D., Swendsen, R. H., Kollman, P. A. (1992) THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J Comput Chem* **13**, 1011–1021.

- [236] Roux, B. (1995) The calculation of the potential of mean force using computer simulations. *Computer Physics Communications* **91**, 275–282.
- [237] Zhang, J., Qin, M., Wang, W. (2005) Multiple folding mechanisms of protein ubiquitin. *Proteins* **59**, 565–579.
- [238] Lee, H. S., Moon, B. K., Choi, S. Y., Kwon, O. S. (2000) Human pyridoxal kinase: overexpression and properties of the recombinant enzyme. *Mol Cells* **10**, 452–459.
- [239] Wilton, D. J., Tunnicliffe, R. B., Kamatari, Y. O., Akasaka, K., Williamson, M. P. (2008) Pressure-induced changes in the solution structure of the GB1 domain of protein G. *Proteins* **71**, 1432–1440.
- [240] Vajpai, N., Nisius, L., Wiktor, M., Grzesiek, S. (2013) High-pressure NMR reveals close similarity between cold and alcohol protein denaturation in ubiquitin. *Proc Natl Acad Sci U S A* **110**, E368–76.
- [241] Zheng, W., Tekpinar, M. (2011) Accurate flexible fitting of high-resolution protein structures to small-angle x-ray scattering data using a coarse-grained model with implicit hydration shell. *Biophysical Journal* **101**, 2981–2991.
- [242] Yang, S., Blachowicz, L., Makowski, L., Roux, B. (2010) Multidomain assembled states of Hck tyrosine kinase in solution. *Proc Natl Acad Sci* **107**, 15757–15762.
- [243] Daily, M. D., Makowski, L., Phillips, G. N., Cui, Q. (2012) Large-scale motions in the adenylate kinase solution ensemble: coarse-grained simulations and comparison with solution X-ray scattering. *Chemical Physics* **396**, 84–91.
- [244] Jamros, M. A., Oliveira, L. C., Whitford, P. C., Onuchic, J. N., Adams, J. A., Blumenthal, D. K., Jennings, P. A. (2010) Proteins at work: a combined small angle X-RAY scattering and theoretical determination of the multiple structures involved on the protein kinase functional landscape. *J. Biol. Chem.* **285**, 36121–36128.