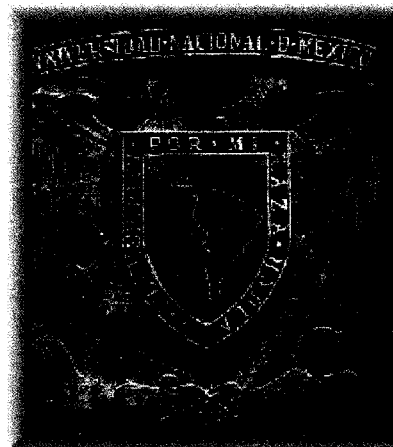
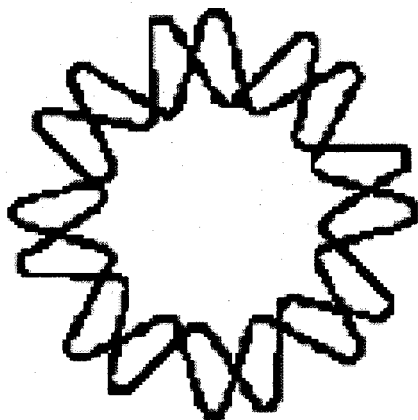


UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

INSTITUTO DE BIOTECNOLOGÍA



**ANÁLISIS DINÁMICO DEL CAMBIO ALOSTÉRICO DE LA GLUCOSAMINA-6-FOSFATO  
DESAMINASA DE *ESCHERICHIA COLI***

Tesis de Maestría en Ciencias Bioquímicas que presenta:

**JONATHAN VALENCIA S.**

Cuernavaca, Morelos. Septiembre de 2006.



Universidad Nacional  
Autónoma de México



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

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

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

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

Autorizo a la Dirección General de Bibliotecas de la UNAM a atender en formato electrónico e impreso el contenido de mi trabajo recepcional.

NOMBRE: Jonathan Valencia Swain

FECHA: \_\_\_\_\_

FIRMA: Jonathan Valencia S

Este trabajo se realizó bajo la dirección del Dr. Eduardo Horjales Reboredo en el Departamento de Medicina Molecular y Bioprocesos, Instituto de Biotecnología. Durante el proyecto de maestría Jonathan Valencia S. recibió una beca del CONACyT y un complemento de beca de la DGEP, UNAM.

Por llevarme de una etapa a otra ayer, hoy y mañana,  
le dedico a mi familia esta, y las que vendrán.

## AGRADECIMIENTOS

Al Instituto de Biotecnología de la UNAM.

Al CONACyT, por la beca otorgada para cursar esta maestría.

A la DGEP de la UNAM, por contribuir con un complemento de beca.

Al programa de Maestría y Doctorado en Ciencias Bioquímicas.

Al Dr. Eduardo Horjales Reboredo, por dirigir este trabajo.

A los integrantes del comité tutorial: Dra. Nina Pastor Colón, Dr. Lorenzo Segovia Forcella y Dr. Eduardo Horjales Reboredo, por su guía durante el desarrollo de este trabajo.

A los miembros del jurado: Dr. Baltazar Becerril Luján, Dr. Eduardo Horjales Reboredo, Dr. Arturo Rojo Domínguez, Dr. Jean Louis Charli Casalonga y Dr. Francisco Barona, por sus aportaciones sobre esta tesis.

A quienes fueron mis profesores en la maestría, por mostrarme otros aspectos interesantes de la biología.

A quienes son o han sido integrantes del grupo del Dr. Eduardo Horjales Reboredo durante mi estancia, por compartir su conocimiento y facilitar mi incursión en el mundo de la Bioquímica, en especial a Yagul Pedraza Pérez, Álvaro Resines Sierra, Dr. Enrique Rudiño Piñera, Dra. Brenda Valderrama Blanco, Eugenio De la Mora Lugo, Paloma Gil Rodríguez, Aisel Valle Garay, Sonia Rojas Trejo y Leopoldo Guereca Gurrola.

A la Unidad de Cómputo, por su apoyo en mis batallas con la tecnología, en especial a Juan Manuel Hurtado Ramírez, Roberto Rodríguez Bahena y Jérôme Verleyen.

A los estudiantes de maestría de mi generación, por darle la bienvenida a un físico.

Al grupo del Dr. Lorenzo Segovia Forcella, por darle sabor a la vida del Laboratorio 8.

A La Cofradía, por llenar lugares tan grandes que de otra forma habrían estado casi vacíos.

A las amistades de pasillo, por romper con la cotidianidad.

A la internet de banda ancha del Instituto de Biotecnología, por el acceso a material puramente académico y estrictamente relacionado con este trabajo.

A Gata y Gata, por darme la bienvenida cada vez que regreso a mi casa.

Especialmente, gracias a Patricia Oliver Ocaño, por su amistad y su compañía durante tan poco y tan largo tiempo :\*

# CONTENIDO

1. RESUMEN	1
2. INTRODUCCIÓN	3
2.1. Contexto de la glucosamina-6-fosfato desaminasa de <i>Escherichia coli</i>	3
2.2. Enzimas alostéricas	4
2.2.1. Modelo alostérico concertado	5
2.3. Estructura protomérica de la GlcN6P desaminasa	6
3. ANTECEDENTES	9
3.1. Flexibilidad de la tapa del sitio activo	9
3.2. Forma T con el sitio activo ocupado no observada	9
3.3. Flexibilidad de la forma T libre de ligantes	10
3.4. Modelo alostérico de la GlcN6P desaminasa de <i>E. coli</i>	12
3.5. GlcN6P desaminasa en el contexto de otros sistemas alostéricos	15
4. HIPÓTESIS	16
5. OBJETIVOS	17
5.1. Objetivos generales	17
5.2. Objetivos particulares	17
6. MARCO TEÓRICO	18
6.1. Simulaciones	18
6.2. Dinámica Molecular	18
6.2.1. Inicialización	18
6.2.1.1. Minimización	19
6.2.2. Calentamiento	20
6.2.3. Equilibrio	20
6.2.4. Producción	21
6.3. Condiciones de frontera periódicas	21
6.4. Potencial de interacción	21
6.5. Ensamble termodinámico	23
6.5.1. Presión constante	23
6.5.2. Temperatura constante	24
6.6. Algoritmo de integración	25
6.6.1. Tamaño de paso	25

6.6.1.1. Algoritmo SHAKE	26
7. METODOLOGÍA	27
7.1. Generación de las simulaciones	27
7.1.1. Construcción del sistema	27
7.1.2. Preparación previa a la fase de producción	29
7.1.3. Imposiciones termodinámicas	30
7.1.4. Propagación del movimiento	30
7.2. Métodos de análisis	30
7.2.1. Cuantificación de la diferencia entre estructuras: RMSD	30
7.2.2. Análisis por proyección	31
7.2.3. Análisis de Fourier	31
7.2.4. Modos cuasiarmónicos	34
8. RESULTADOS Y DISCUSIONES	36
8.1. Resultados del equilibrio	36
8.2. Validación del análisis por proyección	37
8.3. Conformación R desfavorecida en el cristal original	39
8.4. Períodos de oscilación	41
8.5. Formación del sitio activo y del sitio alostérico	46
8.6. Estabilización de la tapa abierta mediante una molécula de agua	49
8.7. Transmisión de los efectos de la tapa al resto de la proteína	51
9. CONCLUSIONES	55
10. REFERENCIAS	56
11. BIBLIOGRAFÍA	60



## ÍNDICE DE FIGURAS

Figura 1. Traza del hexámero de la GlcN6P desaminasa libre de ligantes	3
Figura 2. Vía de degradación de la glucosamina	3
Figura 3. Reacción que cataliza la GlcN6P desaminasa	4
Figura 4. Modelo alostérico concertado	6
Figura 5. División estructural de la GlcN6P desaminasa	7
Figura 6. Transición alostérica de la GlcN6P desaminasa	7
Figura 7. Tres conformaciones alternativas de la tapa del sitio activo en la forma R heterotrópica	9
Figura 8. Factores B anisotrópicos de la traza de un protómero	11
Figura 9. Conformaciones alternativas para la forma T	11
Figura 10. Modelo alostérico de la GlcN6P desaminasa para la transición heterotrópica	13
Figura 11. Modelo alostérico de la GlcN6P desaminasa	14
Figura 12. Contracción de la celda del cristal después del equilibrio	12
Figura 13. Ejemplo de un conjunto de datos con varios períodos	32
Figura 14. Ejemplo de un análisis de Fourier	33
Figura 15. Ejemplo de un análisis de Fourier con sus intervalos de error	34
Figura 16. Energía total de las simulaciones mostrando el equilibrio de cada sistema	36
Figura 17. Estabilización del valor del parámetro de red $a$ para cada simulación	37
Figura 18. Validación del análisis por proyección mediante la RMSD	38
Figura 19. Análisis por proyección sobre el protómero de las simulaciones Crist14 y Crist06	40
Figura 20. Interacciones cristalinas en los conformeros T y R en el grupo espacial $P6_322$	41
Figura 21. Análisis de Fourier sobre el protómero de la simulación Sol28	42
Figura 22. Análisis por proyección sobre distintas zonas de la proteína en la simulación Sol28	43
Figura 23. Análisis de Fourier sobre distintas zonas de la proteína en la simulación Sol28	43
Figura 24. Análisis por proyección sobre el protómero de la simulación Sol10	44
Figura 25. Análisis de Fourier sobre el protómero de las simulaciones Sol28 y Sol10	44
Figura 26. Análisis por proyección sobre la tapa del sitio activo en las simulaciones Sol28 y Sol10	45
Figura 27. Análisis de Fourier sobre la tapa del sitio activo en las simulaciones Sol28 y Sol10	46
Figura 28. Formación del sitio activo a lo largo de Sol28	47
Figura 29. Formación del sitio activo a lo largo de Sol10	47
Figura 30. Formación del sitio alostérico a lo largo de Sol28	48

Figura 31. Formación del sitio alostérico a lo largo de Sol10 _____	49
Figura 32. Tapa abierta del sitio activo estabilizada por Tyr254 mediante una molécula de agua ____	50
Figura 33. Tapa cerrada del sitio activo no estabilizada por Tyr254 _____	51
Figura 34. Efecto de la tapa del sitio activo al resto de la proteína a través de Arg156 _____	52
Figura 35. Modo cuasiarmónico más representativo de Sol28, visto a lo largo del eje de orden tres __	53
Figura 36. Modo cuasiarmónico más representativo de Sol28, visto a lo largo del eje de orden dos __	53
Figura 37. Segundo modo cuasiarmónico más representativo de Sol28 _____	54

## ÍNDICE DE TABLAS

Tabla I. Estados de protonación definidos para las histidinas de cada modelo construido _____	28
Tabla II. Intervalos sobre los que se realizó el análisis de Fourier en Sol28 _____	42
Tabla III. Intervalos de la simulación Sol28 con conformaciones de interés _____	50

# 1. RESUMEN

La glucosamina-6-fosfato desaminasa de *Escherichia coli* es una enzima alostérica involucrada en el catabolismo de amino-azúcares. La unidad biológica es un hexámero, cada uno de cuyos protómeros puede subdividirse en un módulo interno (encargado de la cohesión del oligómero) y un módulo externo (que contiene la tapa del sitio activo). Las estructuras cristalográficas de la enzima con y sin ligantes (estructuras R y T, respectivamente) muestran que la transición alostérica corresponde a una rotación de cada módulo externo alrededor de su respectivo módulo interno. Los resultados de cristalografía sugieren que, gracias a su flexibilidad, la enzima libre de ligantes presenta una oscilación estructural en dirección de la transición alostérica.

El objetivo de este trabajo es observar en simulaciones de Dinámica Molecular el estado cinético de la enzima libre de ligantes en el cristal y en solución, así como estimar la frecuencia de sus oscilaciones. Para ello se generaron cuatro simulaciones (dos dentro y dos fuera de la matriz cristalina) mediante el programa CHARMM empleando el potencial CHARMM27. El ensamble impuesto fue el isotérmico-isobárico, manteniendo una temperatura de 300 K con el algoritmo de Hoover, y una presión de 1 atm mediante un pistón de Langevin. El tamaño de paso fue de 2 fs gracias a que se implementó el algoritmo SHAKE. Se impusieron condiciones de frontera periódicas, generando el hexámero mediante operaciones de simetría. Las interacciones de largo alcance se simplificaron siguiendo un esquema de truncamiento.

El estado de la proteína con respecto a las estructuras T y R se cuantificó proyectando su conformación sobre el conjunto de vectores que conecta ambas estructuras de referencia. En ambas simulaciones del cristal (de 14 ns y 6 ns) este análisis revela que la conformación activa (R) se ve desfavorecida en el grupo espacial P6<sub>3</sub>22 debido a choques estéricos interhexaméricos. En las simulaciones de la molécula fuera del cristal (de 28 ns y 10 ns) sí se observa la oscilación conformacional esperada, y este es su movimiento más representativo. El análisis de Fourier de este movimiento global revela que se trata de una superposición de oscilaciones con frecuencias del orden de cientos de MHz, mientras que algunos de los movimientos de la tapa del sitio activo son del orden de GHz. Esto sugiere que estos últimos son distintos a los globales, a pesar de que se encontró que existe una transferencia de los efectos del estado de apertura de la tapa al resto de la proteína, mediada por Arg156.

En estas dos simulaciones se observa también que el sitio activo requiere de la presencia de algún ligante para permanecer en su conformación activa una vez formado. Sin embargo, no se observa la formación de la cavidad del sitio alostérico en ausencia de ligantes que neutralicen las cargas

positivas de sus paredes. Además, una vez que la tapa del sitio activo se ha abierto, Tyr254 puede estabilizarla en este estado por medio de interacciones electrostáticas a través de una molécula de agua.

En conclusión, este trabajo hace una descripción teórica de algunos aspectos del mecanismo alostérico de la glucosamina-6-fosfato desaminasa.

## 2. INTRODUCCIÓN

### 2.1. Contexto de la glucosamina-6-fosfato desaminasa de *Escherichia coli*

La forma biológicamente funcional de la proteína glucosamina-6-fosfato desaminasa (GlcN6P desaminasa) de *E. coli* (cepa K-12) es oligomérica, y en particular es un homohexámero [1]. Cada una de sus seis subunidades (llamadas protómeros, pues están asociadas y son idénticas) es un monómero NagB de 266 residuos de aminoácidos (29774Da) codificado por el gen *nagB* [2]. Actualmente se conoce su estructura tridimensional en diferentes complejos [3] [4] y libre de ligantes [5] [4] (figura 1).

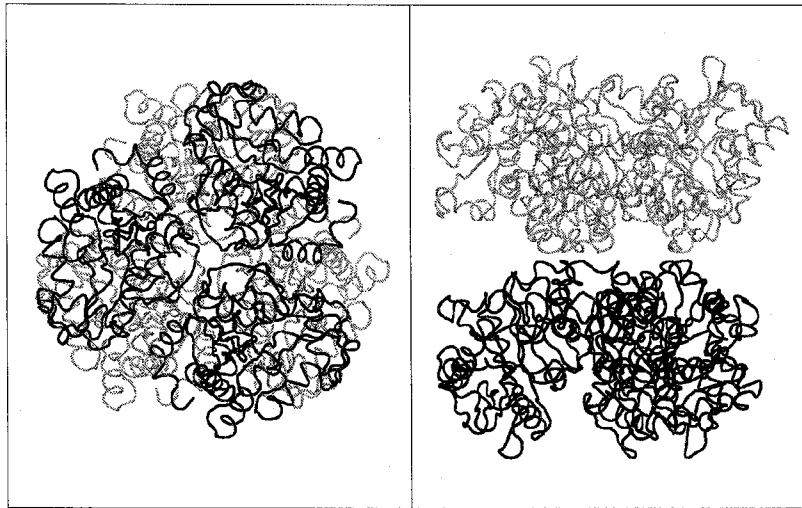


Figura 1. Trazo del hexámero de GlcN6P desaminasa libre de ligantes, vista a lo largo de su eje de simetría de orden 3 (izquierda) y a lo largo de su eje de simetría de orden 2 (derecha). El trímero superior se muestra en color gris y, el inferior, en color negro. La estructura cristalográfica es la reportada en la referencia [4] (imagen generada en VMD [6]).

Esta enzima participa en la vía de degradación de la glucosamina (figura 2).

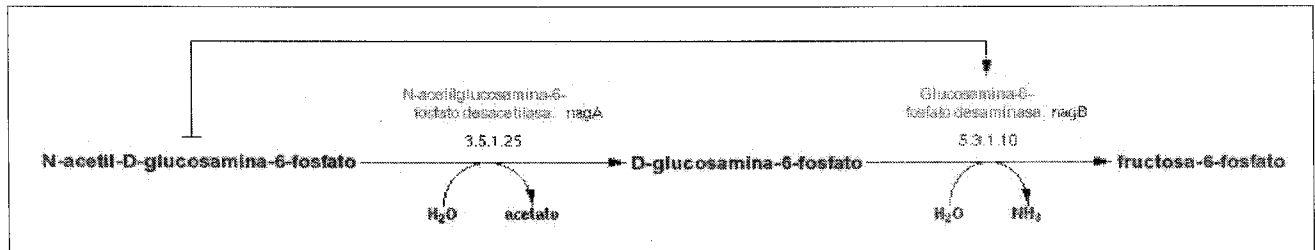


Figura 2. Vía de degradación de la glucosamina. La flecha que conecta la N-acetil-D-glucosamina-6-fosfato con la enzima glucosamina-6-fosfato desaminasa indica que la primera es un activador de la segunda (ver más adelante).

Particularmente, cataliza la reacción de desaminación de la D-glucosamina-6-fosfato, que es el paso regulador de la vía metabólica (figura 3).

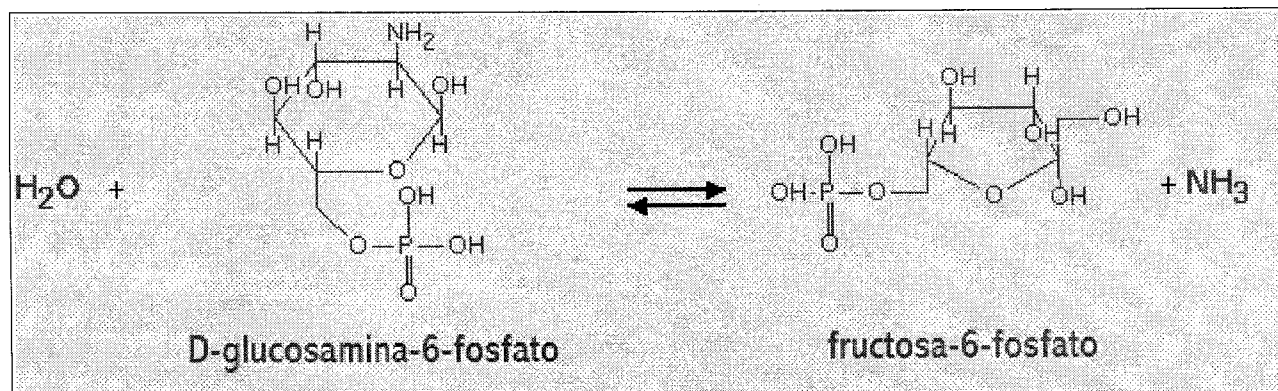


Figura 3. Reacción que cataliza la GlcN6P desaminasa.

Dicha regulación es posible gracias a los mecanismos de modulación de actividad con que cuenta esta proteína (que se describirán más adelante). Así, la GlcN6P desaminasa juega un papel central en el catabolismo de amino-azúcares al ser la única enzima alostérica de la ruta metabólica [4].

## 2.2. Enzimas alostéricas

Las enzimas de este tipo poseen propiedades de regulación intrínsecas dependientes de la unión de una molécula pequeña (ligante) en un sitio de unión de la macromolécula. Esta ligación tiene como efecto aumentar o disminuir la afinidad por otras moléculas en otros sitios.

La relevancia fisiológica de esto se hace evidente cuando la afinidad que es modulada es la de los sitios activos (sitios donde se une el sustrato de la reacción que es catalizada) por el sustrato, de manera que la proteína es llevada a una conformación ya sea más o menos activa, es decir, que puede o no ser capaz de unir la molécula de sustrato y, por lo tanto, llevar a cabo su función biológica. Así, este mecanismo (llamado alosterismo) regula la actividad enzimática en algunas vías metabólicas de los organismos, con un tiempo de respuesta corto.

Las moléculas pequeñas que desencadenan este proceso en las enzimas alostéricas son llamadas moléculas efectoras (o simplemente efectores), y llevan a cabo este papel uniéndose ya sea a un sitio activo o a un sitio alostérico. Este último se encuentra sólo en las enzimas alostéricas, y consiste en un sitio de unión que no participa en la catálisis, sino que tiene como función desencadenar el mecanismo alostérico arriba descrito. Este tipo de enzimas usualmente son oligómeros, y sus sitios alostéricos se encuentran en las cavidades que hay entre los protómeros que las forman.

Los efectores fisiológicos que pueden unirse en los sitios activos son las moléculas de sustrato, mientras que los sitios alostéricos pueden ser afines a otras moléculas que estén involucradas en las reacciones de la vía metabólica de la enzima (o incluso moléculas ajenas a la vía) y, dependiendo de su efecto sobre la actividad de la enzima, pueden catalogarse como activadores o inactivadores.

En el contexto de las enzimas alostéricas, se le llama efecto heterotrópico al que un efector ejerce sobre la afinidad a *otros* ligantes en otros sitios de unión. Si, por el contrario, se modula en otros sitios de unión la afinidad por ligantes que son *iguales* al efector, se le llama efecto homotrópico. En este caso, existe cooperatividad positiva si la afinidad aumenta, o cooperatividad negativa si la afinidad disminuye.

Aún queda por exponer el mecanismo base del alosterismo, es decir, el proceso de comunicación entre los sitios activos y alostéricos de una macromolécula. Existen dos modelos distintos que intentan explicar este problema, y que recurren a la flexibilidad de las proteínas. Uno de ellos, el modelo secuencial o modelo KNF [7], no es consistente con los datos de cooperatividad y activación de la GlcN6P desaminasa [8] ya que implica que, contrario a las observaciones experimentales, la unión de un ligante a un sitio de unión siempre alterará la afinidad en un sitio de unión vecino a través de cambios conformacionales inducidos por el efector. El otro esquema propuesto es el modelo concertado, y se describe a continuación.

### **2.2.1. Modelo alostérico concertado**

Este modelo, también llamado modelo MWC, modelo simétrico o modelo de dos estados fue propuesto en 1965 por Monod y colaboradores [9]. A partir del hecho de que las transiciones alostéricas corresponden a grandes desplazamientos de subunidades poco alteradas (principalmente cambios en la estructura cuaternaria), propone que la afinidad en los sitios de unión en realidad depende de la conformación cuaternaria que presente la proteína (con mayor o menor afinidad por los ligantes, con todas las subunidades en la misma conformación), y que la afinidad además no es modulada directamente por el ligante alostérico, sino que el efecto de este consiste en desplazar el equilibrio entre las distintas conformaciones de la proteína, las cuales coexisten en solución (figura 4). Bajo este contexto, a la conformación con mayor afinidad por los ligantes se le llama conformación activa o conformación R ("relajada", sin constricciones) y, a la de menor afinidad por los ligantes, conformación T ("tensa", asumiendo que está de alguna forma constreñida).



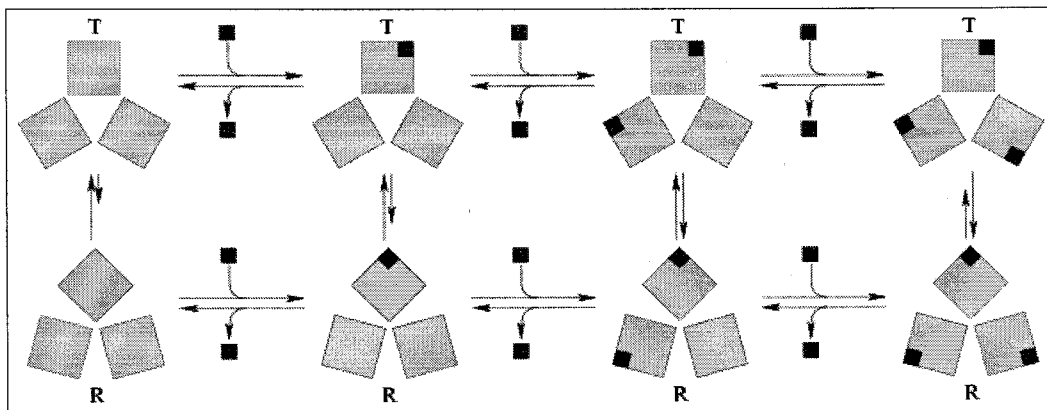


Figura 4. Modelo alostérico concertado, representado para una proteína de tres subunidades. Los cuadros grises representan protómeros, mientras que los cuadros negros corresponden a ligandos activadores. Se representan arbitrariamente los tamaños relativos de las flechas verticales que indican la interconversión entre los conformeros T (menos afin) y R (más afin), y con ellas sólo se pretende ilustrar el desplazamiento gradual del equilibrio hacia R con la adición de activadores.

### 2.3. Estructura protomérica de la GlcN6P desaminasa

De acuerdo con las divisiones asignadas a la estructura de la enzima que han sido propuestas anteriormente [3] [5] [4], cada protómero de esta enzima puede subdividirse en un módulo interno y un módulo externo, los cuales se detallan a continuación (figura 5).

El módulo interno abarca los residuos 137 a 143, 207 a 233 y 244 a 253, y es responsable de la cohesión del hexámero al participar en la mayoría de los contactos interprotoméricos por lo que, como se verá más adelante, tiene poca movilidad. El módulo externo es el complemento del anterior, pues abarca los residuos 1 a 136 (conteniendo el extremo amino terminal), 144 a 206, 234 a 243 y 254 a 266 (conteniendo el extremo carboxilo terminal). En cada protómero ambas secciones participan en la formación del sitio activo.

De acuerdo a las estructuras cristalográficas de la proteína en sus conformaciones T y R, el cambio conformacional alostérico ocurre principalmente en el módulo externo, presentando una rotación de  $11^\circ$  alrededor del módulo interno [5] (figura 6). El módulo externo, además, contiene una tapa de gran flexibilidad (residuos 161 a 184) que cubre el sitio activo, así como dos asas (residuos 75 a 83 y 109 a 112) cuyo movimiento está correlacionado con el de dicha cubierta. Por último, cada módulo externo define en parte dos de los sitios alostéricos del hexámero, pues estas cavidades se encuentran en la región interprotomérica de esta sección.



Figura 5. Trazas de una subunidad del trímero inferior del hexámero de GlcN6P desaminasa libre de ligantes (cfr. figura 1), vista a lo largo del eje de simetría de orden 3 del hexámero (izquierda) y a lo largo del eje de orden 2 del hexámero (derecha). El código de colores es: azul para el módulo interno, amarillo para el módulo externo, magenta para la tapa del sitio activo, y naranja para las dos asas que están relacionadas con la tapa. El sitio activo está el espacio entre la tapa y el resto del portómoro, y se aprecia mejor en el panel derecho. Imagen generada en VMD [6].

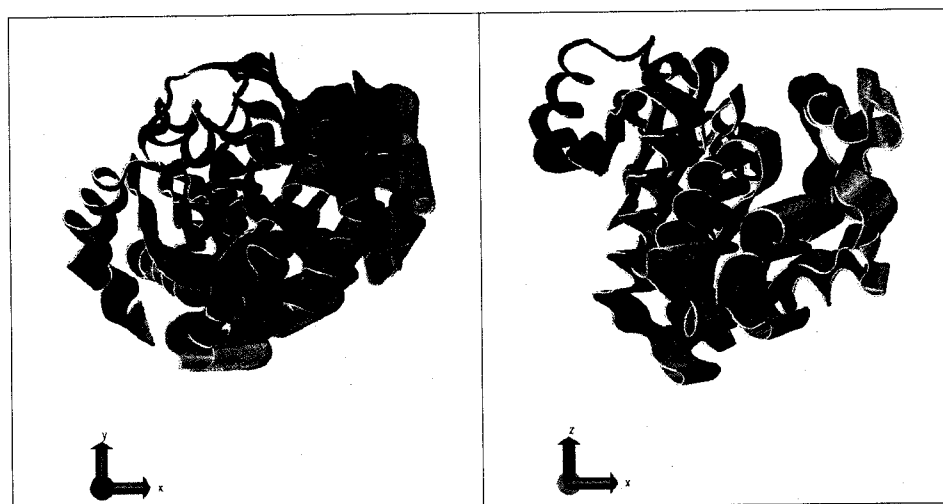


Figura 6. Superposición de las trazas de las conformaciones de la enzima a lo largo de la transición de su conformación menos afin (blanco) a su conformación activa (negro), vistas a lo largo de su eje de orden 3 (izquierda) y de orden 2 (derecha). En las conformaciones intermedias la distancia recorrida por cada  $C^{\alpha}$  es proporcional a la longitud de onda de su color. Imagen generada en VMD [6].

Así, el módulo externo es el responsable del cambio alostérico heterotrópico de la enzima al unir en el sitio alostérico el activador *N*-acetil-glucosamina-6-fosfato (GlcNAc6P), cuya constante de afinidad es  $K_m=[180\pm 50]\mu\text{M}$  [10]. En cambio, en el efecto homotrópico participan ambas secciones al unir glucosamina-6-fosfato (GlcN6P;  $K_m=[5.5\pm 0.2]\text{mM}$  en T,  $K_m=[0.75\pm 0.05]\text{mM}$  en R [10]) o

fructosa-6-fosfato (Fru6P;  $K_m=[0.9\pm 0.1]$ mM en R [10]) en el sitio activo. Este activador y estos dos sustratos son los ligantes biológicos conocidos para la enzima.

### 3. ANTECEDENTES

Existe un análisis previo en términos de flexibilidad y vibraciones atómicas del cambio alostérico de la GlcN6P desaminasa, que utiliza técnicas cristalográficas [4]. En él se describe la relación entre la transición alostérica, los cambios en la flexibilidad molecular y la función biológica de la proteína. A continuación se resumen algunos de sus resultados más relevantes para este trabajo.

#### 3.1. Flexibilidad de la tapa del sitio activo

La movilidad de la tapa del sitio activo de la GlcN6P desaminasa depende de la conformación que presente la enzima. En la forma R con ambos sitios de unión (activo y alostérico) ocupados la tapa está estabilizada en su conformación cerrada tanto por la molécula de sustrato como por la cadena lateral del residuo Glu148, pues este forma puentes de hidrógeno con los residuos Thr166 y Thr163 de la tapa, y con His143 del sitio activo. En cambio, en la forma T no existe un ligante en el sitio activo, y Glu148 apunta al sitio alostérico [5] equilibrando las cargas de esa cavidad, por lo que no contribuye a la estabilización de la tapa, y esta presenta una gran flexibilidad.

En la forma R producto de la activación heterotrópica, es decir, con sólo el sitio alostérico ocupado, se observan propiedades dinámicas intermedias de la tapa ya que no tiene contactos con ningún ligante en el sitio activo, lo cual resulta en tres conformaciones alternativas de la tapa, sólo una de las cuales es estabilizada por Glu148 (figura 7).



Figura 7. Las tres conformaciones alternativas (cerrada, entreabierta y abierta) de la tapa del sitio activo en la forma R de la enzima, con sólo el sitio alostérico ocupado. El código de colores es el de la figura 5. Adaptación de [4].

#### 3.2. Forma T con el sitio activo ocupado no observada

El modelo MWC predice que, en solución y en ausencia de ligantes, sólo una pequeña población de la enzima afin al sustrato coexiste en equilibrio con los conformeros menos afines de GlcN6P desaminasa. Así que, en presencia de moléculas que se unan al sitio activo (GlcN6P o Fru6P),

la forma T libre de ligantes cambiaría a una forma T con el sitio activo ocupado la cual, a su vez, se transformaría en una forma R con el sitio activo ocupado a través de la transición alostérica homotrópica.

Sin embargo, los diversos intentos por obtener un cristal de la forma T con el sitio activo ocupado se han reportado infructuosos [4]. Aunque esto podría deberse a que se trate en realidad de un intermediario inestable que no pueda ser detectado por cristalografía de rayos X, existe también la posibilidad de que en la transición homotrópica el sustrato se una sólo a la forma R libre de ligantes, y no a la forma T con la que coexiste. Esto último es consistente con las observaciones de la transición heterotrópica, en la cual el activador alostérico se une exclusivamente al estado R [3], y también con las observaciones de la flexibilidad atómica, como se presenta a continuación.

### **3.3. Flexibilidad de la forma T libre de ligantes**

Una manera indirecta de observar la flexibilidad de las moléculas de un cristal es a través de los factores  $B$ , los cuales están directamente relacionados con la vibración térmica de los átomos. Generalmente pueden calcularse los factores  $B$  isotrópicos, los cuales dan información sólo de la magnitud del movimiento, pero en algunos casos pueden calcularse los factores  $B$  anisotrópicos, los cuales además describen las direcciones preferenciales de las vibraciones.

Este análisis se aplicó a un cristal de GlcN6P desaminasa en su estado T [4], cuya estructura corresponde a la de código 1FSF almacenada en el banco de datos de proteínas [11] (PDB, por sus siglas más comunes). Se encontraron factores  $B$  anisotrópicos con valores bajos (poca movilidad) en el módulo interno, y con valores altos (mayor movilidad) en el módulo externo, consistentes con un movimiento de rotación paralelo al eje de orden 3 del hexámero, a pesar de estar la proteína en una malla cristalina (figura 8). De manera interesante, el centro de esta rotación coincide con el que se conoce para la transición alostérica [5]. Este movimiento, observado para un protómero en el cristal, se propaga a todo el hexámero a través de los residuos Tyr254 y Thr152 del sitio alostérico, que están involucrados en interacciones interprotoméricas específicas.

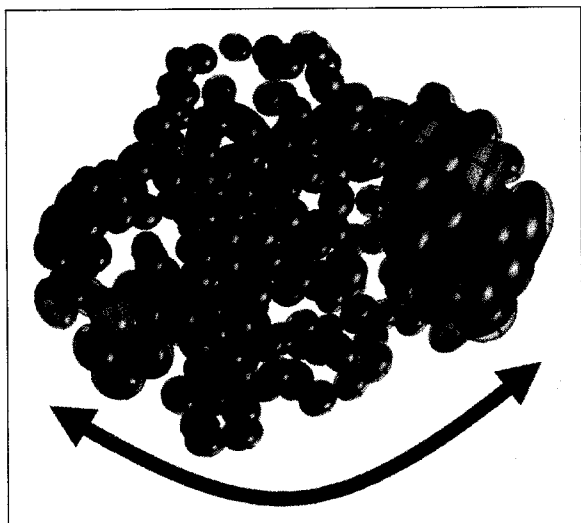


Figura 8. Representación en elipsoides de los factores  $B$  anisotrópicos de los átomos  $C^\alpha$  de un protómero de la GlcN6P desaminasa en su conformación T libre de ligantes (cfr. figura 6). La longitud de cada eje de los elipsoides indica el grado de movilidad en esa dirección para el átomo correspondiente. El movimiento descrito en este análisis se asemeja a la rotación del cambio alostérico en la dirección del desplazamiento colectivo de los  $C^\alpha$  (representada por la flecha azul), el centro de la rotación (indicado por la elipse azul) y el grado de movilidad de las distintas zonas (indicado por el tamaño de los elipsoides y por la longitud de onda de su color: de azul a rojo). Adaptación de [4].

Estos resultados también están apoyados por las conformaciones alternativas encontradas en el mismo cristal (1FSF) para la tapa del sitio activo y el extremo carboxilo terminal (figura 9). En estas zonas los conformémeros vecinos están muy cercanos entre sí (menos de 1 Å) y, sin embargo, presentan factores  $B$  isotrópicos relativamente altos (con valores mayores a 55 Å<sup>2</sup>), lo cual sugiere que en realidad no se trata de un conjunto de conformémeros estáticos, sino de un movimiento continuo de gran amplitud en esas zonas, tal como sugiere el análisis de factores  $B$  anisotrópicos.

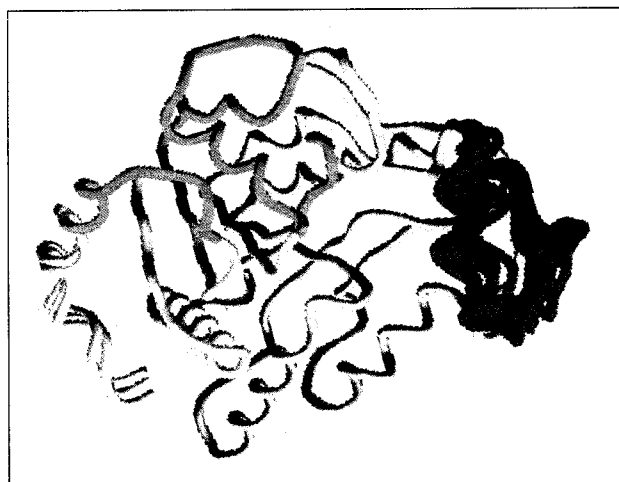


Figura 9. Traza de un protómero de la GlcN6P desaminasa en su conformación T libre de ligantes (cfr. figura 5), en la que se muestran las conformaciones alternativas encontradas (siete para la tapa del sitio activo y tres para el extremo carboxilo terminal). El código de colores es el de la figura 5. Adaptación de [4].

Así, estas evidencias sugieren que el estado T de la GlcN6P desaminasa de *E. coli* presenta la transición alostérica aun sin ligantes unidos a sus sitios activos o alostéricos, y su estructura se alterna entre los conformeros T y R. Dado que los valores de los factores *B* isotrópicos encontrados en los cristales de los conformeros R (con al menos un ligante unido) son menores que los de los conformeros T, el estado activo parece ser menos flexible que el que está libre de ligantes, de forma que el papel de los efectores es fijar a la proteína en su conformación activa disminuyendo su entropía.

### **3.4. Modelo alostérico de la GlcN6P desaminasa de *E. coli***

Estos resultados ponen en evidencia algunas características de la enzima que restringen su comportamiento alostérico a un caso particular del modelo alostérico MWC [9], como se expone a continuación.

- El modelo alostérico concertado predice la existencia de una forma T de la proteína unida a algún ligante alostérico, la cual participa en una posible ruta de activación haciendo de intermediaria entre la forma T libre de ligantes y la forma R con un sitio ocupado (figura 4). Sin embargo, las dificultades para obtener un cristal de la forma T con el sitio activo ocupado sugieren que la GlcN6P desaminasa sólo une ligantes en su estado R, de manera que sólo puede seguir una de las rutas de activación propuestas: de la forma T libre de ligantes a la forma R libre de ligantes, y de esta a la forma R con un ligante unido.
- Según el modelo alostérico concertado, una enzima alostérica que ya ha unido uno o más ligantes puede interconvertirse entre sus conformaciones T y R (figura 4). Como se mencionó, en la GlcN6P desaminasa no se ha conseguido observar la forma T unida a ligantes, de forma que de la figura 4 sólo queda T sin ligantes y la ruta inferior.
- El conformero menos afin se considera tenso (T) y el conformero activo se considera relajado (R) en el modelo simétrico, mientras que en el caso de esta proteína existe no sólo una conformación sino un estado T, pues se trata en realidad de un estado cinético de gran flexibilidad que oscila entre una gama continua de conformaciones, una de las cuales corresponde a la fugaz conformación R libre de ligantes (y que tiene una población pequeña en comparación con el resto de las conformaciones como conjunto, y por lo tanto no es observable por cristalografía). Esta conformación es la que puede unir ligantes en cualquiera de los sitios de unión, de forma que el papel de los efectores es fijar a la enzima en su conformación activa, la cual es menos móvil en comparación con el estado T.

Para reconciliar estas diferencias con la teoría alostérica, Rudiño y col. [4] propusieron un modelo concertado restringido para la GlcN6P desaminasa de *E. coli*, el cual se muestra en la figura 10.

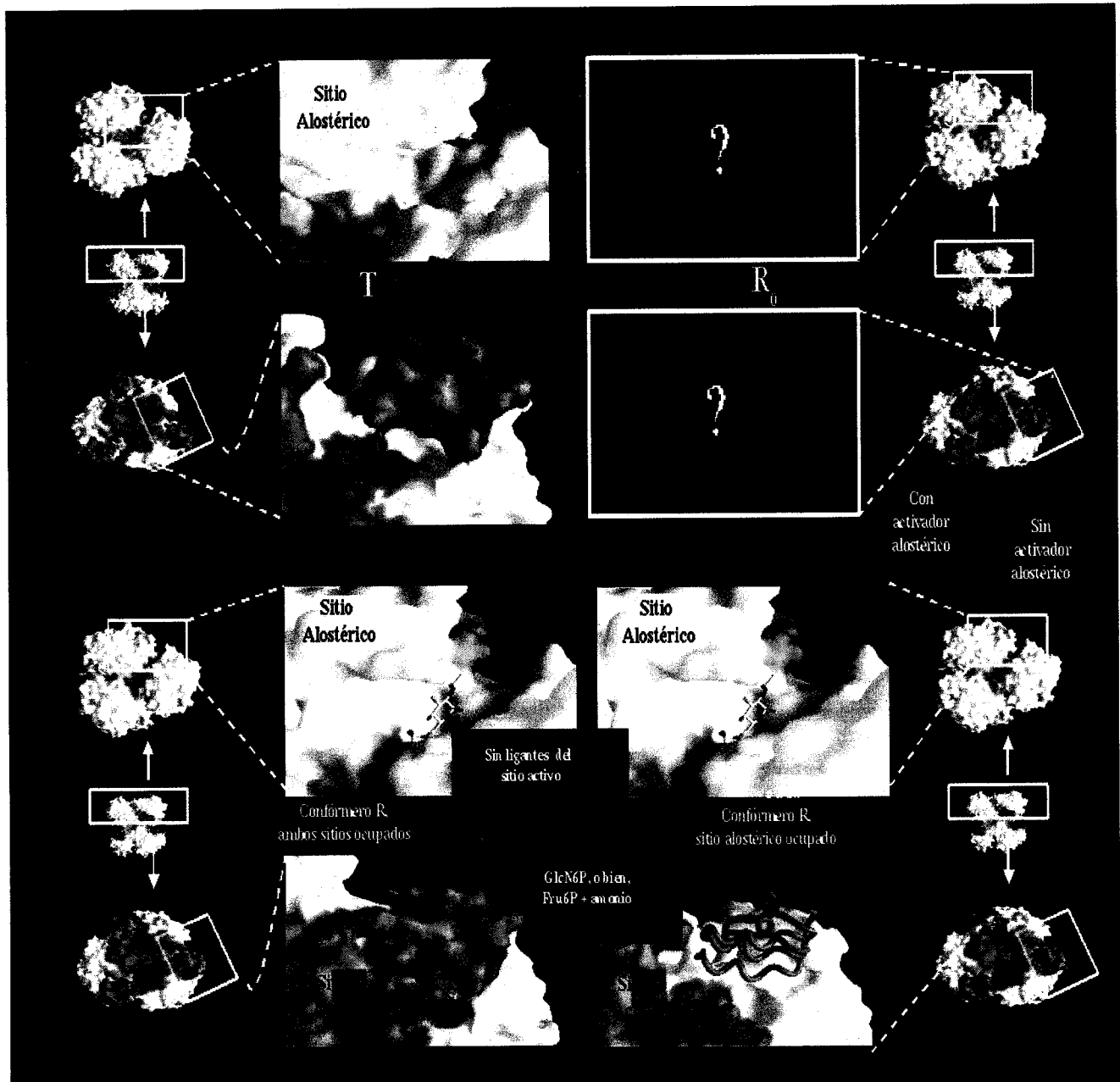


Figura 10. Modelo alostérico de la GlcN6P desaminasa. Se presenta la secuencia obligatoria para la transición alostérica heterotrópica: A, forma T libre de ligantes; B, forma R libre de ligantes (las conformaciones A y B en distintas concentraciones son parte del estado cinético T); C, la forma R con sólo el sitio alostérico ocupado; D, la forma R con ambos sitios de unión ocupados (las conformaciones C y D forman el estado R). La secuencia de la transición alostérica homotrópica es A→B→D. El código de colores es el de la figura 5. Adaptación de [4].



En este modelo, tanto en la transformación homotrópica como en la heterotrópica el proceso de unión a un ligante se da directamente en la forma R libre de ligantes, la cual está en equilibrio con la forma T libre de ligantes por medio de vibraciones atómicas en la dirección de la transformación alostérica.

Para clarificar las correcciones del modelo alostérico de la GlcN6P desaminasa con respecto al modelo MWC general, se presenta el primero (figura 11) en un diagrama similar al del segundo (figura 4).

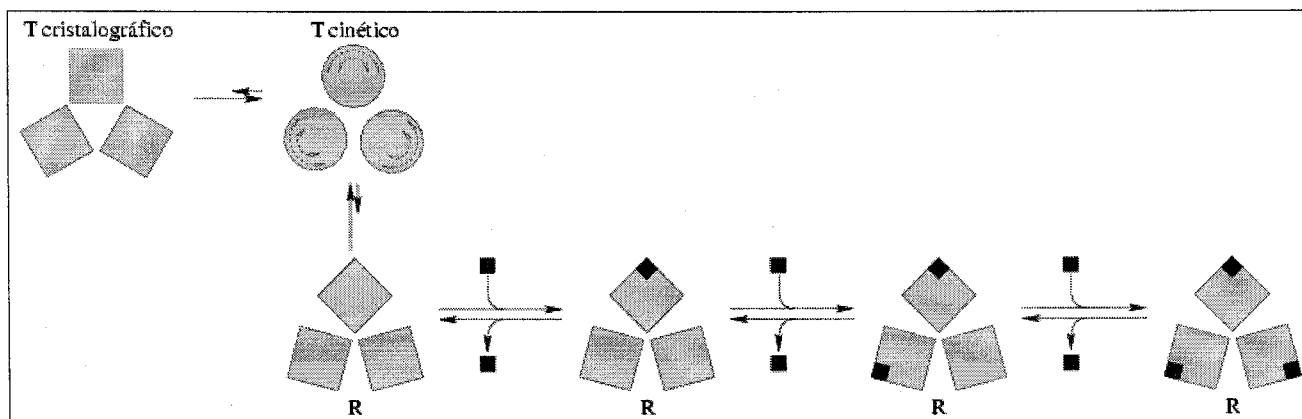


Figura 11. Modelo alostérico de la GlcN6P desaminasa [4], esquematizado para uno de sus trímeros (cfr. figura 4). Los cuadros grises representan protómeros, mientras que los cuadros negros corresponden a ligandos activadores. En ausencia de ligandos, la proteína en solución se encuentra en un estado cinético, sin una conformación única, pero que en uno de los extremos de su oscilación estructural visita la conformación R (la única capaz de unir ligandos).

En este modelo, la proteína en solución a baja concentración de ligandos se encuentra en el estado T cinético (figura 11), en el cual rota de acuerdo a su transición alostérica (alrededor de los módulos internos), oscilando de manera concertada a lo largo de toda una gama de conformaciones que se encuentran en equilibrio.

Dado que la cristalografía de proteínas detecta posiciones promedio de los átomos de un cristal, las estructuras cristalográficas resueltas para el estado T [5] [4] (estado T cristalográfico en la figura 11) son en realidad la estructura a medio camino de la oscilación que lleva los protómeros de una estructura fugaz extrema a otra, una de las cuales no puede ser observada por estos métodos debido a su relativamente baja densidad de población, mientras que la otra es de hecho la forma R sin ligandos (abajo izquierda, en la figura 11), la cual sí puede ser detectada por cristalografía cuando es fijada gracias a la unión de algún ligante. En realidad, es de esperarse que en la matriz cristalina esta oscilación se vea limitada, y se asemeje en dirección y estructura promedio a la que ocurre en solución, pero se diferencie de ella al tener una menor amplitud.

### **3.5. GlcN6P desaminasa en el contexto de otros sistemas alostéricos**

Cada vez se reportan más casos de procesos alostéricos en los que (como en el caso de la GlcN6P desaminasa) las transiciones dinámicas y estructurales pueden describirse como un desplazamiento del equilibrio de poblaciones preexistentes, las cuales se intercambian de manera dinámica aun en ausencia del ligante alostérico [12].

Por ejemplo, se han observado en ausencia de ligantes movimientos relacionados con el cambio conformacional alostérico para diversos sistemas biológicos y mediante diferentes métodos: mioglobina (dinámica molecular) proporcionó uno de los primeros indicios [13], el sistema de chaperonas GroEL-GroES (cálculo de modos normales [14] y dinámica molecular dirigida [15]), aspartato transcarbamilasa (cálculo de modos normales) [16], fructosa-1,6-bisfosfatasa (cristalografía de rayos X, fluorescencia) [17], proteína de señalización NtrC (espectroscopía por resonancia magnética nuclear) [18], glutamato deshidrogenasa [19], proteína de replicación A humana [20] y fosfolipasa A2 de procarionte [21] (cristalografía de rayos X), hemoglobina (dinámica molecular dirigida [22] y cristalografía de rayos X y resonancia magnética nuclear [23]), e incluso la molécula no proteica ARN TAR de VIH (acoplamiento dipolar residual) [24].

También apoyando este modelo, se han reportado casos en los que la unión de un ligante alostérico provoca una disminución en la entropía del sistema: calbindina D<sub>9k</sub> [25], calmodulina [26] y proteína de atenuación de unión a ARN *trp* [27] (resonancia magnética nuclear), así como el antibiótico glucopeptídico vancomicina (dinámica molecular) [28].

## 4. HIPÓTESIS

El cambio conformacional de la glucosamina-6-fosfato desaminasa se da incluso antes de la unión a un ligante en el estado T cinético, siendo la función del efector fijar a la molécula en su conformación activa (conformación R).

Esta transición puede modelarse como un movimiento oscilatorio, el cual puede ser observado en una simulación de Dinámica Molecular.

## 5. OBJETIVOS

### 5.1. *Objetivos generales*

Corroborar la propuesta del estado T cinético de Rudiño y col. [4] observando su movimiento mediante Dinámica Molecular, y encontrar la frecuencia de oscilación de la transición alostérica.

### 5.2. *Objetivos particulares*

- Generar simulaciones de la GlcN6P desaminasa en su forma hexamérica tanto en la red cristalina como en solución.
- Analizar los archivos de trayectoria generados identificando el punto de la oscilación en el que se encuentre la proteína a cada instante de tiempo.
- A partir de las oscilaciones observadas, calcular una frecuencia para la transición alostérica del estado T cinético.

## **6. MARCO TEÓRICO**

### **6.1. Simulaciones**

En situaciones en las que, para analizar un sistema de interés, es necesario manipular o detectar variables que no son observables experimentalmente (por ejemplo, las posiciones de los átomos de una molécula) y no es posible derivar una teoría analítica, puede recurrirse al uso de simulaciones. Estas, al emular la respuesta del sistema real (en base a un modelo teórico) ante ciertos observables (que fuera de este esquema no necesariamente podrían ser manipulados o controlados), generan información relevante al hacer accesibles otros observables.

En el caso de las simulaciones computacionales, la evolución del estado del sistema (trayectoria) es calculada por medio de programas que siguen un algoritmo determinado, el cual depende del método utilizado.

Para estudiar fenómenos secuenciales, es decir, con correlación temporal, es necesario que el método de simulación sea determinista. Esto significa que los estados del sistema generados a lo largo de la trayectoria deben estar conectados en el tiempo, es decir, cada uno se calcula a partir del anterior. El método que más rigurosamente cumple con este requisito es el de Dinámica Molecular, pues calcula la evolución del sistema resolviendo las ecuaciones de movimiento de cada uno de sus componentes a cada instante de tiempo.

### **6.2. Dinámica Molecular**

La Dinámica Molecular es un método computacional que calcula a partir de unas condiciones iniciales dadas para un sistema, su evolución a lo largo del tiempo, en base a un modelo de las interacciones entre sus componentes. Una simulación de este tipo consta de cuatro etapas básicas: inicialización (definir posiciones y velocidades iniciales para todos los átomos del sistema de interés), calentamiento (incrementar la energía cinética hasta alcanzar la temperatura a la que se llevará a cabo la simulación), equilibrio (estabilización de la energía) y producción (parte útil de la trayectoria sobre la que se realiza el análisis), las cuales se describen en las siguientes secciones.

#### **6.2.1. Inicialización**

El comportamiento de una biomolécula generalmente es analizado no en vacío, sino en condiciones fisiológicas, así que debe ser solvatada. Para ello se utiliza algún programa que coloque moléculas de agua alrededor del soluto, respetando el espacio ocupado por este.

Las condiciones iniciales para los cálculos definen las velocidades y las posiciones de los elementos del sistema al inicio de la simulación. Asignar posiciones a todos los átomos puede no ser trivial, como se detalla a continuación, pero una vez que estas han sido definidas las velocidades usualmente se asignan mediante un algoritmo pseudoaleatorio de acuerdo a una distribución gaussiana, a partir de un número generador (semilla) dado. Cuando el sistema de interés es una proteína, las posiciones iniciales de sus átomos por lo general se preparan a partir de las coordenadas de un archivo de coordenadas atómicas del PDB [11].

Dado que esta estructura usualmente es producto de técnicas cristalográficas, en la mayoría de los casos no tiene especificadas las posiciones de sus átomos de hidrógeno, así que estas deben construirse de manera empírica a partir de las de los átomos pesados y las relaciones fisicoquímicas entre ellos, utilizando un algoritmo computacional.

Por otra parte, los átomos que sí hayan sido detectados por cristalografía se encuentran en posiciones promediadas, lo cual puede permitir que algunos átomos estén demasiado cercanos entre sí elevando en gran medida la energía potencial. Estas repulsiones ocasionan que el sistema se caliente rápidamente alcanzando temperaturas no fisiológicas, así que debe recurrirse a un algoritmo de minimización de energía para remediar esto.

#### **6.2.1.1. Minimización**

Este paso consiste en modificar las coordenadas de los átomos del sistema y observando los efectos sobre la energía potencial (los cuales se calculan a partir del potencial de interacción, como se describe más adelante). Se trata de un proceso iterativo que estabiliza la conformación al disminuir su energía potencial. Existen distintos algoritmos de minimización, como el de Descensos Más Pronunciados o el de Gradiente Conjugado, que están entre los más comúnmente utilizados.

La minimización por Descensos Más Pronunciados utiliza el algoritmo más simple de su tipo pues, como su nombre lo sugiere, busca disminuir la energía potencial de la manera más rápida posible. Debido a esta característica, en general no converge a un valor determinado de energía potencial, pero es ideal para la inicialización de estructuras con grandes repulsiones estéricas, o como primera aproximación para obtener una conformación estable.

La técnica de Gradiente Conjugado [29] funciona con un algoritmo similar al de Descensos Más Pronunciados, en el sentido de que ajusta iterativamente las coordenadas en la dirección opuesta al gradiente de la energía potencial. Sin embargo, este método tiene mejores características de convergencia, ya que tiene memoria de los pasos de minimización que ha realizado anteriormente, y combina esa información con la del gradiente actual para calcular el siguiente paso.

La principal desventaja de la minimización por Gradiente Conjugado es que tiene una mayor probabilidad que Descensos Más Pronunciados para generar desbordamiento numérico (generación de números que no pueden ser manejados por la computadora) al tratar con estructuras demasiado inestables, así que la estrategia general para minimizar la energía de una estructura es utilizar primero Descensos Más Pronunciados para aliviar las repulsiones más graves, y corregir los detalles restantes por Gradiente Conjugado.

### **6.2.2. Calentamiento**

El estado inicial del sistema equivale a su estado a temperatura de cero absoluto, pues todos sus átomos están estáticos (a pesar de que en un contexto computacional tienen velocidades asignadas a priori en la etapa de inicialización, las partículas no han comenzado a moverse). Con el fin de llevarlo a una temperatura de interés a la que se desee observar el comportamiento del sistema (temperatura fisiológica, por ejemplo) el proceso de calentamiento, al mismo tiempo que actualiza las posiciones de los átomos con un algoritmo integrador, eleva gradualmente la energía cinética hasta el valor deseado, utilizando incrementos pequeños de la velocidad de las partículas para no desestabilizar el sistema (comenzando por los valores de velocidades asignados en la etapa de inicialización).

En realidad, esta es una etapa opcional de la simulación, ya que algunos sistemas pueden pasar directamente del cero absoluto (etapa de inicialización) a la temperatura final (etapa de equilibrio) sin volverse inestables.

### **6.2.3. Equilibrio**

Una vez alcanzada la temperatura deseada, el sistema puede aún estar fuera de equilibrio y, aunque no sea inestable, este puede presentar un comportamiento distinto al real. Por lo anterior, la información generada en esta etapa no es válida para el análisis.

Para alcanzar un estado en el que puedan obtenerse datos válidos, debe permitirse que el sistema se equilibre por sí mismo evolucionando a lo largo de la simulación. Este proceso es generalmente corto en comparación con la parte útil de la trayectoria (obtenida en la etapa de producción), y su duración depende del parámetro que se adopte para definir el estado de equilibrio del sistema (generalmente la energía total). Esta magnitud será monitorizada a lo largo del tiempo, y se considerará que el sistema ha alcanzado el equilibrio cuando esta se haya estabilizado, es decir, cuando sus fluctuaciones sean relativamente pequeñas.

#### **6.2.4. Producción**

Esta es la etapa en la que se permite evolucionar al sistema equilibrado, generando así la trayectoria que será estudiada. La información cruda que se obtiene son las coordenadas del sistema cada cierto número de pasos de integración, así como otros parámetros como su energía o los parámetros de red de la celda unitaria. Dado que los archivos de salida son demasiado extensos, su análisis debe hacerse con programas de cómputo.

#### **6.3. Condiciones de frontera periódicas**

Las fronteras de un sistema a simular requieren un tratamiento especial en los cálculos, ya que las moléculas tienden a disiparse en el vacío si están en contacto con él. Una forma de evitar esto es imponiendo una fuerza que mantenga las moléculas dentro de una región determinada. Sin embargo, el introducir esta fuerza externa modifica la energía del sistema y produce un modelo menos realista.

La alternativa más común es utilizar condiciones de frontera periódicas, con las cuales el sistema se encuentra rodeado por réplicas de sí mismo llamadas "imágenes" e interacciona con ellas y ya no con el vacío. De esta forma, en caso de que alguna molécula salga del sistema original por una de sus fronteras, una réplica de esta misma entrará de manera concomitante a través de otra zona de la frontera. Visto de otra forma, una molécula que salga del sistema reingresará inmediatamente en otra región.

#### **6.4. Potencial de interacción**

El comportamiento de los átomos de un sistema se modela con un potencial que gobierna las interacciones entre ellos. Este contiene información acerca del costo energético que implica deformar un enlace químico entre dos átomos, un ángulo entre tres átomos enlazados, o un ángulo diedro (ángulo de torsión entre dos enlaces separados por un tercer enlace) o impropio (posición relativa entre dos átomos unidos indirectamente mediante tres enlaces) que involucren cuatro átomos enlazados.

Estas interacciones son referidas como interacciones de enlace, pero no son las únicas que participan entre un conjunto de átomos, y también deben incluirse en el modelo las interacciones de Van der Waals y electrostáticas (fuerzas de Coulomb). Estas ocurren entre pares de átomos, y pueden ser de corto o largo alcance. En este último caso, ambas pierden intensidad conforme aumenta la distancia de separación, pero nunca se anulan.

Lo anterior representa una gran limitación para una simulación, ya que el cálculo de las interacciones de largo alcance de cada átomo con el resto implica un costo computacional demasiado



alto. Por lo tanto, debe modelarse una manera distinta de incluir en los cálculos las fuerzas de Van der Waals y de Coulomb.

En el caso de Van der Waals, los autores del potencial utilizado en este trabajo sugieren un truncamiento de su potencial de acción, es decir, para un par de átomos dado, su interacción de Van der Waals será tomada en cuenta sólo si la distancia que los separa es menor que una distancia de corte previamente definida. Para distancias mayores, en cambio, este término del potencial de interacción será nulo. Esto implica una discontinuidad en la energía así que, en realidad, se modifica el potencial a partir de una distancia poco antes de la de corte, de forma que comience a anularse y tome el valor de cero justo en la distancia de corte, y en adelante.

Por otra parte, para el problema de las interacciones electrostáticas, el método que cuenta con una mayor aceptación es el de las sumas de Ewald [30] [31]. Sin embargo, este método presenta dos limitaciones en el contexto de este trabajo. La primera es que puede generar artefactos en la dinámica de biomoléculas si no existe una gran distancia entre estas y sus imágenes [32], lo cual no es posible para las subunidades de los hexámeros de la proteína estudiada en este trabajo. La segunda es que no es posible eliminar las interacciones entre los oligómeros imágenes aumentando la distancia entre ellos para emular un ambiente fuera de la red cristalina en las simulaciones del hexámero en solución.

Estas dos limitaciones desaparecen si en lugar de las sumas de Ewald se utiliza un método de corte, similar al descrito para las fuerzas de Van der Waals. Este, además de generar un ahorro equivalente en el costo computacional del cálculo de las fuerzas Coulómbicas, genera simulaciones igual de estables si la distancia de corte es suficientemente grande [33]. Aquí no se modifica el potencial (en este caso electrostático) cercano a la distancia de corte, sino que la energía es atenuada uniformemente de forma que se anule a una separación igual a este valor, y continúe siendo nula a distancias mayores. Esta distancia de corte suele ser la misma que se define para el truncamiento de las interacciones de Van der Waals.

Para los cálculos de las interacciones de largo alcance con estos métodos es necesario generar para cada átomo una lista de vecinos, es decir, una lista de los átomos que se encuentren a una distancia menor o igual a la de corte. Debido al movimiento de las partículas a lo largo de una simulación, esta lista deberá ser actualizada constantemente. Dado que este proceso también consume tiempo de cómputo, debe ser agilizado restringiendo los candidatos a sólo aquellos que estén dentro de una distancia umbral definida (mayor a la distancia de corte).

## 6.5. Ensemble termodinámico

Las simulaciones de sistemas biológicos por lo general se realizan en el ensemble isotérmico-isobárico, también referido como ensemble  $NPT$ , en el cual permanecen constantes el número de partículas, la presión y la temperatura. Esto es con el propósito de emular las condiciones del ambiente en que suelen encontrarse las biomoléculas.

Conservar constante la cantidad de átomos de un sistema resulta trivial. Sin embargo, ha sido necesario desarrollar métodos que mantengan constantes funciones de estado como la presión y la temperatura.

### 6.5.1. Presión constante

En el programa utilizado en este trabajo existen varios métodos propuestos para mantener constante la presión de un sistema con condiciones periódicas de frontera, pero todos pueden visualizarse con un esquema en común. En este esquema las ecuaciones de movimiento de las partículas se redefinen en términos del volumen del sistema (el cual se introduce como una nueva variable en los cálculos), de forma que el sistema se comporta como si estuviese acoplado a un pistón. Así, el movimiento de este pistón es el que modifica el volumen del sistema según sea necesario para corregir las fluctuaciones de la presión con respecto al valor de referencia. Este movimiento compensatorio debe estar amortiguado para evitar que las propiedades dinámicas del sistema dependan de la masa del pistón.

De los métodos disponibles, el de acoplamiento débil [34] equivale a tener un pistón sobreamortiguado mientras que, en los métodos de sistema extendido [35] [36], este movimiento está completamente desamortiguado. En cambio, el método del pistón de Langevin [37] se coloca en medio de los dos extremos mencionados, ya que permite un amortiguamiento parcial al utilizar las siguientes ecuaciones de movimiento:

$$\dot{\vec{r}}_i = \frac{\dot{\vec{p}}_i}{m_i} + \frac{1}{3} \frac{\dot{V}}{V} \vec{r}_i \quad \text{Ecuación 6.1}$$

$$\dot{\vec{p}}_i = \vec{f}_i - \frac{1}{3} \frac{\dot{V}}{V} \vec{p}_i \quad \text{Ecuación 6.2}$$

$$\ddot{V} = \frac{1}{W} (P[t] - P_{ext}) - \gamma \dot{V} + R[t] \quad \text{Ecuación 6.3}$$

donde  $V$  es el volumen del sistema,  $P[t]$  es la presión instantánea,  $P_{ext}$  es la presión impuesta,  $W$  es la masa del pistón, y  $\vec{r}_i$ ,  $\vec{p}_i$ ,  $m_i$  y  $\vec{f}_i$  son, respectivamente, la posición, el momento, la masa y la fuerza para la  $i$ -ésima partícula. La ecuación 6.3 es una ecuación de Langevin, ya que involucra un término de ruido ( $R[t]$ , que equivale a una fuerza aleatoria con una distribución gaussiana) y un término de viscosidad ( $-\gamma\dot{V}$ , donde  $\gamma$  es la frecuencia de colisión).

El valor que se le deba asignar a  $W$  depende del tamaño del sistema, pues en tanto más átomos influyan sobre la presión, mayores serán las fluctuaciones que deban ser corregidas y, por lo tanto, mayor inercia deberá tener el pistón de Langevin. Por ejemplo, para un sistema pequeño puede usarse una masa de hasta 500 Da. Sin embargo, distintos valores dentro de este orden no presentan efectos apreciables sobre la dinámica, de manera que para sistemas de 3000 a 30000 átomos el valor de  $W$  puede variar de 200 Da a 1000 Da [37].

La función de  $\gamma$  es amortiguar las oscilaciones de la presión instantánea para un valor de  $W$  dado. Un pistón de Langevin masivo requiere valores pequeños de  $\gamma$ , y esta generalmente se encuentra en el orden de las decenas de  $\text{ps}^{-1}$  [37].

Debido a que el movimiento del pistón es disipativo, la energía disipada por él debe ser retribuida al sistema para evitar que la temperatura disminuya a lo largo de la simulación. Por esta razón el sistema se acopla a un baño térmico a través de la fuerza aleatoria de la ecuación 6.3. Sin embargo, dado que este acoplamiento sólo involucra las variables relacionadas con el volumen, el intercambio de energía es lento, así que debe implementarse algún método adicional para acoplar las partículas a un baño térmico y asegurar que la simulación se realice en el ensamble *NPT* [37].

### **6.5.2. Temperatura constante**

Uno de los métodos más aceptados para mantener la temperatura constante, y que es el sugerido por los autores del algoritmo del pistón de Langevin, es el del termostato de Nosé-Hoover [38] [39] (o termostato integral). Este aborda el problema de la temperatura de manera similar a los métodos de presión constante, pues acopla matemáticamente las variables del sistema a un "pistón térmico" que regula el intercambio de energía cinética entre el sistema y un reservorio infinito de calor para mantener una temperatura de referencia en el sistema.

También de manera análoga, este método involucra un parámetro que influye en la respuesta del pistón a las fluctuaciones del sistema, llamado parámetro de inercia térmica. Así, si la inercia del pistón es demasiado pequeña el sistema tendrá total libertad de disipar o absorber calor a través del pistón y la temperatura no se mantendrá constante, mientras que si, por el contrario, el pistón es

demasiado masivo, el control de temperatura no será eficiente. Su valor debe estar entre 10 kcal ps<sup>2</sup> y 1000 kcal ps<sup>2</sup>, y debe aumentar junto con el tamaño del sistema.

## 6.6. Algoritmo de integración

Una vez que se tiene definida la manera en la que se comportarán las partículas del sistema a lo largo de la simulación, es necesario definir el algoritmo con el que se propagará el movimiento al actualizar las posiciones y velocidades de los átomos, generando así la trayectoria.

Lo anterior implica integrar las ecuaciones de movimiento con un método de diferencias finitas, el cual resolverá el sistema de ecuaciones paso a paso en intervalos discretos de tiempo de longitud  $\delta t$  (llamada "tamaño de paso").

El método de diferencias finitas que es más ampliamente utilizado y, de hecho, el único que puede implementarse en el ensamble isotérmico-isobárico con el programa usado en este trabajo, es el algoritmo Verlet Pídola. Este tiene un error del orden de  $\delta t^2$  gracias a que las posiciones y las velocidades no están definidas en un mismo instante, sino que están desfasadas  $\frac{1}{2}\delta t$ . Las ecuaciones que utiliza para avanzar la velocidad  $\dot{\vec{r}}$  de un tiempo  $t-\frac{1}{2}\delta t$  a  $t+\frac{1}{2}\delta t$ , y la posición  $\vec{r}$  de  $t$  a  $t+\frac{1}{2}\delta t$  son:

$$\dot{\vec{r}}\left[t+\frac{1}{2}\delta t\right]=\dot{\vec{r}}\left[t-\frac{1}{2}\delta t\right]+\frac{f\left[\vec{r}\left[t\right]\right]}{m}\delta t$$

Ecuación 6.4

$$\vec{r}\left[t+\delta t\right]=\vec{r}\left[t\right]+\dot{\vec{r}}\left[t+\frac{1}{2}\delta t\right]\delta t$$

Ecuación 6.5

donde  $m$  es la masa de la partícula y  $f$  es la fuerza que actúa sobre ella de acuerdo con el potencial de interacción. De esta manera, las posiciones y velocidades de todas las partículas calculadas para un tiempo dado son utilizadas para calcular las posiciones y velocidades al siguiente paso de tiempo. La trayectoria se genera repitiendo este procedimiento para cada paso de la simulación.

### 6.6.1. Tamaño de paso

La elección del tamaño de paso es de importancia crítica para todos los métodos de diferencias finitas. Su valor debe ser unas 10 veces menor que el período de la vibración de mayor frecuencia del sistema. Sin embargo, la simulación avanzará más rápido en tanto mayor sea el tamaño de paso, de manera que para hacer los cálculos más eficientes el valor de  $\delta t$  se escoge tan grande como sea posible.

### **6.6.1.1. Algoritmo SHAKE**

En el caso de los sistemas biológicos las vibraciones de mayor frecuencia corresponden a los enlaces entre los átomos pesados y los hidrógenos. Estos movimientos tienen un período de aproximadamente 11 fs, lo cual implica que el tamaño de paso más grande que puede usarse en una simulación es de 1 fs. Sin embargo, estas vibraciones no son de gran importancia para la dinámica de una macromolécula, así que es justificable eliminarlas manteniendo rígidos todos los enlaces de este tipo. Lo anterior es ventajoso ya que, dado que la dinámica ahora presentaría movimientos más lentos, permite aumentar el valor de  $\delta t$  al doble, es decir, 2 fs, con lo cual los cálculos de la simulación son más eficientes.

El algoritmo SHAKE [40] hace precisamente esto, pues aplica un método iterativo que restringe el estiramiento o la contracción de los enlaces con los átomos de hidrógeno con respecto a su estado basal, dentro de un porcentaje de tolerancia.

## 7. METODOLOGÍA

### 7.1. Generación de las simulaciones

A continuación se dan las especificaciones técnicas de los métodos utilizados para llevar a cabo las simulaciones por Dinámica Molecular. En este trabajo se produjeron cuatro simulaciones de distintas longitudes que modelan el hexámero en el estado T cinético: dos en el cristal original, llamadas Crist14 (de 14 ns) y Crist06 (de 6 ns), y dos en solución, llamadas Sol28 (de 28 ns) y Sol10 (de 8 ns).

#### 7.1.1. Construcción del sistema

Para todas las simulaciones se tomó como modelo inicial la estructura cristalográfica de la GlcN6P desaminasa de *Escherichia coli* en su conformación T a 1.9 Å de resolución, con código de identificación 1FSF [4] en el PDB [11].

Para las simulaciones Sol28 y Sol10 la unidad biológica se emuló como un hexámero en solución expandiendo los parámetros de red del cristal original ( $a=b=126.676$  Å,  $c=139.326$  Å) hasta tomar valores que aseguraran una distancia mínima entre los oligómeros suficientemente grande (más de 14 Å) para evitar la interacción entre ellos ( $a=b=151.9$  Å,  $c=167.0$  Å).

Dado que durante la fase de equilibrio el sistema se contrae disminuyendo sus parámetros de red, fue necesario dilatar la celda anticipando que el cristal retomaría el tamaño deseado. Así, para las simulaciones del cristal original Crist14 y Crist06, la celda se expandió 3.88% para que retomara los parámetros de red  $a=b=126.676$  Å y  $c=139.326$  Å (figura 12). Para las simulaciones Sol28 y Sol10 el porcentaje de expansión inicial fue de 3.4%.

Los estados de protonación de las histidinas (con hidrógeno enlazado ya sea en el N<sup>δ</sup>, N<sup>ε</sup>, o en ambos, es decir, con carga positiva) se definieron con base en un análisis visual de los posibles puentes de hidrógeno que formarían de acuerdo a las coordenadas cristalográficas de los átomos pesados (archivo 1FSF). Para construir el modelo para la simulación Sol10, además, se incluyeron en el análisis otros archivos de coordenadas de la GlcN6P desaminasa de *E. coli* (1CD5, 1DEA, 1FQO, 1FS6). En el caso de las histidinas para las que no pudo determinarse su estado de protonación por este método, se definieron arbitrariamente con una carga neutra y con el átomo de hidrógeno enlazado al N<sup>δ</sup>. Las posiciones de los átomos de hidrógeno en cuestión se presentan en la tabla I.

Tabla I. Estados de protonación definidos para las histidinas de cada modelo construido.

Residuo	Átomo enlazado al hidrógeno (Crist14, Crist06 y Sol28)	Átomo enlazado al hidrógeno (Sol10)
His19	N <sup>δ</sup>	N <sup>ε</sup>
His56	N <sup>δ</sup>	N <sup>δ</sup>
His65	N <sup>δ</sup>	N <sup>δ</sup>
His81	N <sup>δ</sup>	N <sup>ε</sup>
His90	N <sup>δ</sup>	N <sup>δ</sup>
His96	N <sup>δ</sup>	N <sup>ε</sup>
His132	N <sup>δ</sup>	N <sup>ε</sup>
His143	N <sup>δ</sup>	N <sup>δ</sup>
His164	N <sup>δ</sup>	N <sup>δ</sup>
His222	N <sup>δ</sup>	N <sup>δ</sup>
His232	N <sup>ε</sup>	N <sup>ε</sup>

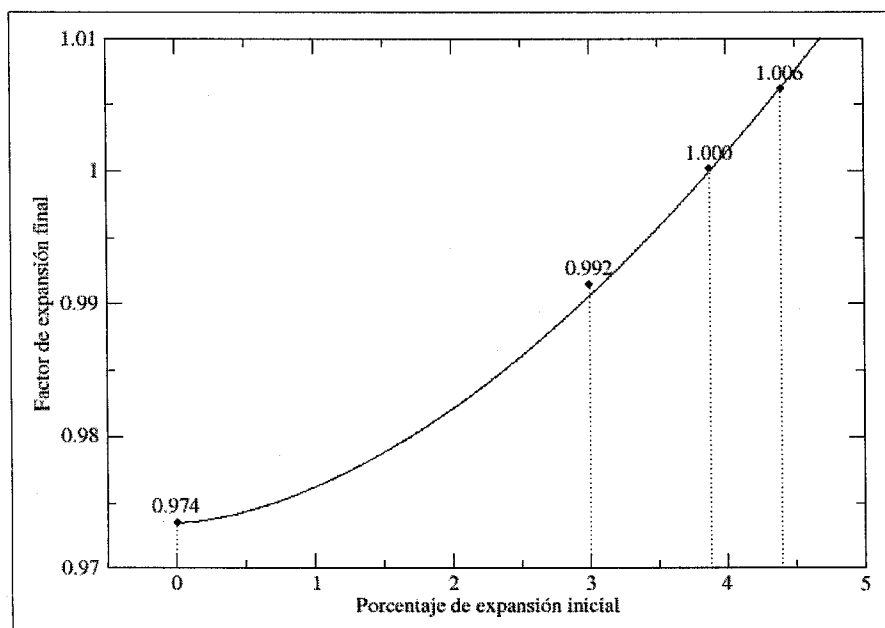


Figura 12. Gráfica de la contracción del parámetro de red  $a$  después del equilibrio de las simulaciones del cristal original. Los valores son relativos al valor original  $a=126.676 \text{ \AA}$ . Los primeros valores calculados fueron para las expansiones iniciales de 0%, 3% y 4.4%. A estos se les ajustó por mínimos cuadrados la función  $y=kx^m+0.974$ , donde  $k$  y  $m$  son parámetros ajustables. De esta se predice que el porcentaje de expansión inicial para el cual se retoma el valor original de  $a$ , es decir, el que corresponde a un factor de expansión 1, es 3.88%.

Así, a los hidrógenos faltantes de la proteína y de las moléculas de agua cristalográfica se les asignaron coordenadas por medio de la subrutina HBUILD [41] de la versión 282b del paquete computacional CHARMM [42] [43]. El sistema fue entonces solvatado en una caja de agua más grande

que la unidad asimétrica del cristal utilizando la subrutina SOLVATE del programa VMD [6]. A estas nuevas moléculas de agua se les eliminaron temporalmente sus átomos de hidrógeno, para así desechar todos los átomos de oxígeno que cayeran fuera de la unidad asimétrica. Como paso final en la construcción del modelo, los átomos de hidrógeno faltantes fueron generados con la subrutina HBUILD. Así, la unidad asimétrica consta de un monómero de la proteína (4173 átomos), así como 4324 moléculas de agua en las simulaciones del cristal original (17145 átomos en total), u 8455 en las simulaciones del hexámero en solución (29538 átomos en total).

Utilizando condiciones de frontera periódicas, los cálculos se realizaron sobre la unidad asimétrica del cristal de cada simulación, generando a partir de ella misma las imágenes cristalinas, formando así el hexámero de GlcN6P desaminasa y la masa de agua del cristal. Cabe señalar que para esto debió utilizarse el "sistema de coordenadas simétrico" de CHARMM.

### **7.1.2. Preparación previa a la fase de producción**

La minimización, al igual que las fases de calentamiento, equilibrio y producción, se llevaron a cabo usando la versión 282b del programa CHARMM [42] [43] con el potencial CHARMM27. Las interacciones energéticas de largo alcance se truncaron siguiendo el esquema sugerido por los autores, es decir, con una distancia umbral para la generación de la lista de vecinos de 14 Å, una distancia de corte de 12 Å, y modificando la energía de Van der Waals de 10 Å a 12 Å.

La minimización de cada simulación se llevó a cabo en etapas. Para las simulaciones Crist14 y Crist06 el sistema se minimizó con 4000 pasos del algoritmo Descensos Más Pronunciados (DMP) y 20000 pasos por Gradiente Conjugado (GC) manteniendo fijos los átomos de la proteína; posteriormente se minimizó el sistema completo con 15000 pasos de DMP y 30000 pasos de GC. Para las simulaciones del hexámero en solución el solvente se minimizó con 1000 pasos de DMP y 2000 pasos de GC, y el sistema sin restricciones se minimizó con 500 pasos de DMP y 7000 (para Sol28) ó 5000 (para Sol10) pasos de GC.

El equilibrio se dividió en una etapa para sólo el solvente (con la proteína fija) y otra para todo el sistema. El calentamiento sólo fue necesario para la simulación Crist14, en la cual las velocidades iniciales se asignaron para una temperatura de 200 K, estabilizándose alrededor de 300 K a los 100 pasos (0.2 ps) de dinámica (los cuales se incluyeron en el equilibrio del solvente). El equilibrio del solvente se extendió 0.2 ns para Crist14, Crist06 y Sol28, ó 0.1 ns para Sol10. Posteriormente se reasignaron las velocidades (correspondientes a una temperatura de 300 K) y se equilibró todo el sistema durante 0.2 ns en Crist14 y Sol28, o durante 0.1 ns en Sol10. Para Crist06, en cambio, en esta última etapa del equilibrio se conservaron fijos los C $\alpha$  de la proteína, extendiéndose 0.4 ns.



Después de la fase de equilibrio nuevamente se reasignaron las velocidades a 300 K para iniciar la etapa de producción, excepto para Sol10. En el caso de Crist14, además, hubo aun otra reasignación de velocidades a los 2 ns después de iniciada la fase de producción.

### **7.1.3. Imposiciones termodinámicas**

El ensamble utilizado en los cálculos es el isotérmico-isobárico, manteniendo una temperatura de referencia de 300 K por medio del algoritmo de Hoover [38] [39] con un parámetro de inercia térmica de 1000 kcal ps<sup>2</sup>, y manteniendo una presión de referencia de 1 atm usando un pistón de Langevin [37] de masa  $W=600$  Da con una frecuencia de colisión  $\gamma=10$  ps<sup>-1</sup> y un baño de temperatura de 300 K.

### **7.1.4. Propagación del movimiento**

La evolución del sistema se calculó con el algoritmo integrador Verlet Pídola. Pudo utilizarse un tamaño de paso de 2 fs, gracias a que los enlaces en los que participan átomos de hidrógeno se mantuvieron fijos con ayuda del algoritmo SHAKE [40], con una tolerancia de  $10^{-5}$  con respecto a los valores de referencia.

## **7.2. Métodos de análisis**

A continuación se presentan los métodos utilizados para analizar las trayectorias generadas durante las simulaciones. Cabe mencionar que los análisis para el sistema en una conformación dada pueden generalizarse para obtener resolución temporal al evaluar todas las configuraciones que haya adoptado este a lo largo de los cálculos.

### **7.2.1. Cuantificación de la diferencia entre estructuras: RMSD**

Dadas dos configuraciones de una proteína (una de referencia y otra por analizar) y, en general, de cualquier sistema de partículas, es posible cuantificar de manera general las diferencias estructurales entre ambas. Esto se logra calculando la "raíz del promedio de las diferencias cuadráticas" (RMSD, por sus siglas más comunes).

Denotando los conjuntos de vectores de posición de las  $N$  partículas de una macromolécula en dos conformaciones no necesariamente iguales con  $\{\vec{x}_i\}_{i=1}^N$  e  $\{\vec{y}_i\}_{i=1}^N$ , para  $i \in [1, N] \subseteq \mathbb{N}$ , se define la RMSD como

$$RMSD \left[ \left\{ \vec{x}_i \right\}_{i=1}^N, \left\{ \vec{y}_i \right\}_{i=1}^N, m_i \right] := \sqrt{\frac{\sum_{i=1}^N m_i (\vec{x}_i - \vec{y}_i)^2}{\sum_{i=1}^N m_i}} \quad \text{Ecuación 7.1}$$

donde  $m_i$  denota el factor con el que se pesa la contribución de la partícula  $i$  al valor total.

Así, si ambas estructuras son idénticas, es decir, si  $\vec{x}_i$  es igual a  $\vec{y}_i$ ,  $\forall i$ , entonces de la ecuación 7.1 resulta que la RMSD es nula. Por otra parte, entre más difieran las posiciones de los elementos, mayor será el valor de la RMSD calculada.

Los cálculos de RMSD de este trabajo se realizaron con la versión 282b del programa CHARMM [42] [43].

### 7.2.2. Análisis por proyección

Para identificar el punto de la oscilación en el que se encuentra la macromolécula a cada instante de tiempo, se proyectó la trayectoria de la simulación sobre el conjunto de vectores de desplazamiento que llevan a los C $^{\alpha}$  de la enzima de la conformación T cristalográfica (código PDB 1FSF, [4]) a la conformación R con el sitio alostérico ocupado (código PDB 1FRZ, [4]). Estos vectores modifican la estructura de la proteína a lo largo de su transición alostérica, transformando su estado medio (T cristalográfico) en uno de los extremos de la oscilación: la estructura R. Aproximando linealmente el desplazamiento que presentan estos átomos en la transición estructural, podemos asumir que este conjunto de vectores también lleva a la enzima desde su otro extremo de la oscilación (estructura a la que referiremos "meta T") a la conformación promedio (T cristalográfica). De manera análoga, los mismos vectores en sentidos opuestos la llevan del conformero R al T cristalográfico, y de este último a "meta T".

Al proyectar la trayectoria sobre este conjunto de vectores se obtiene un número escalar dentro del intervalo  $[-1,1] \subseteq \mathbb{R}$ , y cuyo valor indica la conformación que haya adoptado la macromolécula: -1 para "meta T", 0 para T cristalográfica, 1 para R, o valores intermedios.

El algoritmo del análisis por proyección fue programado en la aplicación Mathematica [44].

### 7.2.3. Análisis de Fourier

Es difícil identificar oscilaciones directamente en una gráfica de proyección de la trayectoria, debido a que en ella se encuentran superpuestos los movimientos de la molécula. Por esta razón, estos valores deben separarse en una suma de oscilaciones, a través de una descomposición de Fourier.

Este análisis toma un conjunto de datos que corresponden a la evolución de un parámetro a lo largo del tiempo, y lo separa en funciones senoidales de períodos distintos, indicando qué tan representativa es cada una para reconstruir la gráfica original. Así, en este caso se obtiene, para un intervalo de tiempo analizado, una medida cuantitativa del número de períodos que hayan presentado las oscilaciones más representativas del movimiento de la enzima.

Para ejemplificar este método, se analiza un conjunto de datos generados a partir de la función

$$f[x] = 2 \text{Sen}[0.95x] + 4 \text{Sen}[3.1x] + \text{Sen}[4.9x] + \text{Sen}[8.1x], \quad \text{Ecuación 7.2}$$

representados en la figura 13.

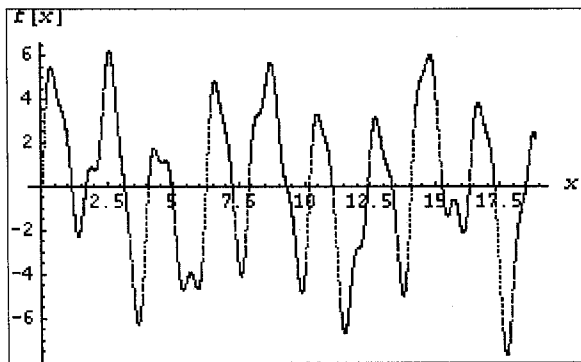


Figura 13. Datos generados a partir de la función  $f[x] = 2 \text{Sen}[0.95x] + 4 \text{Sen}[3.1x] + \text{Sen}[4.9x] + \text{Sen}[8.1x]$ , que es una combinación lineal de funciones periódicas. En este orden, cada término corresponde a una función con período  $2\pi/0.95$ ,  $2\pi/3.1$ ,  $2\pi/4.9$  y  $2\pi/8.1$ , y con amplitud 2, 4, 1 y 1. Imagen generada en Mathematica [44].

La descomposición de estos datos por el método de Fourier encuentra que existen cuatro señales representativas asociadas a oscilaciones que, a lo largo de la serie de datos, presentaron 3, 9, 15 y 24 períodos (figura 14).

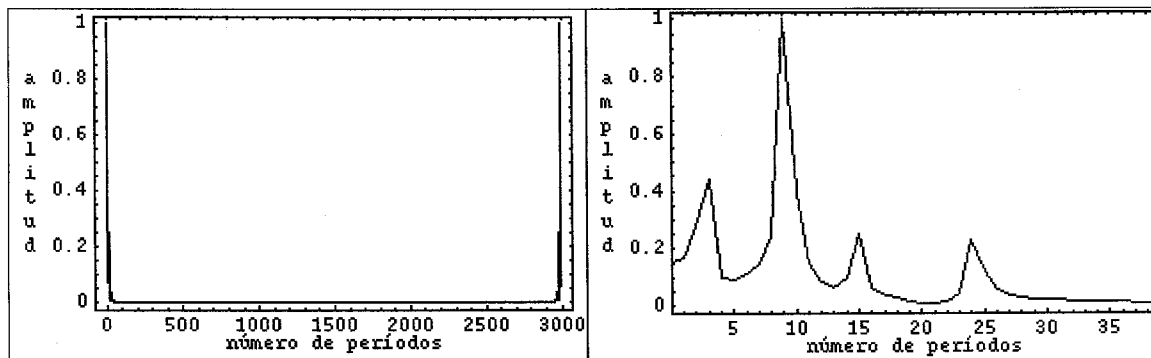


Figura 14. Análisis de Fourier sobre los datos graficados en la figura 13. En la parte izquierda se muestran las señales de todas las cantidades de períodos que hayan podido detectarse en el intervalo de los datos de la figura 13 (la gráfica es simétrica alrededor de los 1500 períodos). En la parte derecha se hace un acercamiento a las señales más representativas (de mayor amplitud). Estas se encuentran en la región correspondiente a las frecuencias más bajas (pocos períodos ocurridos), e indican que el conjunto de datos de entrada está compuesto por cuatro oscilaciones principales, una de las cuales presentó 3 períodos, otra (de mayor amplitud) alrededor de 9 períodos, y otras dos (de menor amplitud) 15 y 24 períodos. Imagen generada en Mathematica [44].

Así, el análisis de Fourier encuentra que los datos de la figura 13 (graficada de 0 a  $6\pi$ ) es una suma de cuatro funciones, de períodos aproximadamente de  $6\pi/3=2\pi$  (correspondiente a la señal de  $2\text{Sen}[0.95x]$ ),  $6\pi/9=2\pi/3$  ( $4\text{Sen}[3.1x]$ ),  $6\pi/15=2\pi/5$  ( $\text{Sen}[4.9x]$ ) y  $6\pi/24=\pi/4$  ( $\text{Sen}[8.1x]$ ). Además, estas funciones también fueron etiquetadas por orden de representatividad, pues en la figura 14 el pico más alto corresponde a la función de amplitud 4 ( $4\text{Sen}[3x]$ ), el de la mitad del tamaño a la función de amplitud 2 ( $2\text{Sen}[x]$ ), y los dos de un cuarto de su tamaño a las funciones de amplitud 1 ( $\text{Sen}[5x]$  y  $\text{Sen}[8x]$ ).

El análisis de Fourier tiene una incertidumbre intrínseca para la longitud de los períodos de cada oscilación, pues sólo considera valores que sean partes alícuotas de la longitud total del intervalo en el que se encuentren los datos analizados. Debido a esto, los valores que arroja para los períodos característicos deben interpretarse no como valores exactos, sino como intervalos de tiempo que contienen al valor exacto. Esto es, dado un punto asociado a  $n$  períodos en el intervalo analizado  $L$  (figura 14) se considera que, en lugar de representar un período de longitud  $T=L/n$ , su valor en realidad puede estar entre  $n-1/2$  y  $n+1/2$ , por lo que el valor del período que representa se encuentra dentro de un intervalo de error entre  $L/(n+1/2)$  y  $L/(n-1/2)$ . Siguiendo esta interpretación, los resultados para el análisis de Fourier de la figura 14 se presentan corregidos en la figura 15 como intervalos dentro de los cuales se encuentra el valor correcto. Además, para verificar que los resultados sean consistentes entre distintos subconjuntos de los datos, estos se calculan y se comparan sobre los datos de 0 a  $2\pi$ , de 0 a  $4\pi$ , y de 0 a  $6\pi$  (figura 15).

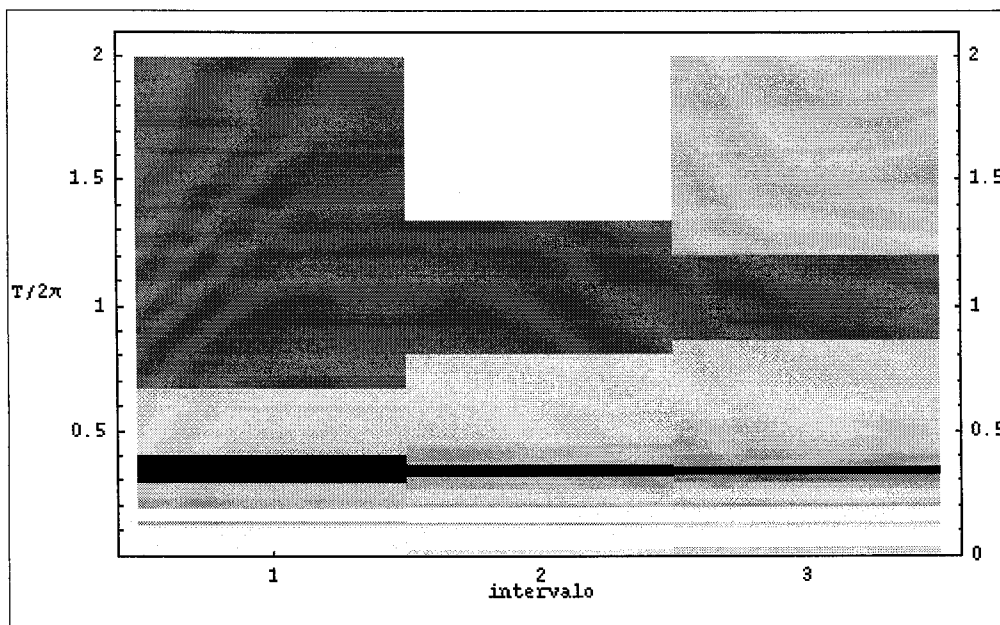


Figura 15. Análisis de Fourier sobre los datos graficados en la figura 13, donde ahora se incluyen en la representación los intervalos de error de los valores de los períodos. Separados por columnas se presentan los resultados del análisis sobre los datos de 0 a  $2\pi$  (intervalo 1), de 0 a  $4\pi$  (intervalo 2) y de 0 a  $6\pi$  (intervalo 3), los cuales deben ser consistentes entre ellos. La longitud de cada estimación del período se representa por la altura de un rectángulo, y la importancia de la oscilación a la que está asociado está dada por el nivel de negritud del recuadro. Cabe notar que el intervalo de análisis con más datos (intervalo 3) arroja resultados con una mejor resolución, es decir, los valores estimados para los períodos están acotados por rectángulos de menor altura. Imagen generada en Mathematica [44].

Como puede observarse, el análisis de Fourier efectivamente arroja resultados consistentes tomando distintos subconjuntos de los datos de entrada, pero la incertidumbre de los valores de los períodos característicos disminuye en tanto cuente con más datos para los cálculos. Así, de acuerdo con dichos valores calculados para el intervalo 3 en la figura 15, se observa que a partir de los rectángulos con mayor negritud se obtienen buenas aproximaciones a los valores de los períodos de las funciones originales. Estos intervalos de error (en múltiplos de  $2\pi$ ) son:  $[0.857, 1.200]$  (que contiene al valor 1.05 del período de la función  $2\text{Sen}[0.95x]$ );  $[0.316, 0.353]$  (que contiene al período 0.322 de la función  $4\text{Sen}[3.1x]$ );  $[0.194, 0.207]$  (conteniendo el período 0.204 de  $\text{Sen}[4.9x]$ );  $[0.122, 0.128]$  (con el período 0.124 de  $\text{Sen}[8.1x]$ ).

En este trabajo el análisis de Fourier se realizó en la aplicación Mathematica [44].

#### **7.2.4. Modos cuasiarmónicos**

Aproximando los movimientos de la proteína como movimientos oscilatorios, es posible descomponer su dinámica en una combinación de oscilaciones (de manera análoga a como el análisis de Fourier descompone una serie de datos en una combinación de funciones periódicas). Esto se logra

construyendo una matriz de correlación de los movimientos relativos entre todas las partículas del sistema, y calculando sus eigenvectores. Cada uno de estos últimos equivale a un conjunto de vectores en el espacio real que indican la amplitud y la dirección en la que oscila cada partícula a lo largo de una oscilación particular.

Los movimientos oscilatorios con cuya suma se aproxima el movimiento real son llamados "modos cuasiarmónicos", y se obtienen tantos como variables de posición existan, es decir, el triple del número de partículas. Es posible cuantificar la importancia de cada uno de ellos, y generalmente sólo unos pocos de los primeros modos cuasiarmónicos son significativos.

Los modos cuasiarmónicos presentados en este trabajo son calculados con el programa *Metaphor* (Ceruso, M. A., trabajo sin publicar).

## 8. RESULTADOS Y DISCUSIONES

Se construyó en la versión 282b del programa CHARMM [42] [43] el cristal de la glucosamina-6-fosfato desaminasa de *Escherichia coli*. Este corresponde a la estructura de código 1FSF del PDB [11], y pertenece al grupo espacial  $P6_322$ . Dado que el hexámero está constituido por imágenes cristalográficas de una de las subunidades, el análisis conformacional del oligómero se reduce al de un protómero. Así, los resultados presentados en este trabajo se derivan de los análisis hechos para uno de los protómeros.

A partir de este cristal y usando el mismo programa, se generaron con el método de Dinámica Molecular cuatro simulaciones del hexámero de GlcN6P desaminasa: dos dentro del cristal original y dos en solución, como se describe en la sección Metodología.

### 8.1. Resultados del equilibrio

El equilibrio de cada sistema se definió con base a la estabilización de su energía total, como se presenta en la figura 16.

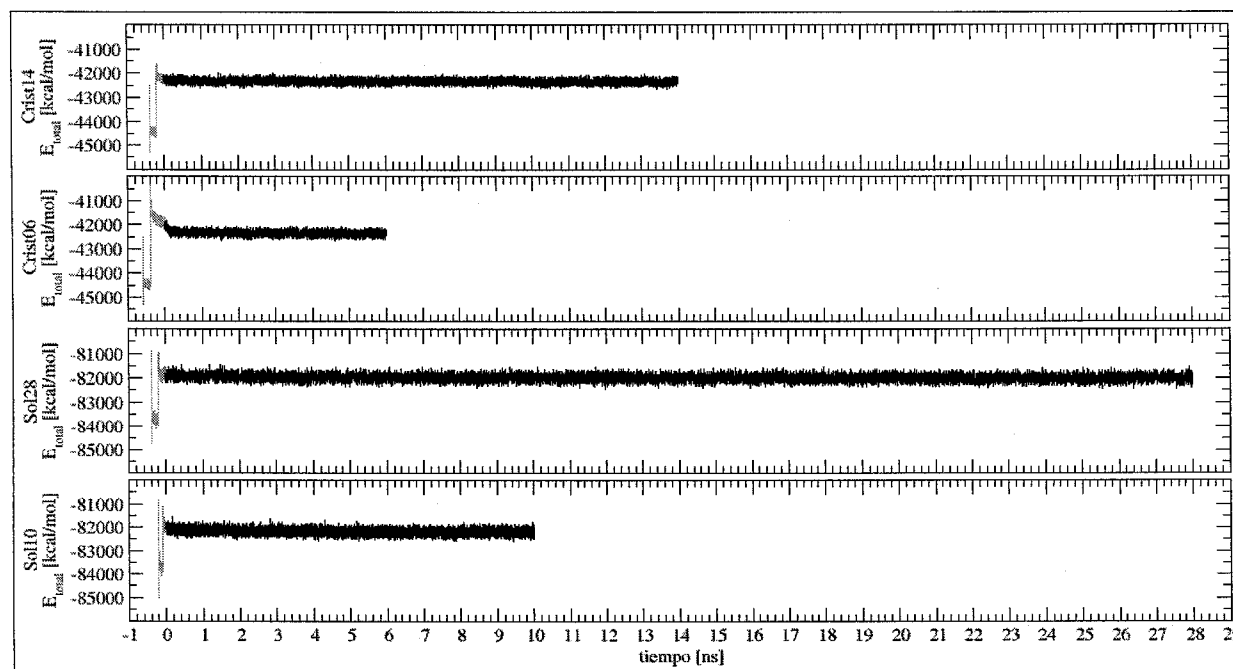


Figura 16. Gráficas de la energía total a lo largo de cada simulación. Cada punto representa un promedio a lo largo de los 0.2 ps anteriores. La fase de equilibrio corresponde a los tiempos negativos, representada en color gris. En cada caso el salto inicial de la energía se debe a la asignación de velocidades a los átomos de la proteína, con lo cual incrementa la contribución de la energía cinética a la energía total del sistema. El equilibrio de cada simulación se alcanza cuando se estabiliza su gráfica de energía, de forma que la etapa de producción comienza a tiempo  $t = 0$  ns (en color negro). Cabe aclarar que, en el caso de Crist06, en este punto se liberan las restricciones sobre los  $C^\alpha$  de la proteína, de manera que el equilibrio se extiende hasta  $t = 0.1$  ns.

Como consecuencia de la contracción del cristal durante el equilibrio, los parámetros de red de la celda de cada simulación se estabilizaron alrededor de los valores deseados:  $a=b=126.676 \text{ \AA}$  y  $c=139.326 \text{ \AA}$  para Crist14 y Crist06, o  $a=b=151.9 \text{ \AA}$  y  $c=167.0 \text{ \AA}$  para Sol28 y Sol10. Dado que  $a$  es igual a  $c$  y proporcional a  $b$ , sólo fue necesario monitorizar  $a$  para asegurar que la celda de cada sistema alcanzara las dimensiones deseadas. Así, se encuentra que a lo largo de la etapa de producción su valor promedio es de  $126.595 \pm 0.114 \text{ \AA}$  para Crist14,  $126.566 \pm 0.116 \text{ \AA}$  para Crist06 (descartando los primeros 0.1 ns),  $151.852 \pm 0.108 \text{ \AA}$  para Sol28, y  $151.859 \pm 0.109 \text{ \AA}$  para Sol10, donde las incertidumbres son las desviaciones estándar. La evolución de  $a$  se muestra en la figura 17.

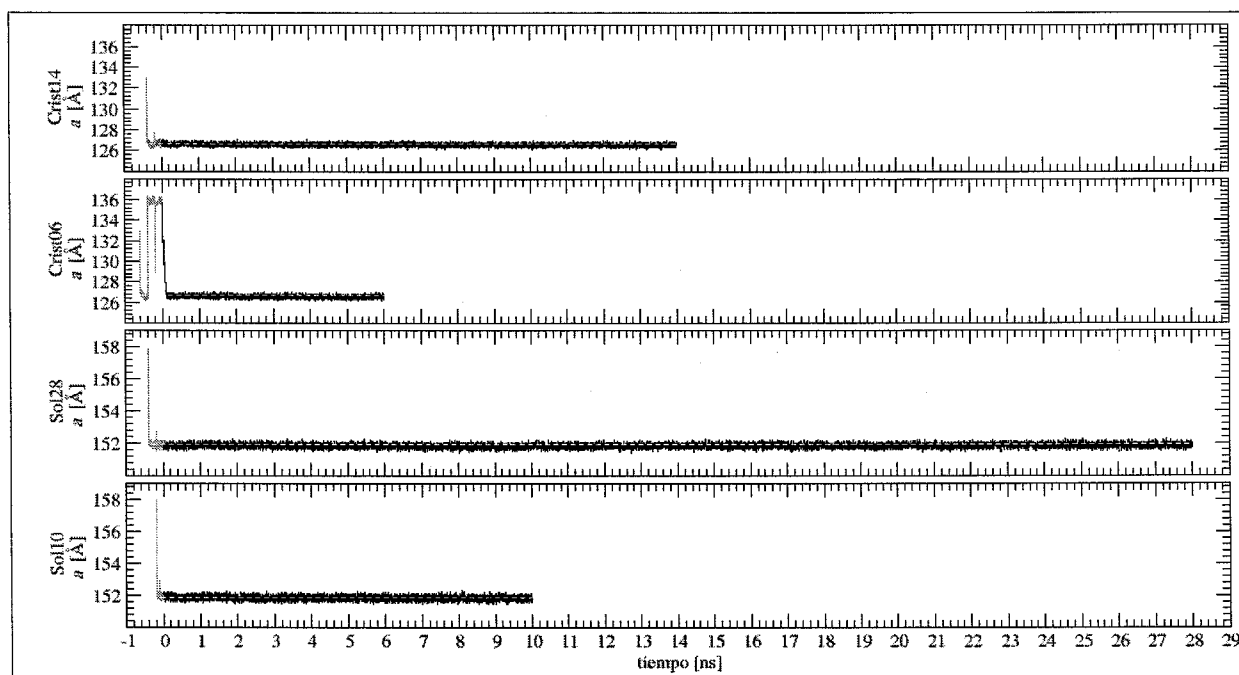


Figura 17. Gráficas de la evolución del parámetro de red  $a$  a lo largo de las simulaciones, el cual es representativo de las dimensiones de la celda cristalina con respecto al tamaño deseado. Cada punto representa un promedio a lo largo de los 0.2 ps anteriores. La fase de equilibrio corresponde a los tiempos negativos, representada en color gris. La línea blanca segmentada indica el valor deseado para cada sistema ( $126.676 \text{ \AA}$  para Crist14 y Crist06, y  $151.9 \text{ \AA}$  para Sol28 y Sol10).

Con base en estos resultados, se concluye que en la etapa de producción de cada simulación se genera información confiable acerca de la dinámica de la GlcN6P desaminasa, en los entornos cristalinos de interés.

## 8.2. Validación del análisis por proyección

Para validar este análisis se compararon sus resultados con los de la RMSD, cuyo cálculo es un método establecido que se toma aquí como referencia. Esta comparación es consistente entre las cuatro



simulaciones generadas en este trabajo, y se ejemplifica con los resultados de la simulación Sol28 en la figura 18.

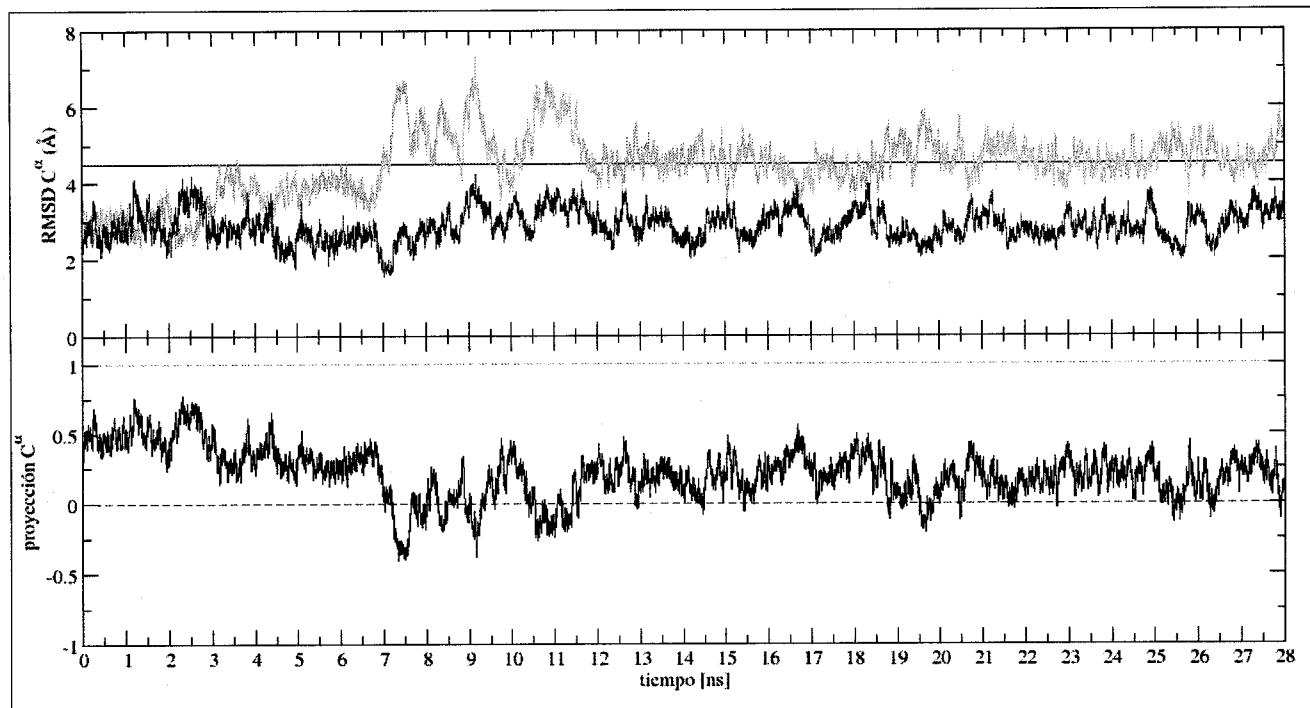


Figura 18. Validación, mediante la RMSD, del análisis por proyección para los  $C^\alpha$  de la trayectoria. En la gráfica superior se representa la RMSD de la estructura de la enzima a lo largo de la simulación (cada 0.2 ps) contra la conformación R (línea naranja continua) y contra la conformación T cristalográfica (línea verde continua), así como el valor de la RMSD entre estas dos estructuras (línea constante azul). En la gráfica inferior se muestra la proyección de la estructura de la enzima a lo largo de la simulación (cada 0.2 ps) sobre el conjunto de vectores que la transforman a lo largo de la transición alostérica (línea negra), así como el valor de la proyección de la conformación R (línea naranja punteada, de valor 1) y de la conformación T cristalográfica (línea verde punteada, de valor 0). La proyección de la conformación meta T correspondería al valor -1.

La validación del análisis por proyección a través de las gráficas de la figura 18 se construye a partir de diversas observaciones:

- Se sabe que la RMSD entre las dos estructuras de referencia (T cristalográfica y R) es de 4.5 Å, así que cuando la proteína adopte una de estas conformaciones en la simulación, deberá presentar una RMSD de 4.5 Å con respecto a la otra (su gráfica en la parte superior de la figura 18 cruzará la línea azul). Efectivamente, los momentos en los que el valor de la proyección es 0 (estructura T cristalográfica) coinciden con aquellos en los que la gráfica de RMSD contra R (naranja) pasan por la línea azul (por ejemplo, a los 7 ns), y cuando la proyección se acerca a 1 (estructura R) la RMSD contra T cristalográfica (verde) se acerca a la línea azul de referencia (por ejemplo, alrededor de 2.5 ns).

- Dado que al acercarse la enzima a la conformación R debe disminuir su RMSD contra R (línea naranja continua) al mismo tiempo que aumente el valor de la proyección (línea negra, acercándose a 1) y viceversa, estas dos gráficas deben estar anticorrelacionadas. Esto se cumple para este análisis a lo largo de toda la trayectoria.
- Lo anterior también debe cumplirse para la RMSD contra T cristalográfica (línea verde continua), pero sólo para las estructuras entre la T cristalográfica y la R (zona de proyección entre 0 y 1), pues cuando se encuentre entre meta T y T cristalográfica (proyección entre -1 y 0) aumentará la RMSD cuando se aleje de T cristalográfica disminuyendo el valor de su proyección (hacia -1), y estas gráficas deberán estar anticorrelacionadas. Nuevamente, el comportamiento esperado coincide con el observado (por ejemplo: correlación antes de 7 ns, y anticorrelación cerca de 11 ns).
- Para valores de proyección entre 0 y 1 el acercamiento de la enzima a una de las dos conformaciones debe causar que se aleje de la otra, y que por lo tanto las gráficas de RMSD estén anticorrelacionadas, mientras que en la zona de proyección entre 0 y -1 su estructura se aleja o se acerca a ambos conformeros por igual, por lo que las dos gráficas de RMSD deben estar correlacionadas. El primer caso se observa por ejemplo antes de 7 ns y, el segundo, alrededor de 11.5 ns.
- La conformación justo a medio camino entre T cristalográfica y R (proyección de 0.5) debe presentar valores iguales de RMSD contra ambas conformaciones de referencia (deben coincidir las líneas naranja y verde), lo cual sí se observa en las gráficas (por ejemplo, al inicio de la trayectoria, o poco después de los 16.5 ns).
- Durante la simulación la máxima similitud estructural con R (mínima RMSD, línea naranja) debe coincidir con los máximos valores de proyección (línea negra, cercana a 1) y, similarmente, la mínima RMSD contra T cristalográfica (línea verde) debe coincidir con valores de proyección cercanos a 0. Este es el comportamiento observado, por ejemplo: alrededor de 2.5 ns para R, o cerca de 7 ns para T cristalográfica.

El valor asociado a la proyección es, entonces, indicativo de la conformación que adopte el estado T cinético de la GlcN6P desaminasa con respecto a sus estructuras T y R cristalográficas.

### **8.3. Conformación R desfavorecida en el cristal original**

En ambas simulaciones del cristal original (Crist14 y Crist06) la proteína presenta una clara tendencia a adoptar conformaciones lejanas a la activa, como lo muestran sus análisis por proyección sobre los C<sup>α</sup> (figura 19).

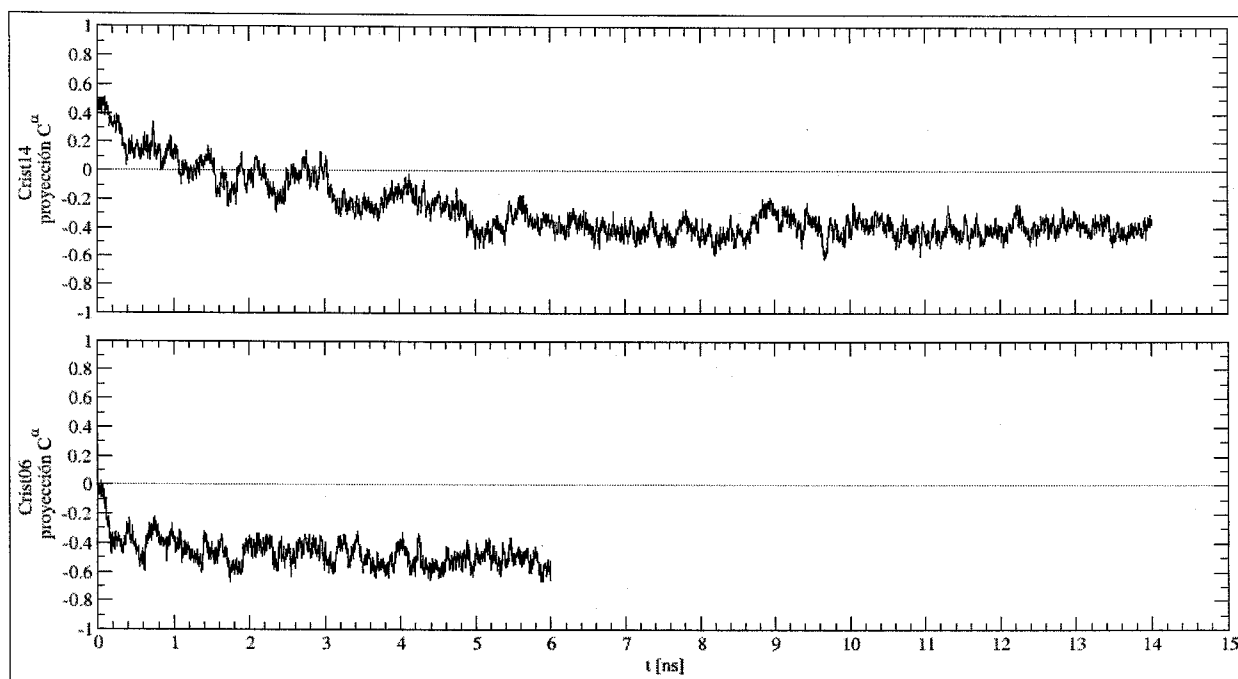


Figura 19. Gráficas de la proyección  $p$  sobre los  $C^\alpha$  de las simulaciones Crist14 y Crist06 en la etapa de producción. La conformación R corresponde a  $p=1$ , T cristalográfica corresponde a  $p=0$  (línea gris), y "meta T" a  $p=-1$ . En ambas simulaciones es evidente que la proteína tiende a adoptar conformaciones entre T y "meta T", lejos de R.

Esta tendencia de la proteína en el cristal original a visitar estados lejanos a R se explica por el hecho de que, en el grupo espacial en el que se realizaron los cálculos (P6<sub>3</sub>22), la transición de la proteína del estado T a la conformación activa implica un choque estérico entre los residuos 95 a 100 de los protómeros de hexámeros vecinos, como se muestra en la figura 20.

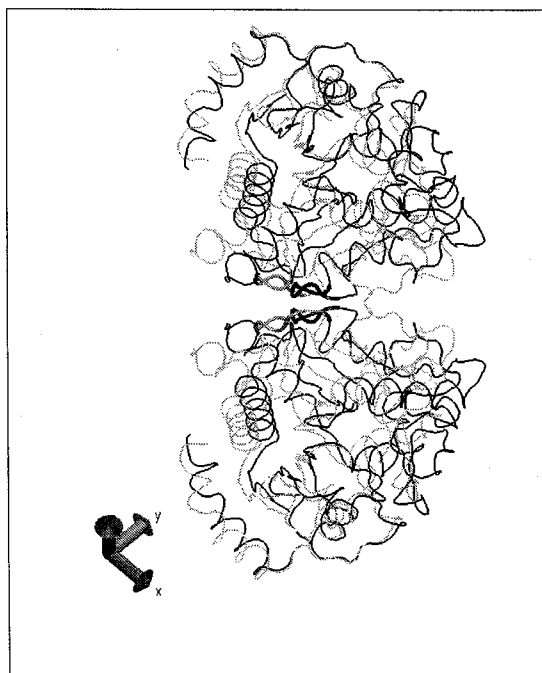


Figura 20. Traza de las estructuras T (gris) y R (negro) cristalográficas de uno de los protómeros del trímero inferior de un hexámero (arriba), y el protómero más próximo del trímero superior del hexámero vecino (abajo). Los residuos más cercanos entre estas subunidades (del 95 al 100) se representan con una línea más gruesa. Los conformeros R fueron trasladados a las posiciones que ocuparían en el grupo espacial  $P6_322$ . Imagen generada en VMD [6].

Así, se concluye que el estado R está desfavorecido en el grupo espacial  $P6_322$ . Esta puede ser la razón por la que los cristales de la proteína en su conformación activa crecen en un grupo espacial distinto ( $R32$ ). Además, también explica el hecho de que los cristales del conformero T remojados en una solución de algún ligante (fructosa-6-fostato 10mM o 2-desoxi-2-amino-D-glucitol-6-fosfato 100 $\mu$ M) no presentan la transición alostérica, sino que conservan la estructura T libre de ligantes [4].

Esta situación no existe en las simulaciones que emulan al hexámero en solución ya que, como se mencionó en la sección de metodología, estos sistemas se construyeron asegurando una distancia interhexamérica mayor a 14 Å. Por lo anterior, los análisis posteriores se realizan sobre Sol28 y Sol10.

#### **8.4. Períodos de oscilación**

Los períodos de las oscilaciones características del movimiento de la proteína se calcularon por medio del análisis de Fourier de las gráficas del análisis por proyección sobre los  $C^\alpha$ . Para sustentar sus resultados, este análisis se realizó a lo largo de distintos intervalos de las trayectorias de las simulaciones Sol28 y Sol10 (cada uno con valores de proyección cercanos en sus extremos), encontrando resultados consistentes. Los datos de Sol28 (figura 18) fueron analizados en las regiones listadas en la tabla II, y se muestran en la figura 21.

Tabla II. Intervalos en los que se realizó el análisis de Fourier sobre los datos de proyección de los  $C^\alpha$  de Sol28 de la figura 18, para sustentar sus resultados.

intervalo	inicio [ns]	final [ns]
1	7.0216	11.5562
2	7.0216	15.4088
3	7.0216	19.5084
4	7.0216	22.7336
5	0.0172	16.9902
6	0.0172	22.9376
7	0.0172	27.3628

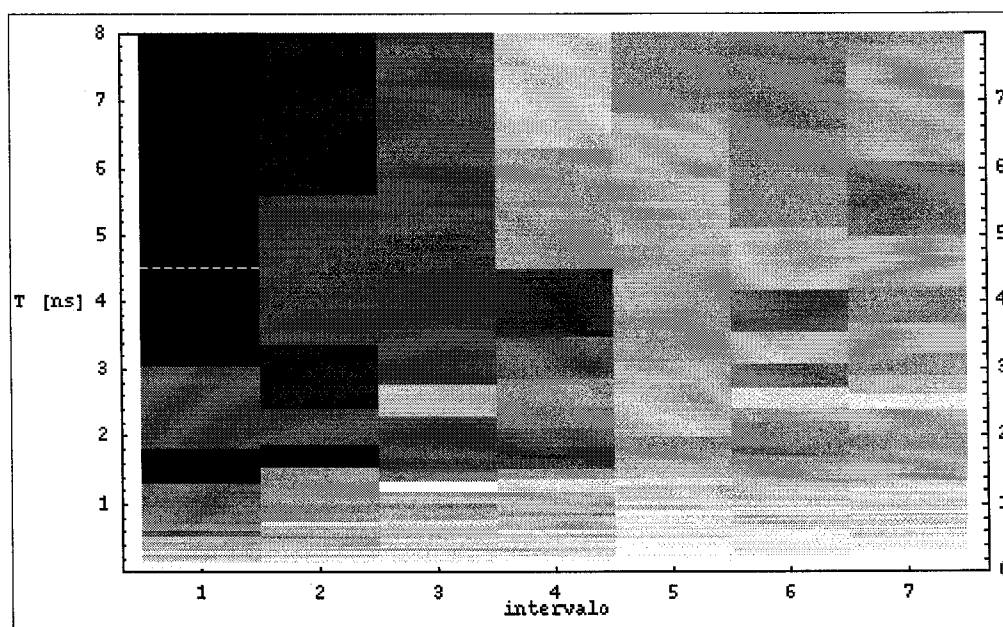


Figura 21. Estimaciones de los períodos de cada oscilación representativa encontrada por el análisis de Fourier en Sol28, para cada intervalo de la tabla II (cfr. figura 15). Las líneas blancas segmentadas, cuando aparecen dentro de la escala, indican la longitud de cada región analizada. Sólo se muestran las señales de los períodos que cuentan con una buena resolución. El propósito de esta figura es mostrar que las estimaciones de los períodos característicos que este método arroja son consistentes independientemente del intervalo de análisis. Imagen generada en Mathematica [44].

Tomando en consideración las zonas de buena resolución para cada intervalo de análisis, puede concluirse que el análisis de Fourier sobre la gráfica de proyección de la figura 15 arroja resultados similares entre ellos. Esto se cumple también para el resto de los análisis de Fourier de este trabajo.

Las estimaciones para los períodos de las oscilaciones características en la simulación Sol28 también fueron comparadas realizando el análisis sobre las proyecciones de los  $C^\alpha$  de distintas regiones de la proteína (graficadas en la figura 22), encontrando las mismas oscilaciones representativas tanto

para el protómero completo como para el módulo interno y el módulo externo (con o sin la tapa del sitio activo), como se muestra en la figura 23.

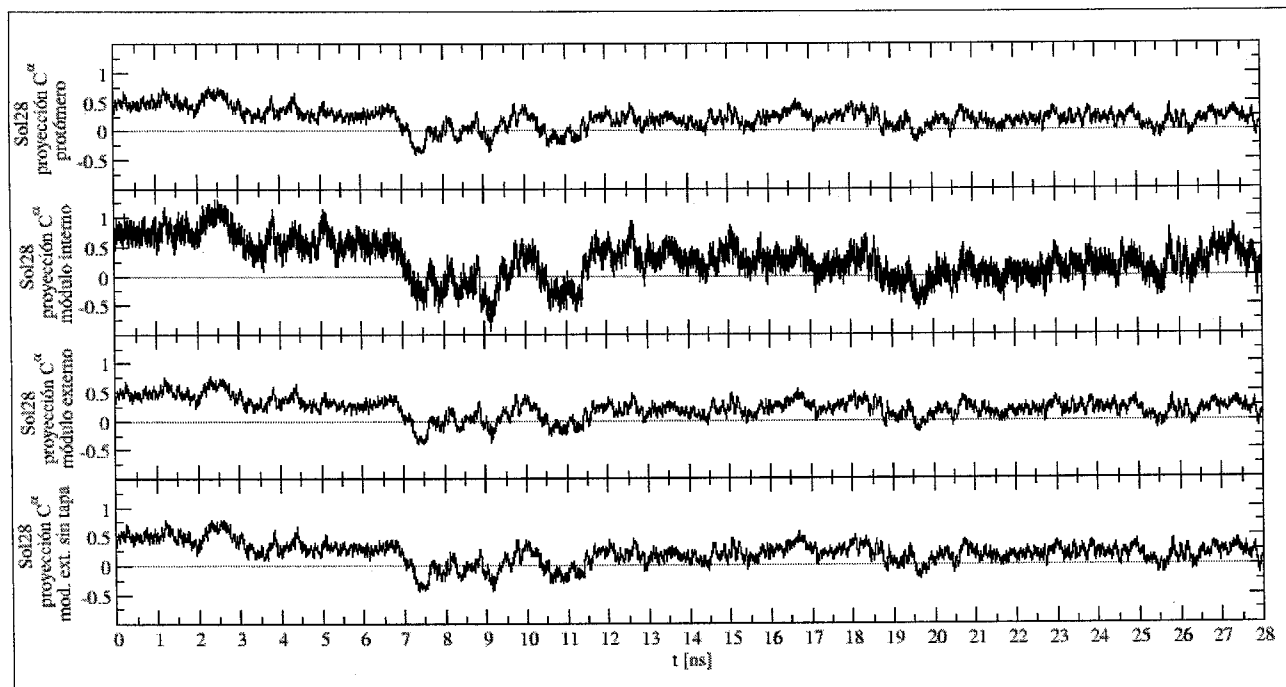


Figura 22. Gráficas del análisis por proyección sobre los  $C^\alpha$  de distintas zonas de la proteína en la simulación Sol28. Las líneas grises indican el valor correspondiente a la conformación T cristalográfica.

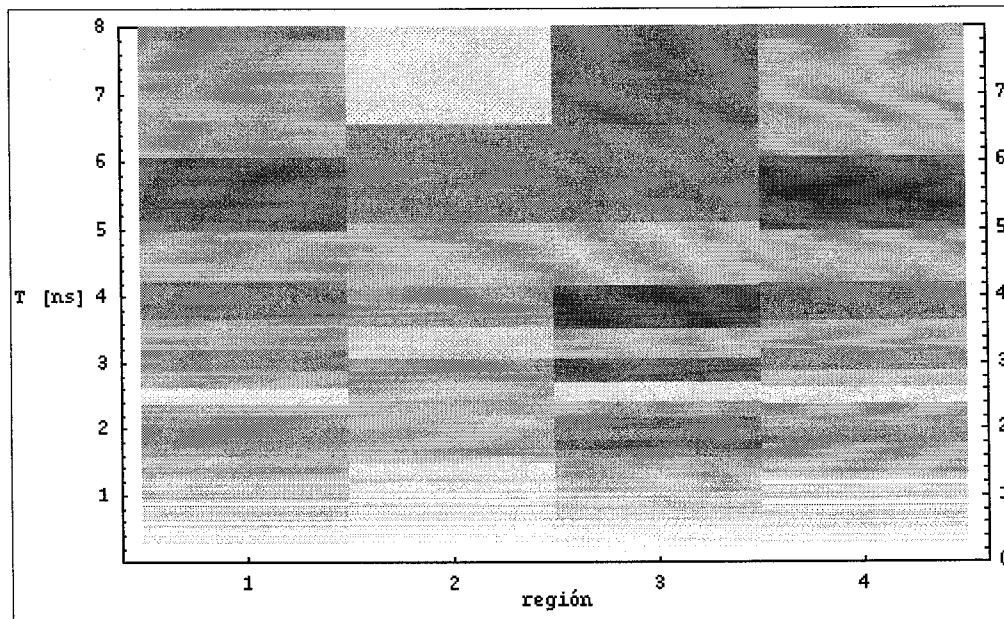


Figura 23. Estimaciones de los períodos de cada oscilación representativa encontrada por el análisis de Fourier, para distintas regiones de la proteína: protómero completo (región 1, cfr. figura 21), módulo interno (región 2), módulo externo (región 3), módulo externo sin tapa (región 4). Los resultados provienen de los intervalos de análisis más largos, y sólo se muestran las señales de los períodos que cuentan con una buena resolución. Imagen generada en Mathematica [44].

Antes de cuantificar los períodos de oscilación de la proteína, se les da un mayor sustento comparándolos con los obtenidos a partir de la proyección sobre los  $C^\alpha$  del protómero de Sol10 (figura 24), encontrando que sus valores son consistentes aun entre estas dos simulaciones, como se muestra en la figura 25.

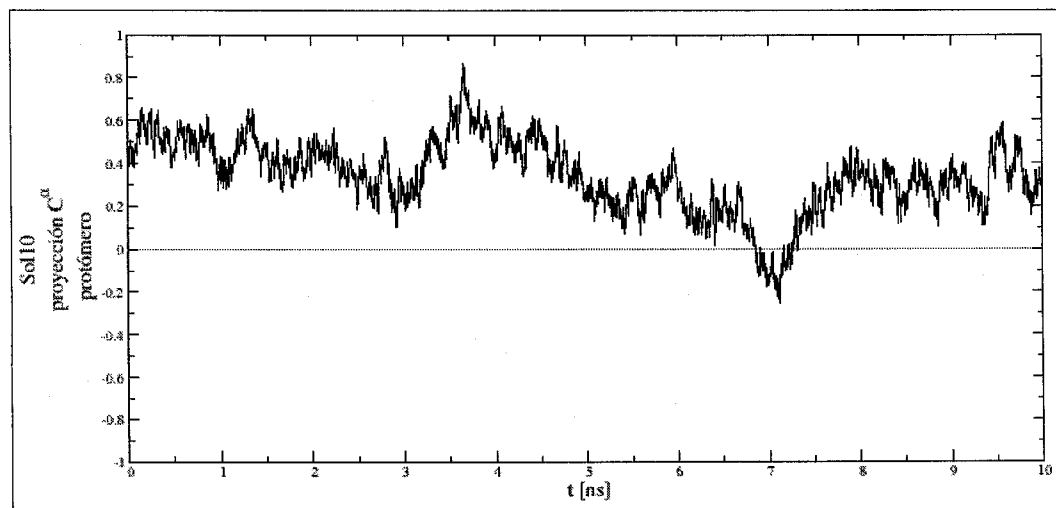


Figura 24. Gráfica del análisis por proyección sobre los  $C^\alpha$  del protómero en la simulación Sol10. La línea gris indica el valor correspondiente a la conformación T cristalográfica.

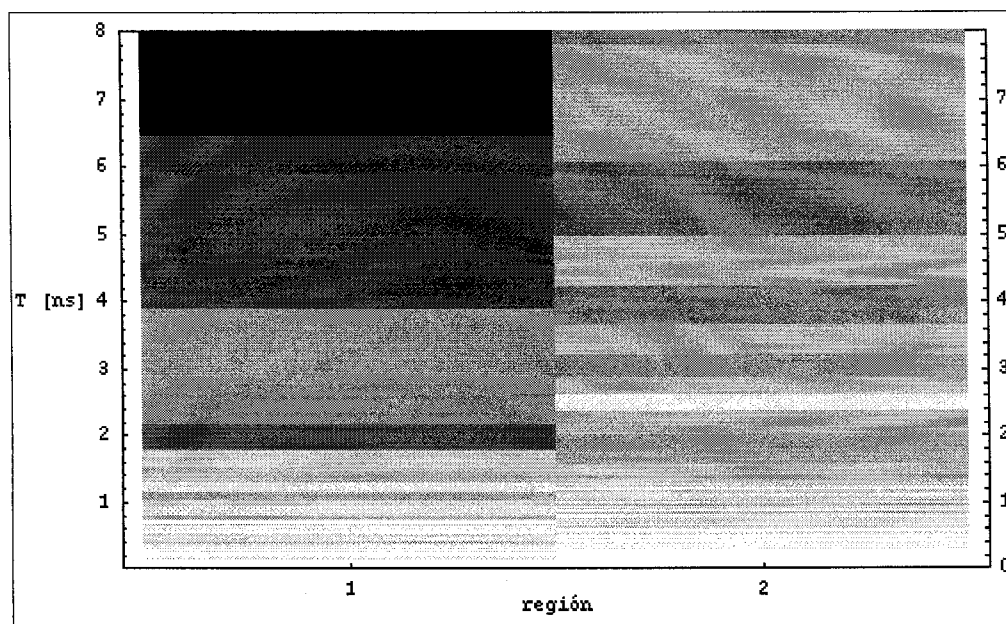


Figura 25. Estimaciones de los períodos de cada oscilación representativa encontrada por el análisis de Fourier, para los  $C^\alpha$  del protómero completo en las dos simulaciones del hexámero en solución: Sol10 (región 1) y Sol28 (región 2, cfr. figuras 21 y 23). Los resultados presentados son los de los intervalos de análisis más largos de cada simulación. A pesar de que, debido al corto tamaño de Sol10 su resolución por encima de  $T=3$ ns es demasiado baja, ambas simulaciones sí presentan estimaciones consistentes de los períodos de las oscilaciones características por debajo de este valor. Imagen generada en Mathematica [44].

Una vez validado de esta forma el cálculo de los períodos de los movimientos representativos de la proteína, pueden cuantificarse estos valores. Así, utilizando los resultados de mayor resolución (intervalo de análisis más largo de Sol28, figura 25), se encuentra que existen oscilaciones representativas del movimiento de la proteína con períodos cercanos a 5.5 ns, 4 ns, 3 ns, 2 ns y 1.4 ns. Los intervalos de error que acotan estos valores son, en orden de representatividad, [4.97, 6.08] ns, [3.65, 4.21] ns, [2.88, 3.22] ns, [1.56, 2.19] ns y [1.33, 1.40] ns.

También se buscaron los movimientos representativos de la tapa del sitio activo mediante análisis similares, encontrando que en general son distintos a los del resto de la proteína. En la figura 26 se muestran las proyecciones sobre sus  $C^\alpha$  y, en la figura 27, la estimación de sus períodos.

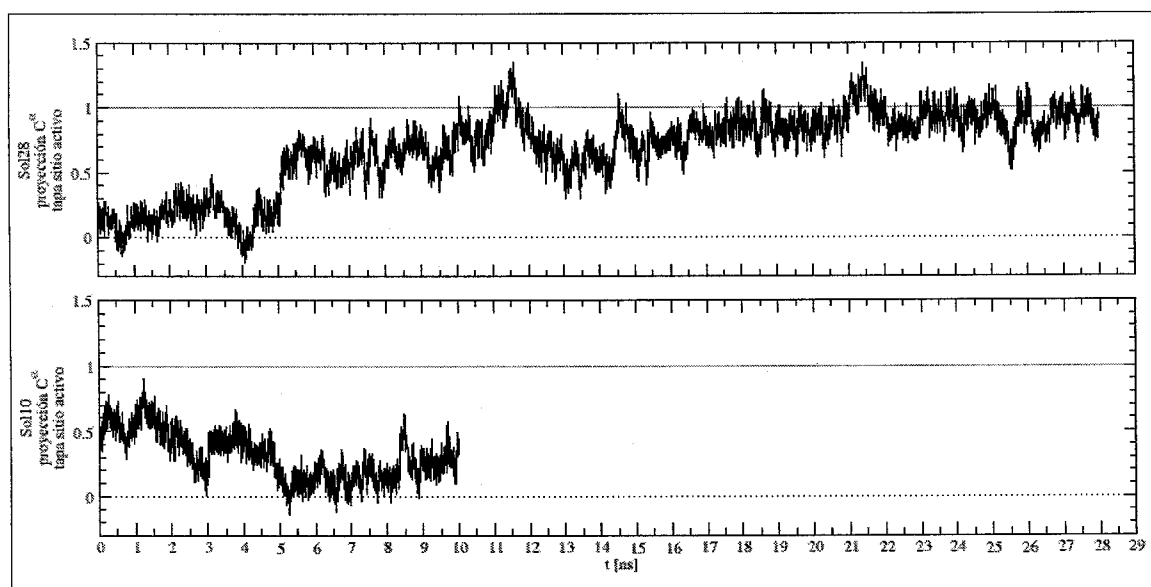


Figura 26. Gráfica del análisis por proyección sobre los  $C^\alpha$  de la tapa del sitio activo en las simulaciones Sol28 y Sol10. En este caso el valor de proyección  $p=1$  representa la tapa abierta (línea gris continua), mientras que  $p=0$  representa la tapa cerrada (línea negra punteada), de acuerdo con las conformaciones de referencia del archivo de coordenadas 1FS5 [4]. A lo largo de cada trayectoria la proteína fue orientada con el módulo externo (sin la tapa) de la conformación de referencia, con el propósito de eliminar los movimientos ajenos a la tapa en estos cálculos.



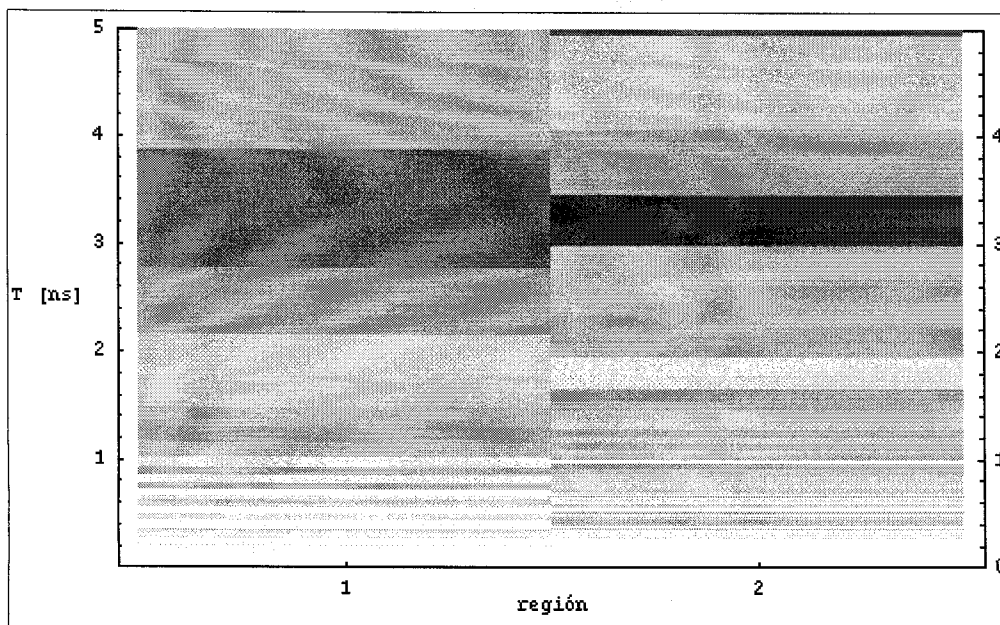


Figura 27. Estimaciones de los períodos de cada oscilación representativa encontrada por el análisis de Fourier, para los  $C^{\alpha}$  de la tapa del sitio activo en las dos simulaciones del hexámero en solución: Sol10 (región 1) y Sol28 (región 2). Los resultados presentados provienen de los intervalos de análisis más largos de cada simulación, y sólo se muestran las señales de los períodos que cuentan con una buena resolución. Imagen generada en Mathematica [44].

A pesar de presentar distintos comportamientos de la tapa del sitio activo, ambas simulaciones coinciden en los valores de algunos períodos representativos: alrededor de 3 ns, 2.2 ns, 1.2 ns y 0.9 ns. De acuerdo al análisis con mayor resolución de Sol28 (figura 27), los intervalos de error que acotan estos valores son, en orden de representatividad, [2.98, 4.06] ns, [1.44, 1.65] ns (este parece ser característico de Sol28), [0.91, 0.95] ns, [0.65, 0.67] ns (característico de Sol28), [1.94, 2.35] ns, [1.21, 1.28] ns y [1.09, 1.14] ns. De estos, los períodos de las dos oscilaciones más lentas son comparables con las encontradas para el movimiento global de la proteína, e incluso podrían estar relacionadas con las de períodos [2.88, 3.22] ns y [1.56, 2.19] ns. Sin embargo, la mayoría de las oscilaciones características de la tapa del sitio activo son más rápidas que las del protómero completo, lo cual es consistente con el hecho de que esta es mucho menos masiva.

### **8.5. Formación del sitio activo y del sitio alostérico**

La cavidad del sitio activo se define a partir de las distancias entre el  $C^{\beta}$  de Arg172 (en la tapa del sitio activo), el  $C^{\beta}$  de Lys208 (en el módulo interno) y el átomo N de Gly42 (en el módulo interno), y se considera que se ha formado cuando estas distancias se acercan de manera concomitante a sus valores correspondientes en la estructura R cristalográfica. La evolución de estas distancias a lo largo de Sol28 y Sol10 se presentan en las figuras 28 y 29, respectivamente.

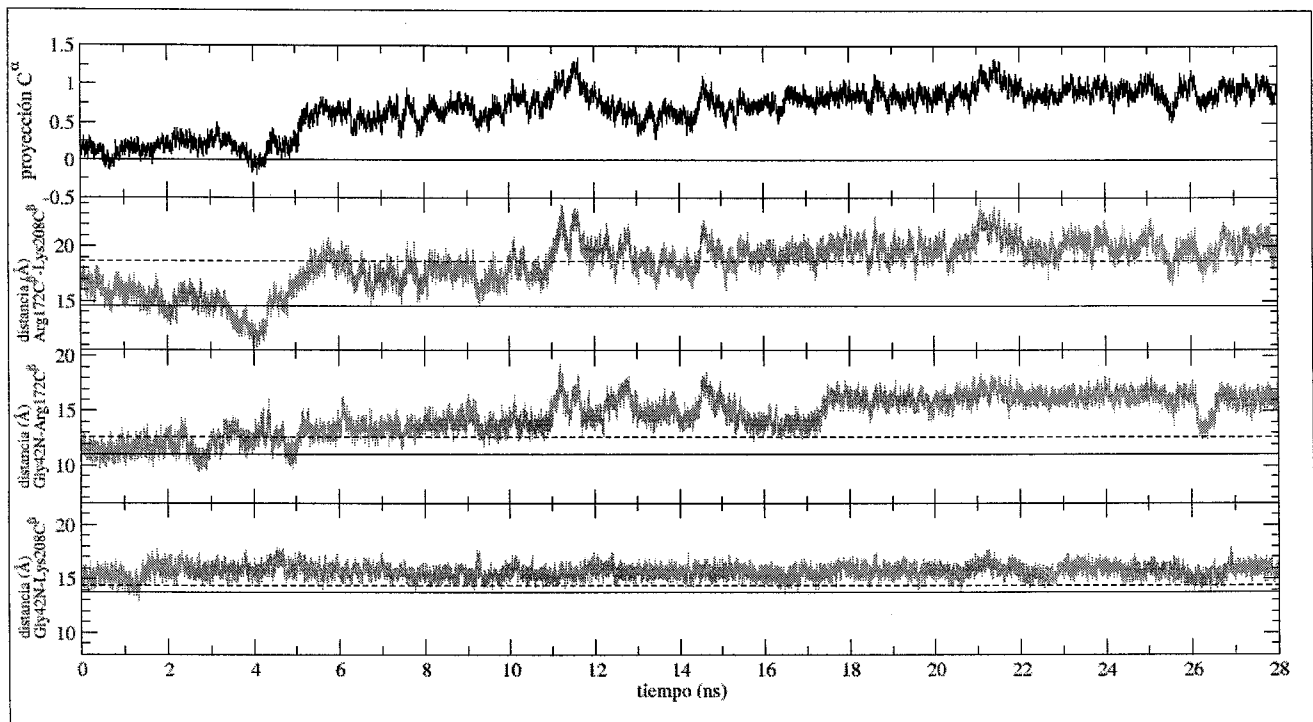


Figura 28. Formación del sitio activo a lo largo de Sol28. El panel superior muestra en negro la gráfica de la proyección de la tapa del sitio activo (cfr. figura 26). En los otros tres paneles las gráficas en gris corresponden a las tres distancias en base a las cuales se define la formación del sitio activo, las líneas negras continuas indican el valor de referencia en R y, las líneas negras segmentadas, el valor de referencia en T.

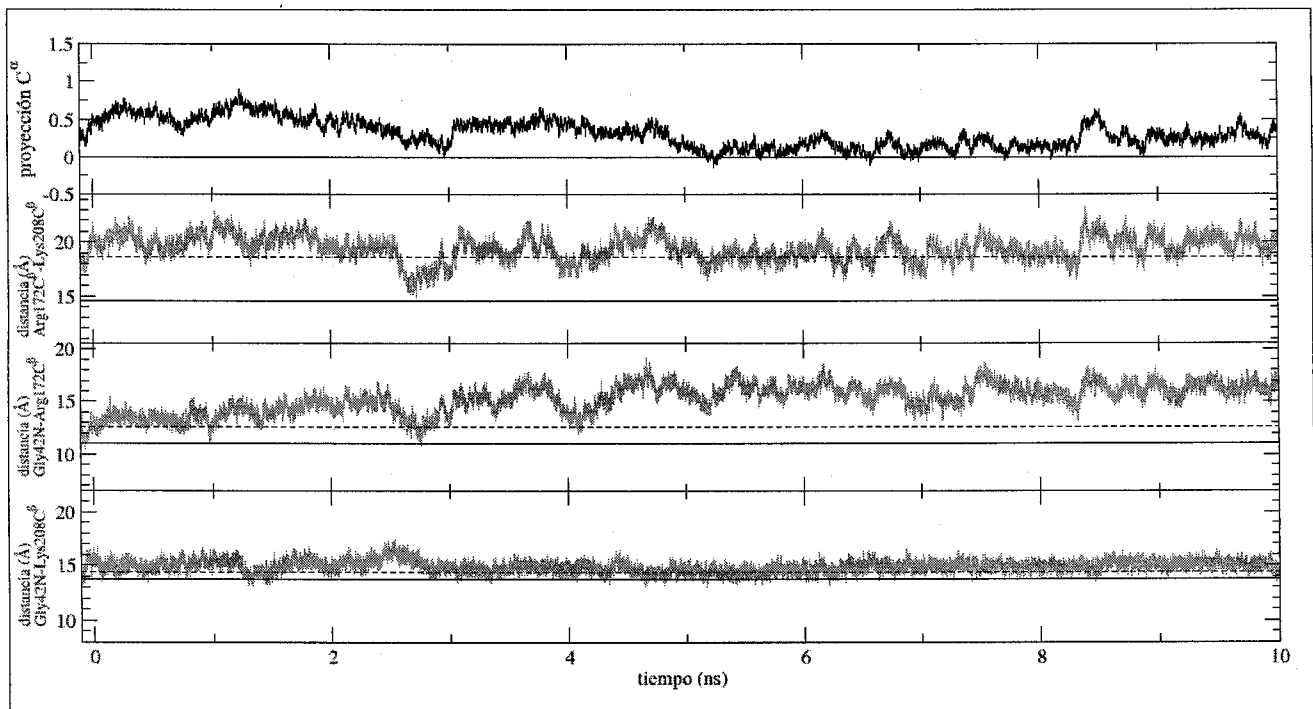


Figura 29. Formación del sitio activo a lo largo de Sol10. El panel superior muestra en negro la gráfica de la proyección de la tapa del sitio activo (cfr. figura 26). En los otros tres paneles las gráficas en gris corresponden a las tres distancias en base a las cuales se define la formación del sitio activo, las líneas negras continuas indican el valor de referencia en R y, las líneas negras segmentadas, el valor de referencia en T.

De estas dos figuras se hace evidente que la formación del sitio activo depende en gran medida del estado de la tapa. A pesar de que este no se formó en Sol10, sí lo hizo en algunas ocasiones al inicio de Sol28. La poca ocupación de este estado de la proteína puede deberse a que se requiere de la presencia de algún ligante para estabilizar la formación del sitio activo.

De manera similar, el sitio alostérico entre dos protómeros vecinos se define a partir de las distancias entre el átomo N de Met1 de uno de ellos y los C<sup>β</sup> de Arg158 y Lys160 de la otra subunidad. También en este caso se considera que está formado cuando estas distancias se acercan de manera concomitante a sus valores correspondientes en la estructura R cristalográfica. La evolución de estas distancias a lo largo de Sol28 y Sol10 se presentan en las figuras 30 y 31, respectivamente.

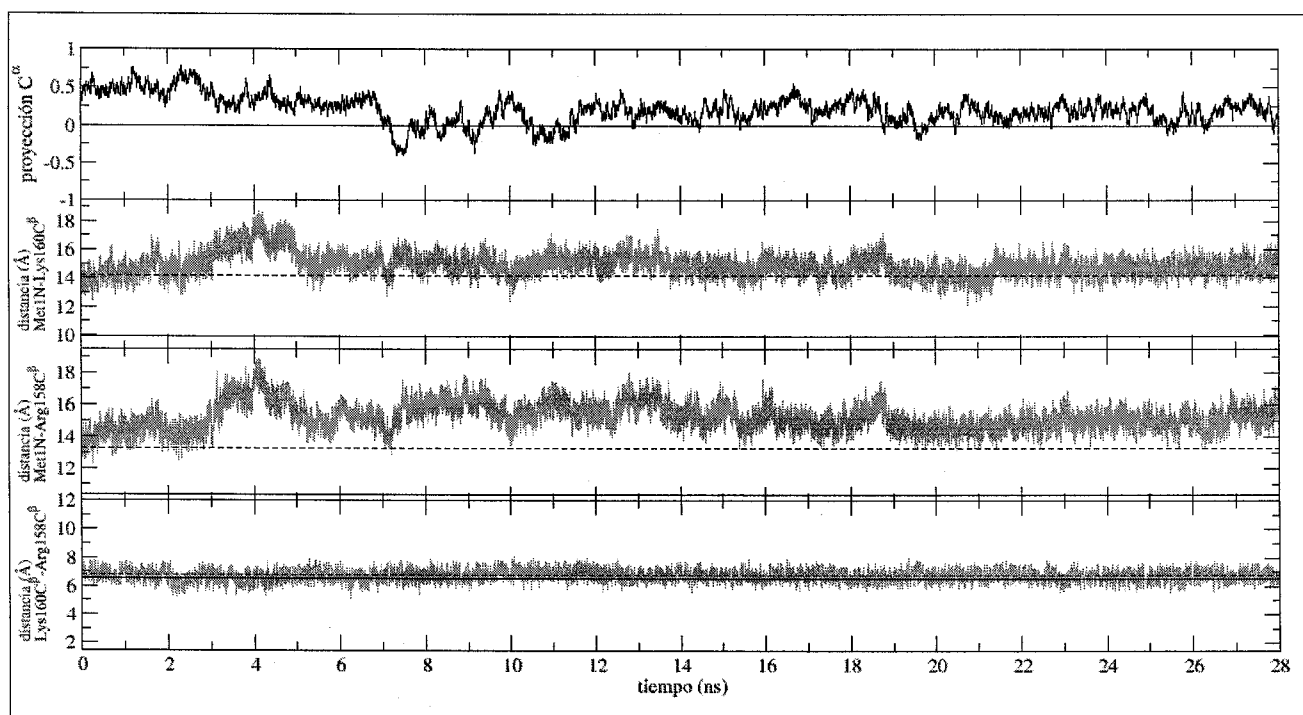


Figura 30. Formación del sitio alostérico a lo largo de Sol28. El panel superior muestra en negro la gráfica de la proyección del protómero (cfr. figura 22). En los otros tres paneles las gráficas en gris corresponden a las tres distancias en base a las cuales se define la formación del sitio alostérico, las líneas negras continuas indican el valor de referencia en R y, las líneas negras segmentadas, el valor de referencia en T.

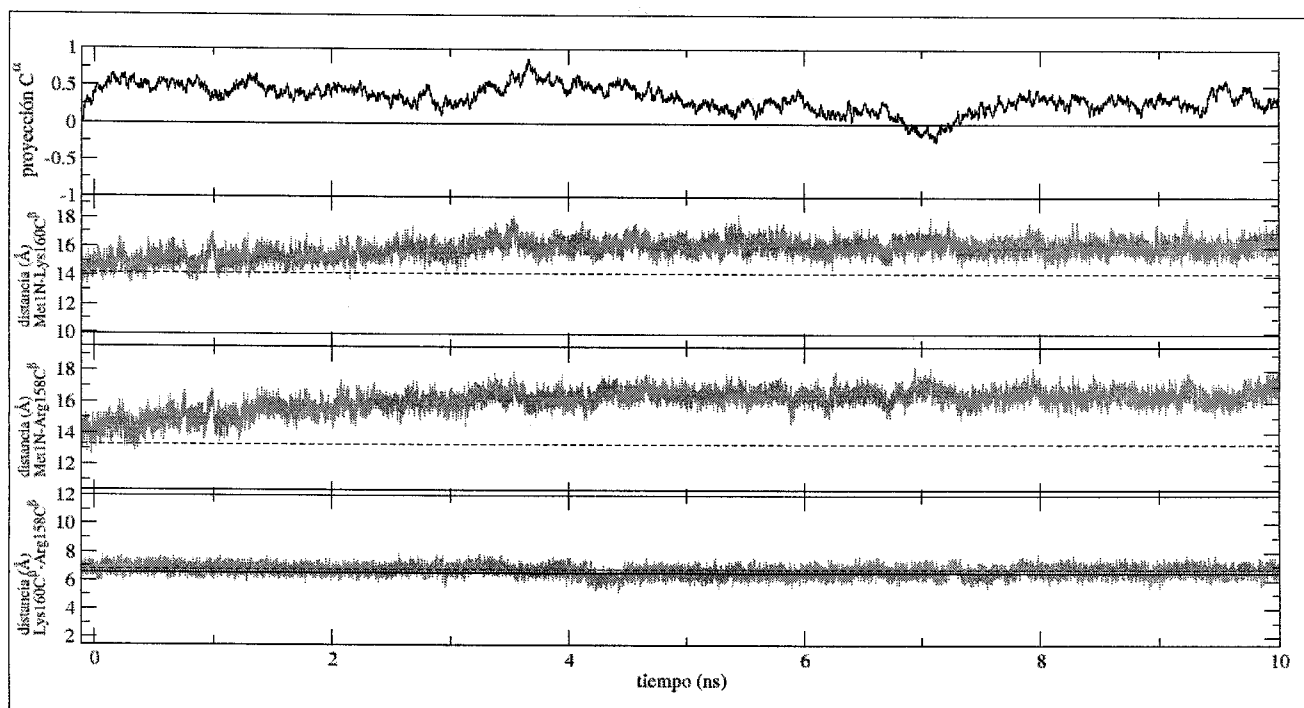


Figura 31. Formación del sitio alostérico a lo largo de Sol10. El panel superior muestra en negro la gráfica de la proyección del protómero (cfr. figura 24). En los otros tres paneles las gráficas en gris corresponden a las tres distancias en base a las cuales se define la formación del sitio alostérico, las líneas negras continuas indican el valor de referencia en R y, las líneas negras segmentadas, el valor de referencia en T.

En el caso del sitio alostérico, este nunca llega a formarse, y las distancias entre los átomos que lo definen permanecen demasiado alejadas de su valor de referencia en R. Esto es consistente con el hecho de que el sitio alostérico tiene una carga neta positiva en su superficie de manera que, por repulsión electrostática, las paredes de esta cavidad permanecen separadas hasta que algún ligante (de carga negativa) la neutralice [5].

### **8.6. Estabilización de la tapa abierta mediante una molécula de agua**

Dado que en Sol28 se muestrea satisfactoriamente el espacio conformacional de la proteína, pudieron observarse las conformaciones abierta y cerrada de la tapa del sitio activo tanto en el estado T como en el R (tabla III). En este análisis visual se encuentra consistentemente una molécula de agua equidistante al oxígeno O de la cadena principal de Thr161 (en la tapa) y el oxígeno OH del grupo fenol de Tyr254 (en el carboxilo terminal) cuando la tapa está abierta, ya sea en la conformación T o cercana a R. De hecho, esta molécula de agua presenta poca movilidad y está orientada electrostáticamente, con la carga negativa de su oxígeno alejándose de los oxígenos mencionados de la proteína y uno de los hidrógenos acercándose a ellos, y con la molécula de agua orientada de forma que

su oxígeno es repelido por el átomo OH de Tyr254, que es ligeramente más negativo que el átomo O de Thr161 (figura 32).

Tabla III. Intervalos de la simulación Sol28 donde se observan los estados abierto o cerrado de la tapa del sitio activo, ya sea cerca de la conformación T o R.

tapa abierta cerca de R	tapa abierta cerca de T	tapa cerrada cerca de R	tapa cerrada cerca de T
12.0246ns a 12.0306ns	11.41ns a 11.43ns	0.5658ns a 0.5806ns	4.7488ns a 4.7782ns
16.7462ns a 16.7564ns	18.73ns a 18.75ns	1.7830ns a 1.7940ns	
18.5760ns a 18.5864ns	20.58ns a 20.60ns	2.6144ns a 2.6418ns	
20.7302ns a 20.7364ns	22.635ns a 22.655ns	2.9736ns a 2.9844ns	
25.8120ns a 25.8190ns	26.01ns a 26.03ns		



Figura 32. Representación en rizoma de los residuos Thr161 (abajo) y Tyr254 (arriba) y moléculas de agua cercanas, según se observa en la simulación Sol28. El panel izquierdo corresponde a una conformación cercana a R (18.5760 ns a 18.5864 ns), mientras que el derecho muestra la conformación T (26.01 ns a 26.03 ns), ambas con la tapa abierta. La molécula de agua de interés se representa en rojo y blanco. Las moléculas de agua cercanas se representan en un solo color cada una. Imagen generada en VMD [6].

Sin embargo, en los intervalos en los que la tapa del sitio activo está cerrada la molécula de agua equidistante a los átomos O de Thr161 y OH de Tyr254 presenta mucha movilidad, cuando llega a estar presente (figura 33).

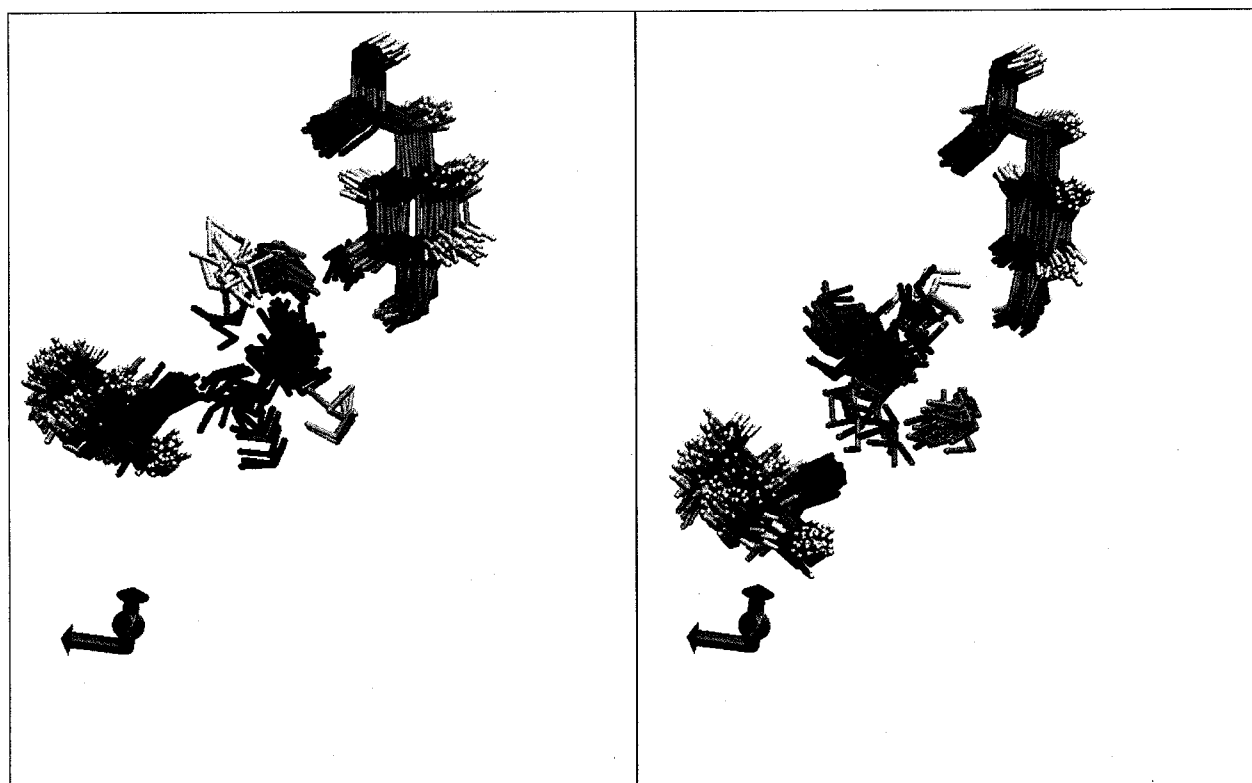


Figura 33. Representación en rizoma de los residuos Thr161 (abajo) y Tyr254 (arriba) y moléculas de agua cercanas, según se observa en la simulación Sol28. El panel izquierdo corresponde a una conformación cercana a R (2.6144ns a 2.6418ns), mientras que el derecho muestra la conformación T (4.7488ns a 4.7782ns), ambas con la tapa cerrada. Las moléculas de agua cercanas se representan en un solo color cada una, para apreciar el grado de movilidad que presentan. Imagen generada en VMD [6].

Estas observaciones sugieren que, una vez que el movimiento de la tapa la abre, la Tyr254 puede llegar a estabilizar este estado a través de alguna molécula de agua cercana.

### **8.7. Transmisión de los efectos de la tapa al resto de la proteína**

El movimiento de apertura y cerrado de la tapa del sitio activo es transmitido al resto de la proteína a través de la Arg156, que está cerca de su base (la tapa abarca los residuos 161 a 184). La importancia de este residuo para la regulación alostérica de la enzima puede inferirse del hecho de que está altamente conservado en su rama del árbol filogenético del gen *nagB*, la cual engloba los ortólogos que sí forman oligómeros (de hecho, todos forman hexámeros) y cuyo efector alostérico es GlcNAc6P, que se encuentran principalmente en proteobacterias pero también en humano (gen *gnp1*) [45].

Se observó que Arg156 interacciona con dos residuos de la hélice  $\alpha$  120 a 128: Glu122 y Arg126. En el estado abierto de la tapa, Arg156 presenta un comportamiento similar independientemente de la conformación alostérica de la proteína, colocándose equidistante a las cadenas laterales de Glu122 y Arg126, atrayendo al primero y repeliendo a la segunda. De esta manera, la hélice 120 a 128 se desplaza alejándose de la tapa, y posiblemente propagando este efecto a otras regiones de la proteína (figura 34).

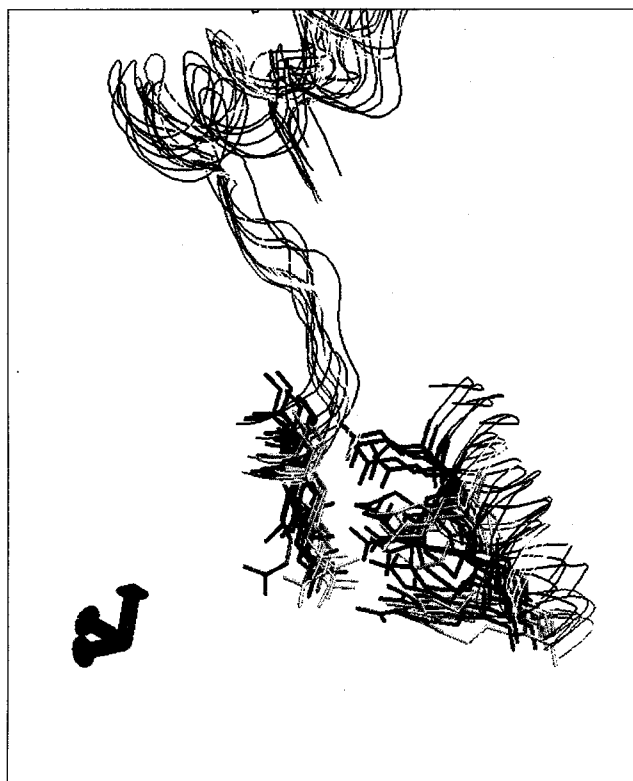


Figura 34. Interacción de Arg156 cerca de la base de la tapa (centro de la figura) con Arg126 (esquina inferior derecha) y Glu122 (arriba de Arg126). Estos residuos se muestran en representación de rizo. Tanto la hélice 120 a 128 como la tapa del sitio activo (residuos 161 a 184) y el segmento que la conecta con Arg156 se muestran en representación de tubo. En gris se muestran las conformaciones cercanas a R y en amarillo las cercanas a T, ambas con la tapa abierta. Las estructuras con la tapa cerrada se muestran en rojo (las cercanas a R), y en azul (las cercanas a T). Cada estructura es el promedio de uno de los intervalos de la tabla III, minimizada posteriormente con 500 pasos del algoritmo DMP limitando armónicamente el movimiento de los  $C^{\alpha}$  con una constante de fuerza de  $24 \text{ kcal/mol/\text{Å}^2}$ . Imagen generada en VMD [6].

Los resultados del análisis de modos cuasiarmónicos sugieren que este efecto puede estar relacionado con la transición alostérica. Esta está presente en la oscilación más representativa de la proteína, pues se observa poca movilidad del módulo interno, y el módulo externo se mueve alrededor de él (figura 35). Sin embargo, en este modo cuasiarmónico se observa también un movimiento de la tapa del sitio activo en una dirección distinta (figura 36), de manera que existe una correlación entre

estas dos oscilaciones. Esto último sólo se observa, sin embargo, en la simulación Sol28, que es la que explora más ampliamente el espacio conformacional de la enzima.

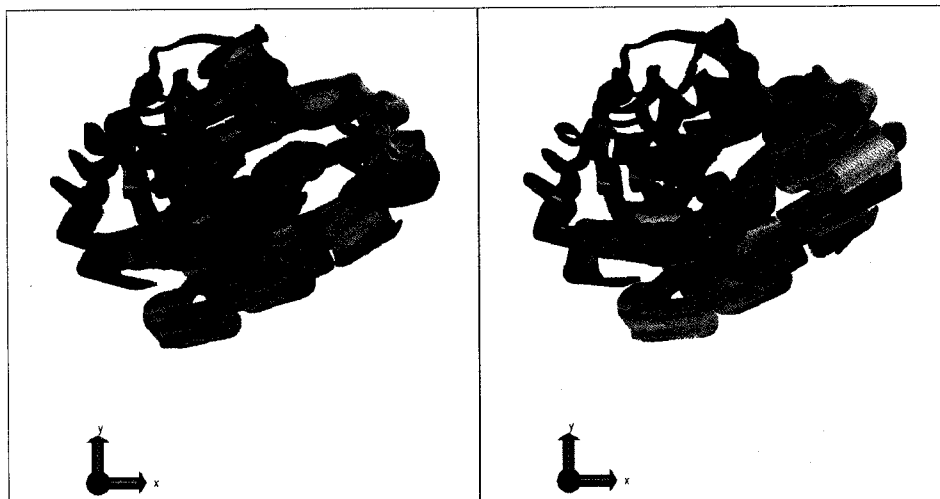


Figura 35. Superposición de las conformaciones de la traza de la proteína a lo largo del modo cuasiarmónico más representativo de su movimiento en las simulaciones Sol28 (izquierda) y Sol10 (derecha), vistas a lo largo del eje de orden tres. El color de cada  $C^\alpha$  indica su movilidad (azul para los menos móviles y rojo para los más móviles, cfr. figuras 6 y 8). Imagen generada en VMD.

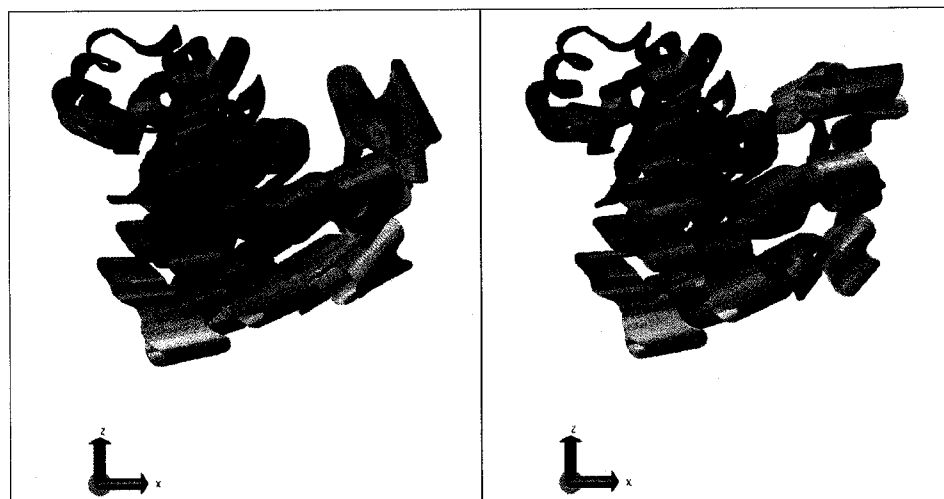


Figura 36. Superposición de las conformaciones de la traza de la proteína a lo largo del modo cuasiarmónico más representativo de su movimiento en las simulaciones Sol28 (izquierda) y Sol10 (derecha), vistas a lo largo del eje de orden dos. El color de cada  $C^\alpha$  indica su movilidad (azul para los menos móviles y rojo para los más móviles, cfr. figura 6). Imagen generada en VMD.

Lo anterior no engloba la dinámica de la tapa del sitio activo, ya que existen también otros modos cuasiarmónicos (de mayor amplitud y, por lo tanto, menor frecuencia) que involucran movimientos de la tapa (por ejemplo, el de la figura 37).



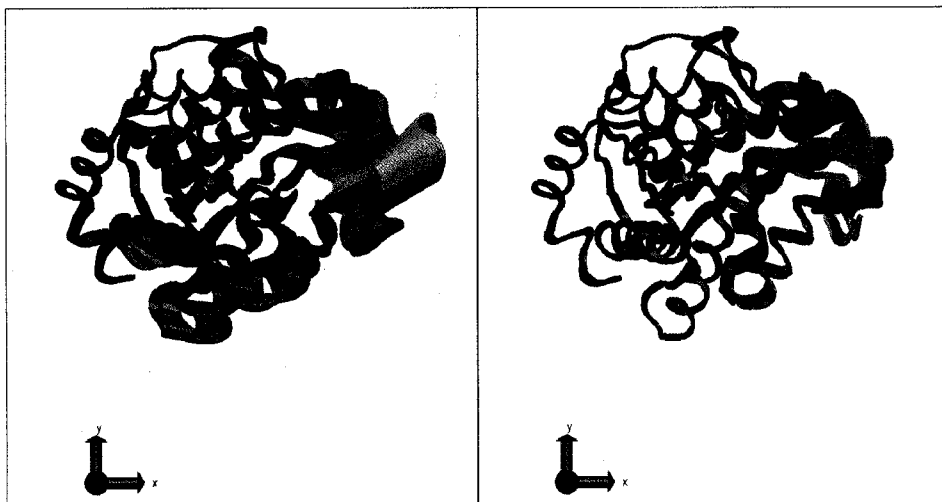


Figura 37. Superposición de las conformaciones de la traza de la proteína a lo largo del segundo modo cuasiarmónico más representativo de su movimiento en las simulaciones Sol28 (izquierda) y Sol10 (derecha), vistas a lo largo del eje de orden tres. El color de cada  $C^\alpha$  indica su movilidad (azul para los menos móviles y rojo para los más móviles, cfr. figura 8). Imagen generada en VMD.

Esto sugiere que la tapa del sitio activo transmite a través de Arg156 los efectos de sólo algunos de sus movimientos. De acuerdo con la descomposición de Fourier sobre el análisis de proyección de la tapa, estos movimientos corresponden a los de menor frecuencia, de manera que las oscilaciones más rápidas no son transmitidas al resto de la proteína.

## 9. CONCLUSIONES

La estructura de la GlcN6P desaminasa de *E. coli* oscila a lo largo de su transición alostérica aun en ausencia de ligantes, lo cual apoya la hipótesis de Rudiño y col. y es consistente con las simulaciones y los experimentos que evidencian un equilibrio conformacional preexistente. Sin embargo, en el cristal las conformaciones cercanas a R se ven desfavorecidas debido a choques estéricos interhexaméricos, por lo que estas deben tener una ocupación baja.

El movimiento más representativo de la dinámica de la proteína se da en la dirección de la transición alostérica. En general, los movimientos de esta enzima son globales y, descartando el caso de la tapa del sitio activo, no son específicos de la zona de la proteína. Además, tienen períodos en la escala de tiempo de ns, o bien, frecuencias del orden de cientos de MHz. Algunas oscilaciones de la tapa del sitio activo son, sin embargo, un orden de magnitud más rápidas, lo cual sugiere que estos movimientos son distintos a los globales. Sin embargo, existe una transferencia mediada por Arg156 de los efectos de los movimientos más lentos de la tapa al resto de la proteína.

La cavidad del sitio alostérico no puede formarse en ausencia de ligantes que neutralicen las cargas positivas de sus paredes. El sitio activo, por otra parte, requiere de la presencia de algún ligante para permanecer en su conformación activa una vez formado. Además, una vez que la tapa del sitio activo se ha abierto, Tyr254 puede estabilizarla en este estado por medio de interacciones electrostáticas con una molécula de agua.

## 10. REFERENCIAS

- 1 Plumbridge J. A., Cochet O., Souza J. M., Altamirano M. M., Calcagno M. L. y Badet B. 1993. *Coordinated regulation of amino sugar-synthesizing and -degrading enzymes in Escherichia coli K-12*. Journal of Bacteriology 175(16):4951 a 4956.
- 2 Bairoch A. y Boeckman B. 1993. *The SWISS-PROT protein sequence data bank, recent developments*. Nucleic Acids Research 21:3093 a 3096.
- 3 Oliva G., Fontes M. R. M., Garratt R. C., Altamirano M. M., Calcagno M. L. y Horjales E. 1995. *Structure and catalytic mechanism of glucosamine 6-phosphate deaminase from Escherichia coli at 2.1 Å resolution*. Structure 3(12):1323 a 1332.
- 4 Rudiño-Piñera E., Morales-Arrieta S., Rojas-Trejo S. P. y Horjales E. 2002. *Structural flexibility, an essential component of the allosteric activation in Escherichia coli glucosamine-6-phosphate deaminase*. Acta Crystallographica Section D 58:10 a 20.
- 5 Horjales E., Altamirano M. M., Calcagno M. L., Garratt R. C. y Oliva G. 1999. *The allosteric transition of glucosamine-6-phosphate deaminase: the structure of the T state at 2.3 Å resolution*. Structure 7(5):527 a 537.
- 6 Humphrey W., Dalke A. y Schulten K. 1996. *VMD - Visual Molecular Dynamics*. Journal of Molecular Graphics 14:33 a 38.
- 7 Koshland D. E., Némethy G. y Filmer D. 1966. *Comparison of Experimental Binding Data and Theoretical Models in Proteins Containing Subunits*. Biochemistry 65(1):365 a 385.
- 8 Calcagno M., Campos P. J., Mulliert G. y Suástegui J. 1984. *Purification, molecular and kinetic properties of glucosamine-6-phosphate isomerase (deaminase) from Escherichia coli*. Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology 787(2):165 a 173.
- 9 Monod J., Wyman J. y Changeux J. P. 1965. *On the nature of allosteric transitions: a plausible model*. Journal of Molecular Biology 12(1):88 a 118.
- 10 Gabriela M. Montero-Morán, Samuel Lara-González, Laura I. Álvarez-Añorve, Jacqueline A. Plumbridge y Mario L. Calcagno. 2001. *On the Multiple Functional Roles of the Active Site Histidine in Catalysis and Allosteric Regulation of Escherichia coli Glucosamine 6-Phosphate Deaminase*. Biochemistry 40(34):10187 a 10196.
- 11 Berman H. M., Westbrook J., Feng Z., Gilliland G., Bhat T. N., Weissig H., Shindyalov I. N. y Bourne P. E. 2000. *The Protein Data Bank*. Nucleic Acids Research 28:235 a 242.

- 12 Kern Dorothee y Zuiderweg Erik R. P. 2003. *The role of dynamics in allosteric regulation*. Current Opinion in Structural Biology 13(6):748 a 757.
- 13 Frauenfelder H., Parak F. y Young R. 1988. *Conformational substates in proteins*. Annual review of biophysics and biophysical chemistry 17:451 a 479.
- 14 Jianpeng Ma y Martin Karplus. 1998. *The allosteric mechanism of the chaperonin GroEL: A dynamic analysis*. Proceedings of the National Academy of Sciences 95(15):8502 a 8507.
- 15 Jianpeng Ma, Paul B. Sigler, Zhaohui Xu y Martin Karplus. 2000. *A dynamic model for the allosteric mechanism of GroEL*. Journal of Molecular Biology 302(2):303 a 313.
- 16 Aline Thomas, Konrad Hinsen, Martin J. Field y David Perahia. 1999. *Tertiary and quaternary conformational changes in aspartate transcarbamylase: a normal mode study*. Proteins: Structure, Function, and Genetics 34(1):96 a 112.
- 17 Scott W. Nelson, Cristina V. Iancu, Jun-Yong Choe, Richard B. Honzatko y Herbert J. Fromm. 2000. *Tryptophan Fluorescence Reveals the Conformational State of a Dynamic Loop in Recombinant Porcine Fructose-1,6-bisphosphatase*. Biochemistry 39(36):11100 a 11106.
- 18 Brian F. Volkman, Doron Lipson, David E. Wemmer y Dorothee Kern. 2001. 2429 - 2433. Science 23(291):2429 a 2433.
- 19 Thomas J. Smith, Timothy Schmidt, Jie Fang, Jane Wu, Gary Siuzdak y Charles A. Stanley. 2002. *The structure of apo human glutamate dehydrogenase details subunit communication and allostery*. Journal of molecular biology 318(3):765 a 777.
- 20 Elena Bochkareva, Visar Belegu, Sergey Korolev y Alexey Bochkarev. 2001. *Structure of the major single-stranded DNA-binding domain of replication protein A suggests a dynamic mechanism for DNA binding*. European Molecular Biology Organization 20(3):612 a 618.
- 21 Yasuyuki Matoba y Masanori Sugiyama. 2003. *Atomic resolution structure of prokaryotic phospholipase A[2]: Analysis of internal motion and implication for a catalytic mechanism*. Proteins 51(3):453 a 469.
- 22 Liliane Mouawad, David Perahia, Charles H. Robert y Christophe Guilbert. 2002. *New Insights into the Allosteric Mechanism of Human Hemoglobin from Molecular Dynamics Simulations*. Biophysical Journal 82(6):3224 a 3245.
- 23 Jonathan A. Lukin, Georg Kontaxis, Virgil Simplaceanu, Yue Yuan, Ad Bax y Chien Ho. 2003. *Quaternary structure of hemoglobin in solution*. Proceedings of the National Academy of Sciences 100 (2):517 a 520.

- 24 Hashim M. Al-Hashimi, Yuying Gosser, Andrey Gorin, Weidong Hu, Ananya Majumdar, Dinshaw J. Patel. 2002. *Concerted motions in HIV-1 TAR RNA may allow access to bound state conformations: RNA dynamics from NMR residual dipolar couplings*. Journal of molecular biology 315(2):95 a 102.
- 25 Lena Måler, John Blankenship, Mark Rance y Walter J. Chazin. 2000. *Site-site communication in the EF-hand Ca<sup>2+</sup>-binding protein calbindin D9k*. Nature Structural Biology 7(3):245 a 250.
- 26 Andrew L. Lee y A. Joshua Wand. 2001. *Microscopic origins of entropy, heat capacity and the glass transition in proteins*. Nature 411(6836):501 a 504.
- 27 Craig McElroy, Amanda Manfreda, Alice Wendt, Paul Gollnick y Mark Foster. 2002. *TROSY-NMR Studies of the 91 kDa TRAP Protein Reveal Allosteric Control of a Gene Regulatory Protein by Ligand-altered Flexibility*. Journal of Molecular Biology 323(3):463 a 473.
- 28 Sutjano Jusuf, Patrick J. Loll y Paul H. Axelsen. 2003. *Configurational entropy and cooperativity between ligand binding and dimerization in glycopeptide antibiotics*. Journal of the American Chemical Society 125(13):3988 a 3994.
- 29 Fletcher R. y Reeves C. M. 1964. *Function minimization by conjugate gradients*. The Computer Journal 7(2):149 a 154.
- 30 Ewald, P. P. 1921. *The calculation of optical and electrostatic lattice potentials*. Annals of Physics (Leipzig) 64:253 a 287.
- 31 Essmann U., Perera L., Berkowitz M. L., Darden T., Lee H. y Pedersen L. G. 1995. *A smooth particle mesh Ewald method*. Journal of Chemical Physics 103(19):8577 a 8593.
- 32 Hünenberger P. H. y McCammon J. A. 1999. *Effect of artificial periodicity in simulations of biomolecules under Ewald boundary conditions: a continuum electrostatics study*. Biophysical Chemistry 78(1):69 a 88.
- 33 Norberg J. y Nilsson L. 2000. *On the Truncation of Long-Range Electrostatic Interactions in DNA*. Biophysical Journal 79(3):1537 a 1553.
- 34 Berendsen H. J. C., Postma J. P. M., van Gunsteren W. F., DiNola A. y Haak J. R. 1984. *Molecular dynamics with coupling to an external bath*. The Journal of Chemical Physics 81(8):3684 a 3690.
- 35 Andersen H. C. 1980. *Molecular dynamics simulations at constant pressure and/or temperature*. The Journal of Chemical Physics 72(4):2384 a 2393.
- 36 Nosé S. y Klein M. L. 1983. *Constant pressure molecular dynamics for molecular systems*. Molecular Physics 50(5):1055 a 1076.
- 37 Feller S. E., Zhang Y., Pastor R. W. y Brooks B. R. 1995. *Constant pressure molecular dynamics simulation: The Langevin piston method*. Journal of Chemical Physics 103(11):4613 a 4621.

- 38 Nosé S. 1984. *A unified formulation of the constant temperature molecular dynamics methods*. The Journal of Chemical Physics 81(1):511 a 519.
- 39 Hoover W. G. 1985. *Canonical dynamics: Equilibrium phase-space distributions*. Physical Review A 31:1695 a 1697.
- 40 Jean-Ryckaert Paul, Ciccotti Giovanni y Berendsen Herman J. C. 1977. *Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes*. Journal of Computational Physics 23(3):327 a 341.
- 41 Brünger A. T. y Karplus M. 1988. *Polar hydrogen positions in proteins: empirical energy placement and neutron diffraction comparison*. Proteins: Structure, Function, and Genetics 4(2):148 a 156.
- 42 Brooks B. R., Bruccoleri R. E., Olafson B. D., States D. J., Swaminathan S. y Karplus M. 1983. *CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations*. Journal of Computational Chemistry 4:187 a 217.
- 43 MacKerell Jr. A. D., Brooks B., Brooks III C. L., Nilsson L., Roux B., Won Y. y Karplus M. 1998. *CHARMM: The Energy Function and Its Parameterization with an Overview of the Program*. The Encyclopedia of Computational Chemistry 1:271 a 277.
- 44 Wolfram Research, Inc. 2004. *Mathematica*, Versión 5.1. Wolfram Research, Inc., Champaign, Illinois.
- 45 Florence Vincent, Gideon J. Davies y James A. Brannigan. 2005. *Structure and kinetics of a monomeric glucosamine 6-phosphate deaminase: missing link of the nagb superfamily?*. Journal of Biological Chemistry 280(20):19649 a 19655.

## 11. BIBLIOGRAFÍA

- CREIGHTON, Thomas E. 1993. *Proteins: Structures and Molecular Properties*. 2ª edición. W. H. Freeman and Company, New York.
- HAILE, James M. 1992. *Molecular dynamics simulation: elementary methods*. 1ª edición. John Wiley & Sons, Inc., New York.
- [President and Fellows of Harvard College]. 1995. *CHARMM Documentation*. <http://www.charmm.org/document/Charmm/c28b2/index.html>.