



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE CIENCIAS

OSCILACIONES GRUPALES ASOCIADAS A CAMBIOS CONFORMACIONALES
EN CRISTALES DE PROTEÍNA

TESIS QUE PARA OBTENER EL TÍTULO DE BIÓLOGO PRESENTA:

EUGENIO DE LA MORA LUGO



DIRECTOR DE TESIS: DR. ENRIQUE RUDIÑO PIÑERA

2005



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

Agradecimientos

A Eduardo por darme la oportunidad de aprender junto con él y todo el grupo. Por sus ideas, asesoría en los comienzos y durante todo el proceso de este trabajo. Por el aprendizaje extracurricular.

A Enrique por compartir su conocimiento y su asesoría. Por su paciencia, disposición, apoyo, buen humor y por la extensa revisión de este escrito.

A los Drs. Adela Rodríguez, Alejandro Fernández y Alejandro Sosa, por sus valiosas aportaciones tras la lectura de esta tesis.

A los Drs. Patricia Joseph y José Luis Puente coordinadores del taller 'La biología a partir de las biomoléculas, nuevos paradigmas y fronteras'. Por su tiempo dedicado a la formación de científicos y personas. Por la oportunidad de formar parte de su proyecto.

A todos los integrantes del Laboratorio de Cristalografía de Proteínas: Sonia, Rosario, Paula, Aisel, Álvaro, Eleuterio, Jonathan, Leopoldo, Mauricio y Yagul.

A Lorenzo por los seminarios, la música y el buen humor de siempre.

Al Instituto de Biotecnología por el apoyo económico para la estancia en Cuernavaca, Morelos, pero sobre todo por integrarnos a su espacio académico.

A la Facultad de Ciencias, espacio de aprendizaje, de diversidad de ideas, de buenos amigos... ('territorio libre y rebelde').

A los profesores de la Facultad: Hortensia González, Humberto Arce, Alicia Negrón, Saúl Cano, Judith Márquez y todos los que me faltó mencionar, por su dedicación y aprecio por la academia.

A mis amigos y compañeros, Miriam, Sandra, Etienne y José Arcadio, con quienes compartí un año lleno de tan buenas experiencias. A Adriana, Elena, Emilio, Martí, Nacxi, Sandra, Marcos, Eduardo y Miguel.

A Isabel. A Alejandro.
A Camilo. A Juliana. A Paula.
A Tata. A Raquel.
A Lua.

'Nuestra atención es la que pone los objetos en un cuarto;
el hábito el que los quita y nos hace sitio.' Marcel Proust.

ÍNDICE

Página

1 Introducción	6
2 Marco teórico	8
Principios de cristalografía	8
Historia	9
Difracción de rayos X	12
Simetría	14
Patrón de difracción y factores de estructura	14
Valores de temperatura isotrópicos y anisotrópicos	17
Criterios de validación	19
Afinamiento por tensores TLS	21
Interpretación de los resultados	25
Relación datos/parámetros	26
Cristales, flexibilidad estructural y oscilaciones	27
Actividad enzimática y dinámica molecular	27
Cambio de conformación local y estados conformacionales	29
3 Objetivos	30
Objetivos generales	30
Objetivos particulares	30
4 Metodología	30
Búsqueda en el PDB	30
Selección de estructuras	31
Sobreposición de estructuras	31
Análisis de valores B	32
Selección de grupos	32
Unificación de criterios	32
Afinamiento TLS	33
Análisis de valores de temperatura anisotrópicos	33
Programas	34
5 Resultados y discusión	35
Búsqueda en el PDB	35
Selección de estructuras	35
O-acetilserina sulfidrilasa	36
Estructura	37
Estructuras de OAS depositadas en el PDB	38
Distribución de valores B	39
Sobreposición de estructuras	41
Cambio conformacional	42
Selección de grupos pseudo-rígidos para el afinamiento TLS	43
Afinamiento por el método de tensores TLS	44
Transportador SecA	49
Estructura	49
Estructuras de SecA depositadas en el PDB	51
Distribución de los valores B	51
Sobreposición de estructuras	52
Cambio conformacional	53
	54

Selección de grupos de grupos pseudo-rígidos para el afinamiento TLS.....	55
Afinamiento TLS.....	58
UDP-N-acetilglucosamina enolpiruvato transferasa (EPT).....	59
Estructura.....	60
Estructuras de EPT depositadas en el PDB.....	61
Distribución de valores B	62
Sobreposición de estructuras.....	63
Cambio conformacional.....	64
Selección de grupos de grupos pseudo-rígidos para el afinamiento TLS.....	65
Afinamiento por tensores TLS.....	69
A manera de conclusión.....	
6 Conclusiones_____	70
7 Perspectivas_____	71
8 Referencias_____	73
9 Índice de figuras_____	77

1 INTRODUCCIÓN

El grial de quienes estudian proteínas es la conexión entre estructura y función¹. Un error común durante el análisis de estructuras es el interpretar los modelos (obtenidos ya sea por cristalografía de rayos X o resonancia magnética nuclear, *NMR*) como estructuras estáticas en las que lo más relevante es localizar a los residuos catalíticos y a los involucrados en la unión de ligandos. Sabemos² que la función de las proteínas depende de procesos dinámicos², como lo son la unión de sustratos y cofactores, liberación del producto y regulación alostérica³. Por eso es cada vez más necesario incluir información dinámica cuando se reporta la estructura de una molécula e intenta, a partir de ella, inferirse su función biológica.

En difracción de rayos X, los valores de estructura, datos primarios de la difracción, contienen la información de cada uno de los puntos que resultan de irradiar con rayos X a un cristal que contiene unas 10^{15} moléculas de proteína⁴. Tras procesar esa información se obtiene una estructura que en realidad es el promedio, en el tiempo, de una población de moléculas, dentro del cristal*, con conformaciones variables tanto en espacio (de celda unitaria a celda unitaria, ver **Principios de Cristalografía**) como en tiempo (producto de las oscilaciones atómicas). *A priori* uno pensaría que las moléculas al encontrarse en estado cristalino pierden su oscilación, pero no es así, incluso, podemos asegurar que éstas se mantienen por al menos tres evidencias:

1. Los cristales de proteína se parecen más a soluciones con alta concentración de solutos⁸ que a un estado sólido, pues contienen entre 30-80% de solvente⁵,
2. La existencia de cristales catalíticos⁶. En especial si asumimos que la dinámica es esencial para la catálisis.
3. Los resultados de difracciones a diferentes temperaturas⁷, que muestran reducción del desplazamiento atómico a bajas temperaturas.

Por ello suponemos que, aún en estado cristalino, las proteínas mantienen el patrón oscilatorio característico de su 'estado celular', eso sí, con desplazamientos de menor magnitud y restringidos en las regiones de contactos cristalinos; en ambos casos causados por la alta densidad del entorno. A pesar de las evidencias anteriores, pocos son los trabajos en que se discute la posibilidad de extraer la información oscilatoria a partir de los datos de difracción de rayos X.

Nuestro grupo ha planteado la posibilidad de obtener información oscilatoria relacionada con la actividad biológica de las moléculas a partir de los datos experimentales generados por difracción de rayos X^{8, 9}. Consideramos que esa información ha sido ignorada; explicada como errores o ruido,

* Tanto por la teoría físicoquímica de las interacciones atómicas para moléculas del tamaño de las proteínas (particularmente en solvente como el agua) que predice movimientos u oscilaciones, como por datos experimentales.

⁸ Entendemos un cristal como un conjunto de átomos ordenados con arreglo periódico en tres dimensiones.

⁹ Asemejándose más al estado de gel. Por la presencia de tantos solutos en la célula podemos decir que el estado en que se encuentran las moléculas de proteína no es tan distinto de su 'estado celular'.

consecuencia de la alta cantidad de moléculas y de sub-estados conformacionales al interior del cristal. Durante el procesamiento de los datos experimentales parte de esos errores es absorbida por los llamados valores de temperatura (**Valores de temperatura isotrópicos y anisotrópicos**). Al graficar dichos valores contra el número de residuo es posible identificar las zonas de mayor flexibilidad en la molécula; generalmente las más apartadas del núcleo estructural. Esto es un indicio de que es posible extraer al menos una parte de la información oscilatoria a partir de la información experimental.

Una forma de extraer la información oscilatoria es a través de un afinamiento anisotrópico del modelo atómico derivado de la difracción (**Difracción de rayos X**). Tras afinar anisotrópicamente cada átomo de la molécula y representarlos en el espacio tridimensional, se observan elipsoides (con ejes mayores en alguna de las direcciones) que permiten identificar grupos de átomos o de residuos de aminoácidos con desplazamientos comunes e independientes. El problema radica en que ese tipo de afinamiento es posible sólo a resoluciones mayores a 1.6 Å, puesto que a resoluciones menores la baja proporción datos/parámetros impide la obtención de este tipo de modelos. Para evitar ese aumento del error generalmente se realiza un afinamiento isotrópico que requiere únicamente cuatro parámetros (3 de posición y uno térmico) en lugar de los nueve que requiere el anisotrópico (tres de posición más 6 térmicos). La diferencia en el número de parámetros térmicos permite pasar de una representación esférica de probabilidad de posición atómica (isotrópica) a una de tipo elipsoidal (anisotrópica) en la que los ejes del movimiento atómico corresponden a los ejes de la elipsoide. Si bien es cierto que el acercamiento isotrópico es una buena aproximación implica una simplificación del problema que impide analizar la direccionalidad del movimiento. El problema se puede resolver por una aproximación de tipo cuasi-anisotrópica en la que en lugar de afinar átomos individuales, se afinan selecciones de grupos de átomos (del tamaño deseado) de manera tal que reducimos considerablemente el número de parámetros requeridos (los mismos cuatro que en el enfoque isotrópico más 6 parámetros por cada grupo seleccionado).

Este trabajo se enfoca en la relación entre patrones oscilatorios y cambios conformacionales causados por la unión de ligandos, inferida del análisis de los datos experimentales de difracción de rayos X (datos cristalográficos). Para demostrar esta relación realizamos un análisis de estructuras depositadas en el PDB¹⁰ (*Protein Data Bank*) resueltas por difracción de rayos X. Comenzamos con una búsqueda de depósitos por palabras clave relacionadas con cambios de conformación, seguida por una selección de calidad de las estructuras bajo criterios cristalográficos estrictos. Para cada molécula que pasó la criba buscamos las estructuras existentes en el PDB que escaparon de la búsqueda por palabra clave y tratamos de reconocer los cambios de conformación mediante la sobreposición de cada una de ellas, enfatizando en las formas libres o nativa (libres de ligandos) y con ligandos. A partir de la distribución de valores de temperatura por residuo de aminoácidos identificamos las regiones de mayor oscilación. Una vez ubicadas en la gráfica bidimensional

regresamos a la representación tridimensional, tratando de relacionar las zonas más móviles con posibles grupos pseudo-rígidos: grupos de residuos que oscilen de manera conjunta. Tras este análisis seleccionamos bajo distintos criterios grupos pseudo-rígidos que fueron afinados por el método de tensores TLS¹¹ (**Afinamiento por tensores TLS**) en el programa Refmac¹² buscando la combinación de ciclos de máxima similitud y tensores TLS que condujeran a la convergencia de los estadísticos de prueba R, R_{free} y figura de mérito (FOM por sus siglas en inglés). Representamos gráficamente los elipsoides derivados de los afinamientos de los pseudo-rígidos con mejores estadísticos de prueba. Sobre la representación tridimensional distinguimos los patrones generales de cada modelo y su posible relación con el cambio de conformación (por lo tanto con la función biológica).

Se presentan tres casos particulares: O-acetilserina sulfidrilasa, transportador SecA y UDP-N-acetilglucosamina enolpiruvil transferasa, cuyos resultados apoyan la tesis, de trabajos previos realizados en nuestro grupo, sobre la utilidad del afinamiento de tensores TLS para extraer información dinámica de estructuras cristalográficas. Esto siempre y cuando se trabaje con estructuras de buena calidad seleccionadas bajo criterios cristalográficos estrictos.

2 MARCO TEÓRICO

La biología estructural es la rama de la biología que busca comprender las bases moleculares de la actividad biológica. Dentro de ella existen varias técnicas que abordan el tema desde diferentes perspectivas. Entre las más socorridas están cristalografía de rayos X, resonancia magnética nuclear (NMR), microscopía electrónica, dicroísmo circular, fluorescencia y simulación de dinámicas moleculares. Este trabajo se enfoca en la cristalografía de proteínas, sobre la que a continuación se presenta una visión general.

PRINCIPIOS DE CRISTALOGRAFÍA

Cristalografía significa descripción de los cristales (del griego *krystallos*, agua solidificada por enfriamiento y *grafos*, descripción); es decir, analizar todo lo concerniente a la forma externa, estructura interna y propiedades físicas, químicas y en algunos casos biológicas de los cristales. En un principio comenzó como una rama descriptiva de la mineralogía que se convirtió en una ciencia aparte a mediados del siglo XIX, cuando el estado cristalino, extendido a moléculas distintas de los minerales, se consideró un estado común de la materia. En las últimas cuatro décadas la cristalografía ha avanzado de una manera tan veloz que ahora es relativamente simple resolver una estructura de proteína (si comparamos con los más de 15 años que tomó a Kendrew y Perutz resolver la estructura de la mioglobina y hemoglobina respectivamente), en gran parte gracias al desarrollo de herramientas computacionales, fuentes de radiación y métodos de colecta de datos.

HISTORIA

Quizá el antecedente más antiguo sobre cristalografía se remonte al origen de la palabra en la cultura helénica. Después de ello hay un vacío histórico que conduce al siglo XVII en el que Kepler, Hooke y Huyghens realizaron independientemente observaciones sobre las propiedades ópticas de los cristales y de la relación entre apariencia externa y ordenamiento interno de la materia¹³. En el siglo XVIII sobresale el establecimiento de las primeras dos leyes de la cristalografía; la de los ángulos constantes por N. Steensen y la de los índices racionales por A. R. Haüy. Ya en el XIX W. H. Miller afina las ideas de Haüy e incorpora a la cristalografía los métodos de la geometría analítica. En 1849 A. Bravais postuló en *La estructura reticular de los cristales* el enunciado que se ha convertido en la base de la cristalografía:

‘Dado un punto p en un cristal existen puntos discretos infinitos, ilimitados en las tres direcciones del espacio, alrededor de los cuales, el arreglo de la materia es igual que alrededor del punto p ’.

Surge entonces la noción de red cristalina tridimensional y las simetrías involucradas enumeradas y clasificadas por Schönflies y Fedorov¹⁴.

Pero las bases de la cristalografía de rayos X comenzaron a establecerse a finales del siglo XIX tras su descubrimiento y posterior descripción (ambas realizadas por Röntgen) de algunas de sus propiedades generales: producción, propagación en línea recta, penetrabilidad de la materia e interacción con placas fotográficas. Faltaba aún demostrar su naturaleza precisa (de partículas o radiación electromagnética) y si eran capaces de interferir, refractar, difractar o reflejar en aparatos ópticos ordinarios. Aún sin conocer su naturaleza, una aproximación teórica considerando las ecuaciones de Lorentz-Maxwell, sugería la formación de ondas electromagnéticas como consecuencia directa del cambio en la velocidad de los electrones¹⁵. De ser así, sería posible reproducir experimentos ópticos comunes con rayos X en sustitución de luz ordinaria. Uno de los más usados consistía en pasar un haz de luz ordinaria por una red o gradilla de difracción; un pedazo de vidrio cubierto por una serie de líneas finas o delgadas, paralelas y equidistantes gravadas sobre su superficie, del que se obtiene un patrón de interferencia característico. Pero ello requería condiciones especiales, desconocidas en ese entonces, debido a la corta longitud de onda de los rayos X.

Varios años después, en 1912, A. J. Sommerfeld logró cuantificar una longitud de onda, difractada a través de una rendija o hendidura en forma de cuña de aproximadamente 0.1nm^{N} . Con estos datos

^N Ahora la mayoría de los difractómetros utilizan el cobre porque produce rayos de espectro de distribución con picos bien definidos a $1.54 \text{ \AA} = 0.154\text{nm}$, producto de una transición en la que el electrón de la capa K es excitado a un nivel superior y emite un rayo X de dicha longitud de onda conforme decae a la capa original.

y apoyado en un artículo publicado por Ewald (alumno de Sommerfeld) sobre la propagación de ondas electromagnéticas en un cristal, M. von Laue realizó el experimento clave que demostró convincentemente la naturaleza de los rayos X y de paso marcó el despegue de la técnica de difracción. La pregunta de von Laue fue sencilla: ¿qué pasaría si la longitud de onda (λ) de los rayos fuera menor a la distancia de repetición en el cristal?³ Tomando como referencia las medidas de Sommerfeld, observó que debería haber interferencia en la trayectoria de los rayos X cuando entraran en contacto con el cristal. En el mismo año de 1912 W. Friedrich y P. Knipping, junto con von Laue, obtuvieron la primera fotografía o patrón de difracción de rayos X producto de irradiar un cristal de sulfato de cobre penta hidratado ($\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$): la famosa fotografía *Beerstein* (taza de piedra, Figura 1)¹⁶. El espaciado regular entre los átomos del cristal causaba la misma interferencia que la producida por las estrechas líneas del enrejado sobre la luz. Ello confirmó la naturaleza electromagnética de los rayos X y permitió determinar su longitud de onda con mayor precisión.



Figura 1 Fotografía *Beerstein*. Difracción del cristal de sulfato de cobre penta hidratado, tomada por Knipping, Friedrich y von Laue.

El mismo año de 1912 W. H. Bragg y su hijo W. L. Bragg propusieron su hipótesis sobre el origen de los puntos en el patrón de difracción y establecieron lo que ahora conocemos como la ley de Bragg. Al año siguiente, los mismos Bragg, trabajaron las estructuras de NaCl, KCl, KBr y KI a partir de fotografías tomadas por von Laue. Es así como comienza a determinarse la estructura de moléculas por difracción de rayos X.

Paralelo al desarrollo de la cristalografía se da el de la enzimología. Por los años en que Röntgen trabajaba su teoría de difracción E. Fisher (1894) mostró que los catalizadores biológicos, a los que llamó enzimas, tenían una alta especificidad por su sustrato. Tras esa observación desarrolló el modelo de llave-cerradura en el que una enzima (cerradura) unía a su sustrato (llave) por complementariedad geométrica¹⁷. Los bioquímicos trataron de resolver el problema de forma y estructura por métodos que aproximaban el tamaño y forma general de las moléculas. Pero una mayor comprensión de la actividad enzimática no sería alcanzada sin 'fotografías' de las proteínas a resolución atómica; imposibles de lograr por métodos directos de imagen como microscopía electrónica o biofísicos indirectos como sedimentación.

³ La distancia típica entre los átomos de un sólido es de 0.1-0.5 nm, igual a $1-5 \times 10^{-8}$ cm, en ese entonces se medía a partir de la densidad del cristal y su peso molecular.

Los primeros intentos de análisis estructural de proteínas comienzan en la década de 1920, cuando ni siquiera se conocía la naturaleza proteica de las enzimas. Para ello fue indispensable el desarrollo de la teoría de cristalización y la obtención de los primeros cristales de proteína (ureasa), por J. Summer en 1926 (algunos dicen que la hemoglobina se cristalizó hace más de 150 años¹⁸). Las impurezas de sus preparaciones hicieron dudar a la comunidad científica sobre la composición química real y fue hasta 1936 cuando J. Nortrop y M. Kunitz demostraron la composición proteica de cristales de pepsina, tripsina y quimiotripsina a través de correlaciones directas entre cantidad de proteína y actividad enzimática cuando se aceptó formalmente la composición química como de bloques unidos de aminoácidos¹⁹. Así, se aclaraba su naturaleza y composición, además de su responsabilidad en el incremento selectivo de la tasa de reacción observado al aumentar la cantidad de enzima en los sistemas bioquímicos.

A pesar del rápido desarrollo, no es hasta la década de los 1930s que podemos considerar el inicio formal de la cristalografía de proteínas. Los primeros intentos de difractar cristales de proteína resultaron en patrones complejos producidos por la mala calidad de los mismos. Bernal y Hodgkin, en 1934, difractaron cristales en su líquido madre (la solución en la que crecieron) y descubrieron una mejora considerable de los patrones de difracción²⁰. En la década de los 40, M. Perutz, estudiante de Bragg, y J. Kendrew fueron de los primeros en entender la importancia de conocer estructuras proteicas y se embarcaron en la tarea de resolver la estructura de las proteínas mejor descritas en la actualidad; hemoglobina y mioglobina, respectivamente²¹. Por las complicaciones (sobre todo las de cuantificar las intensidades y resolver el problema de las fases manualmente) de extrapolar una técnica desarrollada para moléculas pequeñas el trabajo se postergó hasta 1957, cuando fueron publicadas las dos estructuras; primero la de mioglobina²² y tiempo después la de hemoglobina²³. La resolución de estructuras cristalinas acabó con la larga discusión de si las proteínas tenían una estructura estable o eran más bien geles. Tras las primeras publicaciones comenzó a estudiarse la dinámica molecular, en particular la de unión del oxígeno a ambas moléculas transportadoras.

Las mejoras en las técnicas cristalográficas alcanzaron, en la década de los 1970's, un punto en el que la necesidad de fuentes de rayos X más poderosas y de energía controlable se hizo cada vez mayor. Las fuentes se hicieron disponibles en la forma de sincrotrones (anillos de almacenamiento de electrones) de los que se pudo derivar rayos más intensos y con longitud de onda variable, aunque su aplicación a la cristalografía de proteínas comenzó hasta los 80's. Las líneas de los sincrotrones son las instalaciones en las que cristalógrafos y otros investigadores acceden a las fuentes de alta intensidad que necesitan para su investigación. Con el paso del tiempo ha aumentado el número de líneas dedicadas a experimentos de cristalografía hasta llegar alrededor de las 50: 20 en los 5 sincrotrones de Estados Unidos y cerca de treinta en países como Brasil, Alemania, Francia, Gran Bretaña, Japón, Tailandia y Corea del Sur²⁴. A la par del desarrollo de fuentes más poderosas de

energía fue necesario el de técnicas de protección para evitar el decaimiento de los cristales. Desde muchos años atrás los cristalógrafos habían sido conscientes de la conveniencia de coleccionar datos a temperaturas muy bajas, pues, teóricamente, la baja temperatura reduciría las oscilaciones aumentando el orden del cristal y mejorando la difracción. La cristalografía a bajas temperaturas se denomina criocristalografía y las sustancias utilizadas (en que se bañan los cristales) para prevenir la formación de hielo se llaman criopreservador, entre los más utilizados están el glicerol, azúcares como glucosa y polímeros como el polietilenglicol²⁵. El desarrollo de los sincrotrones y las metodologías de crioprotección ha permitido incrementar el número de estructuras resueltas a resolución atómica ($\geq 1.2 \text{ \AA}$) que permiten describir con mayor detalle los modelos estructurales.

Como consecuencia del crecimiento exponencial de estructuras moleculares resueltas surge el Banco de Datos de Proteínas (PDB) que en la actualidad* cuenta con 30,857 estructuras totales, 26,314 (85%) resueltas por difracción de rayos X y el 15% restante por resonancia magnética nuclear (NMR)²⁶.

DIFRACCIÓN DE RAYOS X

Bajo condiciones específicas una gran variedad de moléculas (incluidas las proteínas) se solidifican para formar cristales, estructuras tridimensionales ordenadas mantenidas por uniones no covalentes. En cristalografía lo que se obtiene de los cristales irradiados es una representación de la densidad electrónica molecular en la celda unitaria promedio. Con esta imagen se intenta encontrar en el espacio tridimensional cada átomo de la molécula, descrito por las coordenadas cartesianas x , y y z .

El proceso para la solución de una estructura o modelo a partir de rayos X se puede dividir en nueve subprocesos: cristalización, colecta de datos, procesamiento y determinación de simetrías, cálculo de fases, interpretación de la densidad electrónica (generación y afinamiento de un modelo) y aunque no es parte de la determinación estructural el análisis de la relación estructura-función es, quizá, la parte más importante.

La cristalización es, para muchos, el cuello de botella en el proceso de resolver una estructura cristalográfica. Se logra explorando el espacio, o intervalo de concentraciones proteína-precipitante, que hay entre solubilidad y precipitación. Para ello se forman gotas constituidas por las soluciones en que se encuentra la proteína y la precipitante a diferentes concentraciones. Luego deben optimizarse las condiciones en que aparezcan estructuras cristalinas para obtener cristales de buena calidad. El tamaño mínimo que debe tener un cristal para difractar es de $50 \mu\text{m}$ en su longitud más corta. La calidad del cristal es importante, veremos más adelante, para el número de observaciones y la relación datos/parámetros.

* hasta el 10 de mayo del 2005.

Durante la colecta de datos (medición de las difracciones u obtención del patrón de difracción) lo más importante es la estrategia para obtener la mayor cantidad posible de observaciones (difracciones) que permitan determinar la estructura. Éste depende directamente del grupo espacial o simetría del cristal, la cual se calcula durante las primeras difracciones. También se debe cuidar que el sistema de enfriamiento y la humedad del ambiente no provoquen la formación de hielo alrededor del cristal, así como el decaimiento de la capacidad de difracción del mismo durante la medición.

Después de la colecta de datos hay que procesarlos y confirmar la simetría del cristal. Durante el procesamiento se obtienen y ordenan los índices h,k,l e intensidades de cada difracción en el patrón (**Patrón de difracción y Factores de estructura**). *Grosso modo*, la simetría se obtiene al analizar la posición de las reflexiones y a partir de ellas deducir el arreglo (red) en que se encuentran. La posterior búsqueda de ausencias sistemáticas permite afinar la selección del grupo espacial al comparar los datos experimentales con las ausencias sistemáticas^f que se conocen (han sido calculadas) para cada simetría.

El cálculo de fases^g, es uno de los problemas centrales en la solución de estructuras moleculares por rayos X. Se pueden obtener por diferentes métodos, pero los más utilizados son derivados de átomos pesados y reemplazo molecular. El método de derivados de átomos pesados consiste en comparar el patrón de difracción de cristales nativos con los de cristales a los que se agrega sales metálicas por difusión, con la finalidad de encontrar la posición de los átomos pesados y a partir de ellas aproximar las fases. En el caso del reemplazo molecular se utiliza una estructura homóloga que se rota y traslada para determinar la orientación relativa de los ejes moleculares y su posición en el espacio con respecto a los elementos de simetría cristalográficos. Conociendo la orientación y posición es posible determinar un juego de fases iniciales a partir de las que se calcula un mapa de densidad electrónica¹⁰⁹. El modelo se va ajustando (afinando) sobre el mapa iterativamente hasta obtener al que mejor explica el patrón de difracción.

La interpretación de la densidad electrónica incluye la generación y afinamiento de un modelo a partir de las difracciones. Por afinamiento entendemos el método en el que un modelo es modificado para reducir la discrepancia entre los datos experimentales y el modelo²⁷, respetando siempre las restricciones geométricas de los enlaces peptídicos. En un visualizador gráfico se despliegan el mapa de densidad calculado y el modelo. Posteriormente se mueven los residuos hasta que el mayor número de átomos esté dentro de densidad, enfatizando en los átomos de cadena principal y de mayor número de electrones (S, N y O).

El objetivo de obtener una estructura macromolecular es el encontrar las bases moleculares que la relacionan con su función. En general se localizan sitio activo, aminoácidos catalíticos y en caso de presentarse, el sitio de unión de otros ligandos (como cofactores). Para su ubicación lo más

^f Los puntos que no encontraremos en un patrón de difracción de un cristal con simetría x .

^gRetraso de algunos componentes del haz inicial, completamente en fase, causados por la interacción con los átomos del cristal.

común es determinar estructuras en diferentes estados; por ejemplo, con y sin ligandos (sustratos, análogos a sustratos y cofactores). Las variantes mostrarán con más detalle la participación de los residuos de aminoácidos en la unión de ligandos, los cambios de conformación localizados en regiones específicas de la molécula y las interacciones atómicas importantes para la actividad.

SIMETRÍA

La simetría es uno de los temas esenciales para entender la teoría que hay detrás de la cristalografía y un paso clave en la determinación de una estructura obtenida por difracción de rayos X.

Antes de abordar la simetría hay que explicar dos conceptos básicos. La celda unitaria es la unidad cristalina básica; un paralelepípedo definido por tres longitudes (a , b y c) y el mismo número de ángulos (α , β y γ) que reproduce el volumen total del cristal al trasladarla en las tres dimensiones. La red cristalina (*lattice*) es el arreglo de puntos sobre los vértices, caras o centro de la celda unitaria, tales que alrededor de cada uno se encuentre el mismo ambiente y orientación que el que rodea a los demás. Dentro de la celda unitaria el mínimo agregado de moléculas que puede yuxtaponerse en otras entidades por operaciones de simetría es la unidad asimétrica. En el más simple de los casos la unidad asimétrica es una sola molécula de proteína.

Existen 230 grupos espaciales que describen la simetría de la celda unitaria. Se representan por una letra mayúscula (el tipo de red) seguida de tres números (con subíndice) que indican los operadores de simetría ejecutables que reconstruyen el cristal²⁸. Las operaciones simétricas más simples son: rotación, traslación y reflexión. Cuando se combinan dan origen a otras más complejas como son centros de simetría, ejes de rototranslación y planos de deslizamiento. Debido a que las proteínas son moléculas formadas mayoritariamente por un solo tipo de isómero óptico²⁹, ni los planos de espejo ni otro tipo de simetrías que las contengan pueden presentarse en cristales proteicos. Los únicos posibles son rotaciones, translaciones y rototranslaciones. Las limitantes simétricas presentes en proteínas reducen el número de grupos espaciales posibles a 65²⁹. La descripción amplia de los 230 grupos, agrupados en 32 clases, 7 sistemas cristalinos y 14 redes cristalinas de Bravais se encuentra en las Tablas Internacionales para Cristalografía de Rayos X³⁰.

PATRÓN DE DIFRACCIÓN Y FACTORES DE ESTRUCTURA

Para facilitar la comprensión de la primera parte del proceso de colecta de datos es común utilizar la analogía, permitida por la naturaleza electromagnética de los rayos X, con una obtenida por reflexión de luz visible, por ejemplo la de un microscopio. En un microscopio se ilumina el objeto con un rayo de luz. La radiación que impacta el objeto es dispersada, absorbida y reflejada, esta última

²⁹ Compuestas sólo por estereoisómeros L.

es colectada por el lente objetivo y recombinado para formar la imagen de la muestra. En cristalografía el cristal es irradiado por un haz de rayos X (paralelos) que también se absorbe y difracta al contactar los átomos al interior del cristal. La diferencia radica en la falta de una lente para coleccionar y recombinar los rayos difractados que, en este caso, son recibidos en un detector (anteriormente papel fotográfico) sensible a la radiación X. La carencia de lente se debe a que no hay sustancias capaces de enfocar rayos X. La teoría de difracción de la luz está tan desarrollada que es posible recombinar los rayos difractados de manera analítica con la ayuda de una computadora. Para ello hay dos problemas por resolver: cuantificar las intensidades y obtener las fases. El primero se resuelve experimentalmente con detectores sensibles capaces de cuantificar la intensidad para cada punto del patrón de difracción. Experimentalmente se pierde la información de las fases; sin las cuales es imposible determinar estructuras

Las teorías desarrolladas principalmente por Bragg y von Laue se basan en el hecho de que los electrones dispersan o difractan los rayos X. Los rayos difractados se suman (interferencia constructiva) bajo condiciones particulares del interior cristalino formando los puntos del patrón de difracción. El problema se facilita si visualizamos los átomos del cristal como planos, tal como lo hizo Bragg. Los planos se designan por tres índices hkl (de Miller) que identifican un conjunto de planos paralelos. Entonces cada punto o reflexión del patrón se caracteriza por tres índices enteros h, k, l más una intensidad $I(h, k, l)$. Los índices especifican el número de planos por celda unitaria en el inverso de las tres direcciones del espacio tridimensional: h , en la dirección x ; k , en la dirección y y l , en z , mientras que la intensidad proporciona la amplitud de onda³¹.

La suma de ondas o interferencia constructiva ocurrirá entre rayos dispersados por planos paralelos de índice h, k, l , en la estructura interna del cristal, si y sólo si los rayos X de longitud de onda λ inciden y reflejan sobre los planos a un ángulo θ que cumple la condición (ley de Bragg)³²:

$$2d_{hkl} \sin\theta = n\lambda \quad ,$$

es decir, si los rayos están desfasados en un entero (n) de λ . Si el desfasamiento es distinto a un entero de λ entonces la interferencia será destructiva y el rayo se cancelará en ese ángulo particular.

Esta ley predice la posición en el espacio de cualquier rayo difractado. Es clave notar que mientras menor sea la separación de los planos (representada por d) mayor será el ángulo de difracción; el ángulo θ es proporcional a $1/d$.

Al contacto con la materia los rayos X tienen tres formas de interacción³³: el efecto fotoeléctrico y dos tipos de difracción, coherente e incoherente. El efecto fotoeléctrico consiste en la absorción de fotones que provoca el desprendimiento de electrones del orbital más externo del átomo. El retorno al estado basal es acompañado de la emisión de un fotón con longitud de onda característica del átomo, emisión absorbida que no contribuye al patrón de difracción. La difracción coherente o clásica

(de Thomson) es consecuencia de la incidencia de la radiación sobre electrones libres. La influencia del campo electromagnético fluctuante de la onda incidente induce al electrón a oscilaciones de la misma frecuencia. La carga oscilante produce un rayo secundario con la misma λ pero desfasado 180° . Se llama coherente porque todos los rayos dispersados de un electrón tienen la misma relación de fase respecto del rayo incidente³⁴. La incoherente (de Compton), consiste en el desprendimiento de un electrón de su átomo, causada por el fotón incidente, que produce un fotón difractado de menor energía³⁵. La interacción dominante, al menos en el intervalo de energías utilizadas en cristalografía, es el efecto fotoeléctrico, pues involucra más del 80% del total de interacciones. Consecuentemente la mayor parte de los rayos X que interaccionan con el cristal¹ depositan en él su energía, produciendo daño por radiación o decaimiento del cristal, una de las principales limitantes en los experimentos de difracción.

En un átomo los electrones están enlazados en estados energéticos bien definidos y ocupan un volumen finito. Si queremos medir la dispersión causada por un átomo primero debemos considerar la distribución espacial de los electrones. El factor atómico de dispersión calcula la onda total dispersada por un átomo mediante la suma de las contribuciones individuales sobre el volumen atómico, representada por la ecuación:

$$F(S) = \int_{\text{vol. atómico}} \rho(r) \exp(2\pi i r \cdot S) dv ,$$

en donde $\rho(r)$ es la densidad electrónica y $2\pi r \cdot S$ la fase (con S como la diferencia del vector que representa la radiación incidente, s_0 , menos el que representa la dispersada, s , $S = s_0 - s$). Cuando extendemos este concepto a un cristal la expresión que se obtiene es la ecuación de factor de estructura (demostrada matemáticamente):

$$F_{hkl} = \sum_{j=1}^n f_j \exp 2\pi i (hx_j + ky_j + lz_j)$$

El factor de estructura es una serie de Fourier² que describe a la reflexión hkl , en la que cada término es la contribución de un átomo tratado como una esfera de densidad electrónica. Así la contribución de cada átomo j a F_{hkl} depende del elemento que es, lo que determina f_j , la amplitud de la contribución y de su posición (x_j , y_j , z_j) en la celda unitaria que establece la fase de su contribución.

¹ Menos del 5% del total irradiado.

² Una transformada de Fourier describe a una onda compleja como la suma de ondas simples.

En su libro sobre cristalografía, Rhodes³⁶ escribió: ‘ *Seemingly simple processes like the binding of small ligands to proteins often involve conformational changes to states not revealed by crystallographic analysis*^o. Refiriéndose a que los modelos cristalográficos muestran estructuras estáticas que representan una conformación estable de una molécula dinámica. La tesis principal de este trabajo es mostrar que los modelos cristalográficos contienen información sobre la dirección de los movimientos (u oscilaciones) generales que ocurren en la proteína y que están directamente relacionados con los cambios de conformación estabilizados por la unión de ligandos. Para ello necesitamos obtener valores de temperatura anisotrópicos (direccionados), que representan mejor la incertidumbre en la ubicación de los átomos.

Existen dos principios físicos para la incertidumbre posicional de los átomos durante la difracción. El primero es el movimiento térmico u oscilatorio que se refiere al desplazamiento o vibración de un átomo (de sus electrones) alrededor de su posición de descanso o mayor probabilidad. El otro es el desorden cristalino, es decir, los átomos o grupos de átomos equivalentes que no ocupan una posición idéntica en cada celda unitaria, unidad asimétrica o, incluso, en los distintos monómeros de la unidad asimétrica. El factor de temperatura B_j calculado durante el afinamiento refleja la presencia de ambos elementos físicos de incertidumbre en el átomo j dificultando el distinguir uno del otro.

Si B_j fuera puramente una medida de movimiento oscilatorio isotrópico (*isos*, igual; *tropos*, movimiento o desplazamiento), el caso más simple de movimiento armónico^Y con igual magnitud en todas las direcciones, el valor de B_j dependería de la magnitud en la siguiente forma:

$$B_j = 8\pi^2 \{u_j^2\} = 79 \{u_j^2\}^* ,$$

En donde el término $\{u_j^2\}$ es el desplazamiento medio cuadrado de la posición promedio del átomo j . Un factor $B = 79 \text{ \AA}^2$ implica oscilaciones de 1 \AA sobre la posición promedio. Los valores considerados en las regiones menos móviles corresponden a oscilaciones de $0.25-0.5 \text{ \AA}^2$ (B_s de entre $5-20 \text{ \AA}^2$). Sin embargo los valores térmicos obtenidos durante el procesamiento de datos experimentales son muy altos, tanto que es difícil explicarlos como pura vibración atómica por lo que deben también reflejar desorden³⁷.

Debido al tamaño finito de la nube electrónica alrededor del núcleo, el poder de dispersión de un átomo disminuye conforme aumenta el ángulo de dispersión. Este decaimiento contiene información valiosa que puede cuantificarse si se añade un término ángulo-dependiente a las contribuciones

^o Procesos relativamente simples como la unión de pequeños ligandos involucran cambios de conformación a estados no revelados por el análisis cristalográfico.

^Y El movimiento armónico simple es un movimiento periódico de vaivén, en el que un cuerpo oscila de un lado a otro de su posición de equilibrio, en dirección determinada, y en intervalos iguales de tiempo.

* Algunos autores lo definen como $B = 8/3\pi^2 \langle u_j^2 \rangle$.

atómicas individuales de dispersión f que al sumarse dan las intensidades calculadas F_{calc}^2 . La corrección tiene la siguiente forma:

$$f = f_0 \exp(-2\pi^2 \langle u^2 \rangle \mathbf{h}^T \mathbf{h}) = f_0 \exp[-8\pi^2 \langle u^2 \rangle (\sin^2 \theta / \lambda^2)] ,$$

donde $\langle u^2 \rangle$ es la amplitud media cuadrada de la vibración del átomo, \mathbf{h} un vector del red cristalina recíproca, θ el ángulo de dispersión y λ la longitud de onda.

Al promediar en el tiempo la nube de electrones, esta es considerablemente mayor en un átomo que vibra que en uno en reposo. Debido a que el parámetro u es dependiente de la temperatura se le denomina parámetro térmico o factor de temperatura. Sin embargo, no depende sólo de la temperatura sino que es consecuencia de la variación estocástica de la posición promedio del centro atómico de una celda unitaria a otra, dependiendo su magnitud de los contactos cristalinos, la región de la proteína (interna o externa) en que se encuentra y de las interacciones con átomos circundantes. Por ello también se le denomina parámetro o factor de desplazamiento.

El parámetro B describe una nube electrónica uniforme dispersa en todas las direcciones por lo tanto u es isotrópico. Sabemos que los modos oscilatorios de los átomos enlazados (todos los de la proteína) no son de este tipo, aun así, el B isotrópico es una buena aproximación al comportamiento de oscilación de los átomos reales de una proteína en una red cristalina; describe adecuadamente el modelo obtenido a partir de las reflexiones.

Un modelo mejor requeriría relajar la insensibilidad de dispersión hacia una dirección sin perder la simplicidad computacional. Para eso se puede expandir u a un tensor simétrico \mathbf{U} expresado en una matriz de 3×3 ³⁸:

$$\mathbf{U} = \begin{pmatrix} \mathbf{U}^{11} & \mathbf{U}^{12} & \mathbf{U}^{13} \\ \mathbf{U}^{21} & \mathbf{U}^{22} & \mathbf{U}^{23} \\ \mathbf{U}^{31} & \mathbf{U}^{32} & \mathbf{U}^{33} \end{pmatrix}$$

De esta forma el término se vuelve anisotrópico:

$$f = f_0 \exp(-2\pi^2 \langle u^2 \rangle \mathbf{h}^T \mathbf{U} \mathbf{h})$$

Los seis componentes independientes del tensor \mathbf{U}^{ij} (en negritas) son los parámetros del átomo que describen su distribución probabilística de densidad electrónica en una Gaussiana tridimensional; un elipsoide cuyo centro es el punto de mayor probabilidad (**Figura 2**).

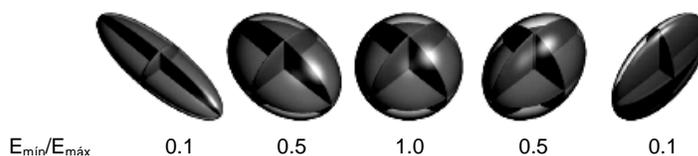


Figura 2 Representación de parámetros atómicos isotrópicos y anisotrópicos. E_{\min} y E_{\max} representan, respectivamente, los valores de los ejes más corto y más largo del elipsoide. Un valor $E_{\min}/E_{\max}=1.0$ equivale a una esfera (isotrópico), mientras que valores distintos de 1.0 muestran diferentes grados de anisotropía.

CRITERIOS DE VALIDACIÓN.

La resolución limitada y la información inexacta de las fases con las que lidian los cristalógrafos de macromoléculas hacen que la construcción y afinamiento de un modelo basado en difracción de rayos X no sea una tarea sencilla³⁹. Es decir, los errores en el proceso son casi inevitables y es tarea del cristalógrafo reducirlos lo más posible antes del análisis, publicación y depósito de la estructura. Esos errores repercuten o son absorbidos por los valores de temperatura, tanto, que aumentos drásticos en los valores B son medidas indirectas del decaimiento cristalino.

Aunque no es un criterio estadístico de validación, la resolución de una estructura es el primer dato que se revisa pues nos indica el nivel de detalle del modelo. Conceptualmente la resolución de una imagen es la distancia mínima distinguible entre dos puntos. En cristalografía está relacionada con la distancia entre los planos h,k,l ; retomando la Ley de Bragg:

$$\text{si } 2d_{hkl} \text{ sen}\theta = n\lambda \Rightarrow d_{hkl} = n\lambda/2\text{sen}\theta ,$$

la resolución máxima es la distancia menor entre planos d_{hkl} , contenida en el patrón de difracción, que a su vez es inversamente proporcional al ángulo de incidencia. Por lo tanto los puntos bien definidos que se encuentren más alejados del centro del patrón de difracción serán los que contengan información más detallada y determinen el límite de resolución. La resolución depende, entre otras cosas, de la calidad del cristal, la fuente de radiación y las oscilaciones atómicas (Figura 5).

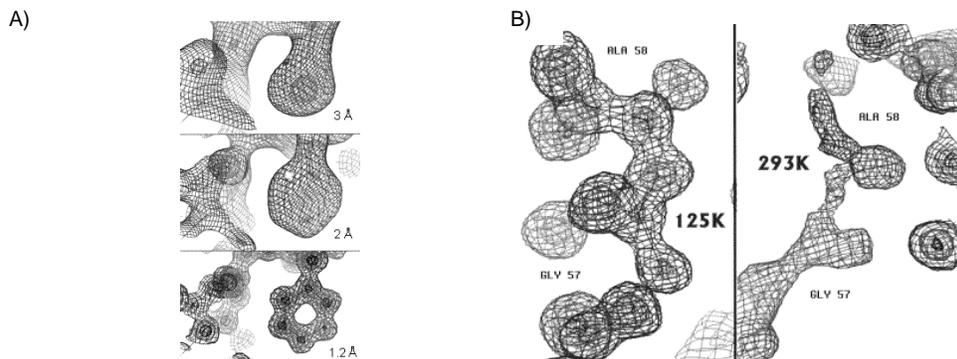


Figura 3 Resolución en mapas de densidad electrónica. A) Densidad electrónica de una fenilalanina a resoluciones 3 Å, 2 Å y 1.2 Å. B) Mejora en la definición dependiente de la temperatura. A 293°K se observa menor definición causada por las oscilaciones atómicas. *Mapas como mínimo a 1σ .

Entre los criterios matemáticos o estadísticos de validación de una estructura resuelta por difracción de rayos X, sobresale el valor R. Durante la búsqueda de localización espacial de los átomos de la molécula, el criterio de evaluación que se utiliza es la correspondencia entre las amplitudes esperadas (teóricas) de los factores de estructura calculados del modelo y las amplitudes contenidas en los datos experimentales (el patrón de difracción). La correspondencia entre los datos teóricos y los experimentales es lo que se conoce como valor R, también llamado índice residual; un criterio de la mejoría de fases, o bien, de la correlación entre el modelo y los datos experimentales durante el proceso de afinamiento de la estructura. El valor R es la sumatoria de la diferencia entre las amplitudes observada y calculada de los factores de estructura, dividida por la sumatoria de la amplitud observada, para cada una de las reflexiones:

$$R = \frac{\sum_{h,k,l} | |F_{\text{obs}}(h,k,l) - F_{\text{calc}}(h,k,l) | |}{\sum_{h,k,l} |F_{\text{obs}}(h,k,l)|}$$

Entonces un valor bajo de R indica una alta congruencia entre el modelo y la molécula del cristal, mientras que un valor alto de R significa baja correspondencia entre el modelo y los datos experimentales; el modelo no ejemplifica lo que hay en el cristal. El valor R cuantifica lo prededecible del modelo actual sobre los datos experimentales. El valor ideal de R (R=0), es imposible de obtener por razones experimentales que ya han sido explicados (desorden y oscilaciones principalmente). Se consideran como indicadores de buena correspondencia valores de $R \leq 20\text{-}25\%$. Como estimados iniciales de fases se aceptan valores de 40-30%. Un valor de 60% equivale a comparar un modelo con un conjunto de factores de estructura al azar, lo que en otras áreas se denomina la zona de incertidumbre.

Otro criterio de evaluación es el índice de residuales libre o valor R_{free}^{40} , así llamado porque es independiente de R. Este es un estadístico más exigente y revelador de la calidad del modelo que se calcula con un conjunto reducido (generalmente de 5-10%) de reflexiones apartadas desde el inicio del cálculo de R:

$$R_{\text{free}} = \frac{\sum_{h,k,l \in T} | |F_{\text{obs}}(h,k,l) - F_{\text{calc}}(h,k,l) | |}{\sum_{h,k,l \in T} |F_{\text{obs}}(h,k,l)|}$$

Donde $\in T$ significa que pertenece al conjunto de los factores de evaluación omitidos del proceso de modelamiento. Este estadístico evalúa, en cada etapa del afinamiento, la capacidad de predicción del modelo sobre un subconjunto de intensidades cuantificadas que no se incluyen en el afinamiento. También se ha utilizado para detectar errores en las fases, optimizar procesos de afinamiento, prevenir sobreajuste de estructuras y modificar mapas de densidad⁴¹. Aunque en sus orígenes no se depositaba en el PDB, en 1995 el 33% de las estructuras cristalográficas lo reportaron contra el 92% en el año 2000⁴¹, se ha impuesto porque muchos cristalógrafos lo consideran un buen criterio sobre la calidad del modelo, la convergencia de los procesos y sobre todo como indicador de sobreafinamiento de las estructuras.

La figura de mérito (FOM) es una función que mide la concordancia entre los datos y el modelo a ajustar para parámetros particulares escogidos. En cristalografía se realiza sobre las fases y se obtiene por:

$$\text{FOM} = ((\cos \varphi)^2 + (\sin \varphi)^2)^{1/2} \quad ,$$

en donde φ simboliza las fases^{12, 42}.

Además de los criterios mencionados anteriormente existen otros indicadores de convergencia del modelo como son las desviaciones medias cuadradas (*root mean square deviations*, r.m.s.d.) de los valores estándares calculados y permitidos para longitudes y ángulos de enlace. Estos valores indican si el modelo es química, estereoquímica (compuesto sólo por estereoisómeros L) y conformacionalmente (enlaces peptídicos planos, con ángulos Φ y Ψ dentro de rangos permitidos) válido, es decir, si cumple con las restricciones físico-químicas impuestas. En el caso particular de los ángulos Φ y Ψ , se utilizan los diagramas de Ramachandran para distinguir residuos cuyos ángulos de torsión se encuentran fuera de los permitidos y por tanto muestran posibles errores durante el afinamiento⁴³.

Es importante resaltar que todos los criterios de validación deben ir mejorando durante el proceso de afinamiento hasta alcanzar la convergencia, de no ser así indicaría un sobreafinamiento[∇] o errores en la construcción del modelo.

AFINAMIENTO POR TENSORES TLS

Antes de pasar al afinamiento por tensores TLS creo apropiado explicar, aunque sea superficialmente, lo que es un tensor. Un tensor de rango n ésimo en el espacio de m dimensiones es un objeto matemático que tiene n índices y m^n componentes y obedece ciertas reglas de transformación⁴⁴. Son generalizaciones de escalares (que no tienen índices), vectores (que tienen un solo índice) y matrices (de dos índices), a un número arbitrario de índices. Se utilizan para resolver problemas en áreas de la física como elasticidad, mecánica de fluidos y relatividad general.

El parámetro total de desplazamiento atómico (U) se compone de la suma de cuatro contribuciones separadas en general anisotrópicas⁴⁵:

$$U = U_{\text{cristal}} + U_{\text{TLS}} + U_{\text{interna}} + U_{\text{átomo}}$$

El término U_{cristal} representa la anisotropía global del cristal y es sólo un parámetro de desplazamiento anisotrópico aplicado a todo el contenido de la celda unitaria. El que más nos interesa es U_{TLS} , que representa translaciones y libraciones[§] de cuerpos pseudo-rígidos al interior de la unidad asimétrica. La tercera contribución, U_{interna} , incluye varios tipos de movimientos

[∇] En particular cuando la diferencia entre los valores de R y R_{free} aumenta durante el proceso de afinamiento.

[§] Movimientos oscilatorios que un cuerpo, ligeramente perturbado del equilibrio, efectúa hasta recuperarlo poco a poco.

colectivos intramoleculares tales como libración sobre ángulos de torsión particulares o modos normales internos de una molécula. Por último $U_{\text{átomo}}$ considera desplazamientos locales de átomos individuales.

Por evidencia de varias fuentes: afinamiento estructural⁴⁶, análisis de modos normales⁴⁷ y dinámica molecular⁴⁸, podemos afirmar que los desplazamientos medios cuadrados de grupos de átomos seleccionados en estructuras macromoleculares es representable en términos de desplazamientos grupales. Bajo la aproximación armónica, dichos desplazamientos de grupos se pueden describir en términos de tensores de traslación (T), libración (L) y rototranslaciones (S). También se ha mostrado que los parámetros TLS representan la mayor contribución de los desplazamientos atómicos sobre el movimiento global⁴⁵.

El primero en proponer un tratamiento para describir el movimiento de una molécula, en estado cristalino, por cuerpos pseudo-rígidos fue Cruickshank (1956)⁴⁹. Lo hizo mediante dos tensores simétricos, uno para la libración (L) y otro para la traslación (T). Tras su utilización se consideró poco adecuado porque, bajo su propuesta, los ejes de libración no están restringidos a intersectarse en un punto dado (e. g. un centro de simetría). Para resolver este problema Shomaker y Trueblood⁵⁰ añadieron un tercer tensor (S) necesario para tomar en cuenta la correlación cuadrática promedio de traslación y libración. Cuando se incluye a S en el análisis, el ajuste de los tensores atómicos individuales de vibración (U_s) es independiente del origen, aunque los componentes de T y S varíen con él. Para abolir esta arbitrariedad y permitir que los datos se puedan repartir de manera uniforme es necesario establecer algunas condiciones que se aplican después de haber encontrado T, L, y S para cada origen inicial arbitrario. Si el sitio de la molécula es centrosimétrico, entonces S se iguala a cero y el análisis se reduce al de dos tensores de Cruickshank. También S se desvanece o es constante para otro tipo de simetrías.

Para comprender de lo que trata el afinamiento por tensores TLS es necesario primero revisar el análisis de movimiento de cuerpos rígidos. El más general de los desplazamientos de un cuerpo rígido fijado en un punto es una rotación sobre algún eje que pasa por el mismo punto (Teorema de Euler)⁵¹. El desplazamiento que corresponde a cualquiera de los otros puntos \mathbf{r} relativamente fijos es un vector lineal en \mathbf{r} que se puede expresar como $\mathbf{D} \cdot \mathbf{r}$; siendo \mathbf{D} el tensor rotacional de desplazamiento o matriz de rotación. El conflicto está en expresar D en términos del ángulo $\lambda \equiv |\boldsymbol{\lambda}|$ y la dirección del eje de rotación $\boldsymbol{\lambda}/\lambda$, para agregar a $\mathbf{D} \cdot \mathbf{r}$ cualquier traslación \mathbf{t} del punto fijo para obtener el desplazamiento total \mathbf{u} desde el punto \mathbf{r} y evaluar los componentes $u_{ij} \equiv u_i u_j$ de sus tensores de vibración al promediar sobre todos los aspectos del movimiento, siendo λ y \mathbf{t} funciones dependientes del tiempo y/o posición en la red cristalina.

Entonces, en una molécula rígida el desplazamiento rotacional instantáneo de un átomo en \mathbf{r} relativo al punto fijo está dado por:

$$\mathbf{D} \cdot \mathbf{r} = \boldsymbol{\lambda} \times \mathbf{r} (\text{sen } \lambda)/\lambda + [(\mathbf{r} \cdot \boldsymbol{\lambda}) \boldsymbol{\lambda} - \lambda^2 \mathbf{r}] (1 - \text{cos } \lambda)/\lambda^2 \quad [\text{a}]$$

En donde $\lambda \equiv |\lambda|$ es el ángulo de rotación sobre un eje a través del origen y paralelo a λ : el primer término a la derecha es componente de $D \cdot r$ perpendicular a λ y r , mientras que el segundo – con $(r \cdot \lambda) \lambda - \lambda^2 r$ idéntico a $\lambda \times (\lambda \times r)$ – es el componente restante, en el plano λ y r , perpendicular a λ y pequeño comparado con el primero si λ es pequeño (**Figura 4**).

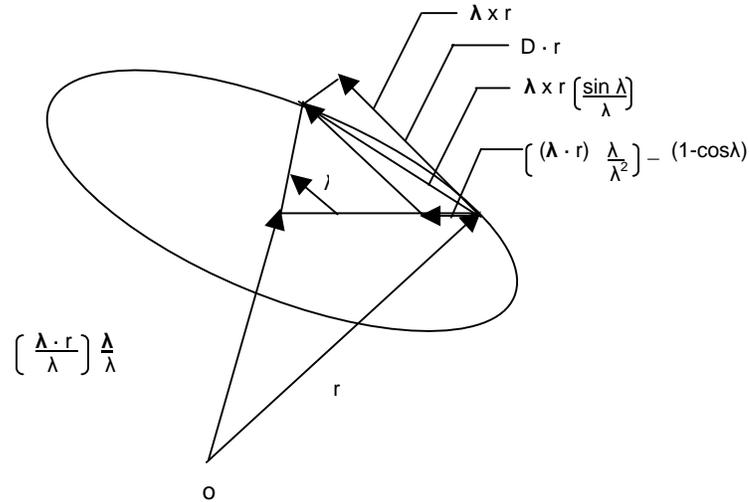


Figura 4 Representación gráfica de los términos en la ecuación [a].

Si la molécula también se traslada, el vector (columna) de distancia t el desplazamiento total (u) del átomo en la posición original r se denota:

$$u = D \cdot r + t \quad [b],$$

Para desplazamientos cortos el primer término en [b] se puede linearizar con respecto a la amplitud de rotación:

$$u \simeq \lambda \times r + t \quad [c],$$

con λ siendo un vector a lo largo del eje de rotación y de la magnitud del ángulo de rotación y \times denotando producto cruzado. Así la contribución de un átomo de un cuerpo rígido a los parámetros anisotrópicos de desplazamiento se puede derivar a partir del cuadrado del desplazamiento dado en la ecuación [c], esto es:

$$uu^T = tt^T + t\lambda^T \times r^T - r \times \lambda t^T - r \times \lambda \lambda^T \times r^T \quad [d],$$

T denota el vector renglón (el transpuesto del vector columna).

Si promediamos todos los desplazamientos en tiempo y espacio tenemos:

$$U \equiv \langle uu^T \rangle = T + S^T \times r^T - r \times S - r \times L \times r^T \quad [e],$$

con $T \equiv \langle tt^T \rangle$, $L \equiv \langle \lambda \lambda^T \rangle$ y $S \equiv \langle \lambda t^T \rangle$. En este contexto el producto cruzado $L \times r^T$ da una matriz en que el i ésimo renglón es el producto cruzado del i ésimo renglón de L y r^T .

La ecuación [e] describe el desplazamiento medio cuadrado en términos de los tensores T , L y S . Si consideramos un conjunto de puntos (átomos) $\{r\}$, correspondientes al resto de las posiciones atómicas dentro del rígido, U es el desplazamiento medio cuadrado de cada uno de los átomos y

puede identificarse como el parámetro anisotrópico de desplazamiento que ocurre en el factor Debye-Waller en la expresión para factores de estructura. Los tensores simétricos T y L describen respectivamente la traslación y libración media cuadrada del cuerpo rígido, el tensor no simétrico S describe la correlación media cuadrada entre desplazamiento traslacional y libracional. La expansión del término derecho de [e] no incluye el rastro de S, por lo que sólo aporta 8 parámetros independientes a [e], los que junto con los independientes de T y de L suman en total 20 parámetros por grupo, correspondientes a 3 amplitudes rotacionales, 3 amplitudes traslacionales, 3 correlaciones de rototraslación (sólo dos de ellas determinables), 9 ángulos de orientación de los principales sistemas de ejes y tres coordenadas del nuevo origen.

Dado que un conjunto de parámetros anisotrópicos de desplazamiento (ADPs) se puede utilizar para determinar los parámetros TLS para cada cuerpo rígido por medio de un ajuste de mínimos cuadrados, se puede partir, inversamente, de [e] para derivar parámetros anisotrópicos y por lo tanto factores de estructura calculados a partir de los parámetros TLS durante el afinamiento. Algo que se hace comúnmente es añadir a U_{TLS} los valores B isotrópicos para obtener el parámetro total de desplazamiento usado con el factor de estructura calculado⁵². De esta forma el desplazamiento medio cuadrado de cada cuerpo es descrito por 20 parámetros independientes más uno isotrópico por átomo, en lugar de los 6 parámetros/átomo que significa el afinamiento anisotrópico. El número total de parámetros depende del número de cuerpos seleccionados. En conclusión, la ventaja del afinamiento por TLS es que reduce considerablemente el número de parámetros (de seis por átomo a veinte por grupo) con la interpretación de cuerpos que se mueven independientemente, proporcionando información sobre la dirección del movimiento de los cuerpos (**Figura 5**).

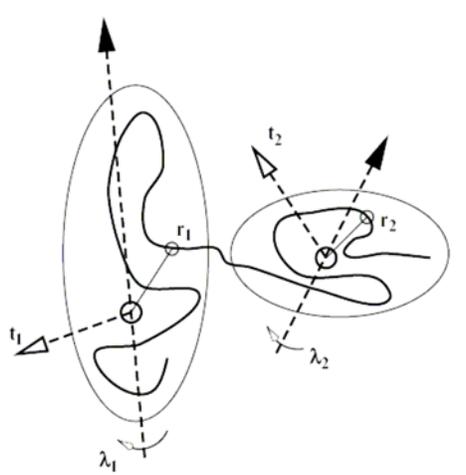


Figura 5 Ejemplo de una proteína dividida en dos cuerpos rígidos. t representa la traslación λ la libración. Las posiciones de reposo están representadas por r_1 y r_2 .

En realidad los grupos asociados con tensores de cuerpos rígidos no son del todo rígidos porque presentan una variedad de movimientos internos a los grupos (incluidos en U_{interna}), consecuentemente a partir de ahora los llamaremos cuerpos pseudo-rígidos.

Hay varios trabajos en los que se ha utilizado el afinamiento por tensores TLS para afinar estructuras de diferentes moléculas seleccionando distintamente los grupos pseudo-rígidos. Desde la división de un dodecámero de DNA en ribosa, fosfatos y bases nitrogenada⁵³, hasta moléculas completas⁵⁴, pasando por cadenas laterales⁵⁵ y subdominios estructurales⁵⁶.

INTERPRETACIÓN DE LOS RESULTADOS

Como en todos los programas de afinamiento modernos, el archivo de salida de `refmac`[†] proporciona tanto estadísticas globales (R , R_{free} , coeficiente de correlación y figura de mérito, ver § **Criterios de validación**) como información geométrica. Además de los estadísticos anteriores, se obtienen los parámetros TLS afinados contenidos en los archivos *logfile*, TLSOUT y en el encabezado de XYZOUT. Los valores B afinados también se incluyen en el XYZOUT y son valores residuales que se afinan separadamente tras la determinación de los parámetros TLS. En éstos no está incluida ninguna contribución de los TLS, por lo que no representan la desviación media cuadrada de los átomos⁵². El archivo TLSOUT contiene la selección de los grupos y los registros ORIGIN, T, L y S de cada uno. Los registros T y L enlistan los seis elementos de los tensores simétricos **T** y **L**. De igual forma el registro S proporciona los ocho elementos determinables del tensor **S**. Los valores **T** y **S** dependen del origen, calculado por el programa e impreso en el registro ORIGIN. Las unidades de **T** son Å^2 , las de **L** grados² y de **S** $\text{Å}\cdot\text{grados}$. La contribución del tensor de traslación es aditiva sobre todos los ADP's (Parámetros Anisotrópicos de Desplazamiento) del grupo (ver ecuación [e]). Los valores del tensor de libración dependen directamente del tamaño del grupo pseudo-rígido. El tensor de rototraslación también depende del tamaño del grupo pero de una forma menos clara, normalmente tiene valores bajos si los grupos son relativamente grandes (dominios o moléculas completas). Generalmente los valores de los parámetros TLS son buen indicador del desorden general del grupo, por ejemplo, en cristales que tienen más de una molécula por unidad asimétrica es común encontrar distintos valores de B para cada una, diferencia que se refleja de manera directa en los valores relativos de los parámetros TLS.

Los archivos TLSOUT y XYZOUT son, ahora, los archivos de entrada para el programa `tlsan1`⁵⁷. El archivo de salida de este programa es el archivo de coordenadas con valores anisotrópicos de B que, ahora sí, incluyen la contribución de los TLS. Al visualizar los elipsoides atómicos tenemos una imagen más clara de los desplazamientos de los cuerpos rígidos. En particular podemos tratar de distinguir desplazamientos dominantes relacionados con la función. Sin embargo, se debe tener precaución cuando se interpretan los resultados TLS, pues tienden a sobreestimar los desplazamientos de cuerpos rígidos, impidiendo la discriminación entre desplazamientos estáticos y dinámicos⁴⁰.

[†] La descripción de este y todos los programas utilizados en esta tesis se encuentra al final de la sección de Metodología.

RELACIÓN DATOS/PARÁMETROS

Como ha escrito Merrit³⁸, la complejidad de un modelo a afinar depende del cociente del número de observaciones (datos experimentales) sobre el número de parámetros del modelo (complejidad). Al aumentar la relación datos/parámetros[Ⓞ] se mejora la calidad del modelo estructural afinado, producto de mejores estimaciones para los parámetros individuales. En el caso específico de la cristalografía el número máximo de observaciones es determinado por la calidad del cristal y el desorden, que limitan la resolución de la difracción.

La cristalografía de moléculas pequeñas el número de observaciones permite el afinamiento libre de un modelo atómico completo de la estructura. Pero cuando pasamos a proteínas la situación se complica porque el número de datos, en la mayoría de los casos, no justifica un procesamiento del mismo tipo. La solución al problema es reducir los parámetros, con el inconveniente de que los modelos resultantes de la reducción son modelos simplificados. Bajo este principio operan algunas técnicas cristalográficas ampliamente utilizadas como las constricciones sobre simetrías no cristalográficas⁵⁸, el promedio de densidad⁵⁹ y el afinamiento de torsión espacial⁶⁰.

La proporción datos/parámetros se puede aumentar si se incrementa la cantidad de observaciones experimentales. El subir la resolución de 2 Å a 1 Å significa multiplicar por ocho el número de intensidades cuantificables. Incrementos de este tipo permiten mejorar el modelo si se incluyen parámetros adicionales que describen cualidades más sutiles o complejas de la estructura. Así es posible realizar un afinamiento anisotrópico para cada átomo, incluyendo seis parámetros térmicos en lugar de uno solo (isotrópico). Sumando los tres parámetros de posición tenemos un aumento de cuatro a nueve parámetros. La mejoría se refleja en valores más bajos de R y R_{free} , en mapas de densidad más limpios y menores incertidumbres en las posiciones atómicas.

Como el aumento de la cantidad de datos no es proporcional a la resolución, existe un rango de resolución conocido como zona gris (entre 1.6-1.2 Å), en el que se duda si incrementar el número de observaciones es suficiente para aumentar los parámetros. Alrededor de los 1.6 Å el cociente datos/parámetros para un afinamiento anisotrópico es el mismo que para un isotrópico a 2 Å. El comportamiento del isotrópico se mantiene bien durante el afinamiento por las fuertes restricciones geométricas impuestas a los parámetros de posición x, y, z. Para mantener un comportamiento del mismo tipo en el anisotrópico a ~1.6 Å es necesario mantener restricciones del mismo tipo sobre los seis parámetros térmicos U^{ij} , mismos que se pueden remover al otro extremo de la zona gris, alrededor de 1.2 Å.

Si bien la relación datos/parámetros es dependiente de la simetría del cristal y del contenido de átomos en la unidad asimétrica la **Tabla 1** nos muestra un promedio general de ésta a diferentes

[Ⓞ] La relación aumenta cuando aumentan las observaciones o datos, cuando disminuyen los parámetros o cuando se combinan ambas situaciones

resoluciones, suponiendo un contenido de agua del 50%⁶¹. El cálculo se hizo a partir de estructuras del PDB.

Tabla 1 Relación observaciones/parámetros de acuerdo a la resolución alcanzada y el afinamiento utilizado

Resolución	Observaciones/parámetros (x,y,z y B isotropico)	Observaciones/parámetros (x,y,z y B anisotrópico)
3.5Å	0.5	0.2
3.0Å	0.8	0.4
2.5Å	1.4	0.6
2.0Å	2.8	1.2
1.5Å	6.2	2.8

El afinamiento TLS ha sido implementado en `refmac` para una descripción simple de los desplazamientos anisotrópicos cuando las observaciones no son suficientes para afinar cada átomo individualmente.

CRISTALES, FLEXIBILIDAD ESTRUCTURAL Y OSCILACIONES

Uno de los principales cuestionamientos que se le han hecho a la cristalografía de rayos X es si los cristales representan una verdadera aproximación a las moléculas biológicas o si observamos estructuras diferentes de las que ocurren en el ambiente celular.

Aunque el contenido de solvente en un cristal varía dependiendo del grupo espacial y de la cantidad y tamaño de las moléculas dentro de la celda unitaria, la mayoría de los cristales de proteína contienen entre 50% y 80% de solvente.

Por otra parte, el análisis comparativo entre estructuras obtenidas por resonancia magnética nuclear y por difracción de rayos X ha mostrado alta similitud entre las estructuras, a excepción de zonas de empaquetamiento que se pueden detectar durante el análisis de la estructura. A las dos anteriores hay que añadir la retención de la función o actividad, aún en estado cristalino, como otra evidencia contundente de que las estructuras de proteínas no se modifican significativamente en este estado[∞]. No solo hay varios reportes de cristales catalíticos⁶² sino que se utilizan en diferentes procesos industriales.

ACTIVIDAD ENZIMÁTICA Y DINÁMICA MOLECULAR

La información que se obtenía de la cristalografía parecía ser solamente la posición de los átomos y por lo tanto la estructura general. El avance científico interdisciplinario ha hecho notar que igual o

[∞] Excepto en el caso de las proteínas de membrana.

más importante que la estructura es la variación temporal de las interacciones interatómicas en la función de una proteína, es decir su dinámica.

Se acepta que la flexibilidad estructural es esencial para la catálisis enzimática y para la actividad biológica molecular en general. Pero el concepto de flexibilidad hace referencia a movimientos moleculares en una escala de tiempo; desde fluctuaciones térmicas en picosegundos hasta cambios conformacionales en la escala de milisegundos (o menores) involucrados en la función cinemática. El concepto de dinámica surge del griego *dynamos* que significa fuerzas, es decir lo relativo a las fuerzas⁶³. En este tema en particular se conocen las fuerzas que mantienen las estructuras biológicas y gobiernan los movimientos al interior de las macromoléculas. Entre ellas encontramos a los puentes salinos, puentes de hidrógenos, fuerzas de van der Waals y a las interacciones hidrofóbicas⁶⁴. Todas las fuerzas anteriores son de tipo débil porque sus energías de asociación son parecidas a la energía térmica a temperaturas normales. Las fuerzas fuertes (valga el término), como las presentes en los enlaces covalentes, son de menor relevancia en la función biológica porque no son interacciones que varíen en el tiempo, durante la desnaturalización o cuando se disocia un complejo proteína ligando.

Las propiedades dinámicas de las estructuras de proteínas son un componente importante que contribuye a la función biológica de las mismas. Ejemplo de ello son las fluctuaciones locales de conformación. La abundancia de interacciones atómicas no enlazadas y enlazadas por hidrógeno son las que confieren la estabilidad a proteínas globulares en solución y las mismas que facilitan la fluctuación. A través de los varios grados de libertad permitidos para una cadena polipeptídica, los cambios locales en la estructura afectan sólo a un pequeño subconjunto de las interacciones débiles totales en una proteína plegada. Ello permite a la estructura flexibilidad local significativa o considerable sin perturbar su topología o estructura general⁶⁵. Esta información dinámica es esencial para entender, a un nivel más profundo de descripción, la función biológica de cada proteína.

La dinámica se puede abordar por medio de diferentes técnicas⁶⁶ incluidas las simulaciones en computadora (mayoritariamente analizadas por modos normales)⁶⁷. Cada una de ellas proporciona información diferente sobre la dinámica y sobre las escalas de tiempo en que ocurre, por ejemplo la espectroscopía de Mössbauer detecta el movimiento de un solo núcleo de prueba (Fe), la dispersión de neutrones (*neutron scattering*) un aspecto más global, mientras que la difracción de rayos X promedia sobre horas, a diferencia de las escalas de tiempo exploradas por Mössbauer (10^{-7} s) o la de dispersión de neutrones (10^{-12} - 10^{-8} s) dependiendo de la resolución del instrumento. En la actualidad las simulaciones *in silico* están limitadas a una escala no mayor de 10^{-8} s. Otras formas de abordar la dinámica son: fluorescencia, infrarrojos, absorción óptica, espectroscopía de resonancia magnética nuclear⁶⁸ (NMR) y cinéticas de intercambio de hidrógeno⁶⁹. La técnica restringe la naturaleza de la muestra y viceversa: la dispersión de neutrones requiere de muestras con baja hidratación, la cristalografía cristales, la resonancia trabaja con soluciones pero con moléculas relativamente

pequeñas (menores de 200 aminoácidos), Mösbauer necesita muestras sólidas, mientras que las simulaciones permiten construir entornos^o pero están restringidas en la escala de tiempo.

Una cadena polipeptídica es intrínsecamente flexible ya que muchos de los enlaces covalentes que ocurren en su esqueleto están rotacionalmente permitidos. Por lo tanto estos grupos deben constituir elementos dinámicos fundamentales en una molécula de proteína. Los movimientos relativos de estas colecciones análogas definidas de cuerpos rígidos pueden definir componentes principales de la dinámica proteica.

Bajo todas las consideraciones anteriores la cristalografía de rayos X ha aportado las siguientes observaciones importantes respecto a las fluctuaciones que existen en el interior de una proteína:

- Flexibilidad atómica individual o grupal de acuerdo a los parámetros térmicos de los modelos cristalográficos.
- Flexibilidad de dominios completos propuesta para regiones sin o de baja densidad electrónica o que difiere en diferentes redes cristalinas.
- Cambios de conformación en presencia de ligandos.
- Variabilidad conformacional.

CAMBIO CONFORMACIONAL LOCAL Y SUB-ESTADOS CONFORMACIONALES

Desde el modelo de llave-cerradura propuesto por Fisher⁷⁰ en 1890, que sugería el ajuste completo y directo del sustrato en el sitio activo de la proteína, la idea de la catálisis ha cambiado mucho. En 1958 Koshland⁷¹ propuso la idea de un mecanismo de unión de ligando por cambio de conformación local. Bajo este modelo la unión del sustrato es acompañada o seguida de un cambio de conformación que alinea los grupos catalíticos en su respectiva orientación correcta. Un aspecto esencial del modelo de cambio de conformación local es la flexibilidad estructural, que como hemos mencionado es un concepto ampliamente aceptado en el campo de las proteínas.

Podemos visualizar los cambios de conformación como estabilizaciones de sub-estados conformacionales y no como transformaciones provocadas por eventos particulares como la unión de un ligando. Una proteína en una conformación de equilibrio puede existir en un gran número de sub-estados; las moléculas en diferente sub-estado tienen la misma estructura general pero difieren en configuraciones locales debido a que ocurren rotaciones sobre un enlace sencillo o modificaciones en la distribución de los puentes de hidrógeno. Como la transición de un sub-estado a otro depende de la superación de una barrera energética, el comportamiento dinámico de las proteínas es dependiente de temperatura. A temperaturas suficientemente altas la molécula se mueve rápidamente de un sub-estado a otro, rompiendo y reconstituyendo los enlaces no covalentes y cambiando continuamente la estructura. A bajas temperaturas, en cambio, mucha de esta relajación

^o Es decir, establecer las condiciones (cantidad de agua, iones, etc.).

conformacional se pierde y cada molécula permanece en un sub-estado, arrestada por la barrera energética. En un experimento clave Frauenfelder et *al.*⁷² utilizaron la difracción de rayos X de cristales de la misma proteína (incluso el mismo cristal) a diferentes temperaturas, para mostrar los cambios en los valores Deby-Waller producidos por las fluctuaciones de temperatura. Cumpliendo lo que se esperaba los desplazamientos eran mayores en las difracciones hechas a mayor temperatura (300 K), en las regiones externas de la proteína y en los residuos polares y cargados, debido a las fluctuaciones en el solvente y en la capa de hidratación.

3 OBJETIVOS

OBJETIVO GENERAL

a) Generar información sobre desplazamientos atómicos asociados a cambios conformacionales en estructuras de proteína determinadas por difracción de rayos X.

OBJETIVOS PARTICULARES

a) Obtener del banco de datos de proteínas (PDB, Protein Data Bank) estructuras de proteínas resueltas por cristalografía de rayos X que permitan realizar afinamientos por tensores TLS (traslación, rotación y rototraslación)

b) Seleccionar los cuerpos pseudo-rígidos bajo diferentes criterios y realizar el afinamiento TLS sobre las estructuras seleccionadas.

c) Analizar con criterios cristalográficos (R , R_{free} , FOM) y por medio de la representación de elipsoides térmicos el resultado de los afinamientos y relacionar las oscilaciones con la función biológica.

4 METODOLOGÍA

BÚSQUEDA EN EL PDB

La búsqueda de datos en el PDB (www.rcsb.org/pdb) se realizó por palabras clave relacionadas con cambios conformacionales que permitieran demostrar la relación entre oscilaciones de grupos pseudo-rígidos en un cristal y la función biológica de la proteína. Las palabras seleccionadas fueron,

conformational change, induced fit, motion, open conformation, oscillation, flexibility, two conformations.

En la página principal del PDB se tecleó la palabra seleccionando la opción texto completo. Esto es, el buscador entra a cada uno de los archivos depositados y separa los depósitos que incluyan en alguna parte del texto la palabra tecleada. A partir de los depósitos que contengan la palabra se crea una tabla de salida, en el mismo servidor, con información general de la molécula (nombre, organismo, resolución, fecha). Como primer criterio de selección de calidad se eliminaron de la tabla a aquellos con resolución menor de 2.5 Å.

SELECCIÓN DE ESTRUCTURAS

Los resultados se tabularon y excluyendo a los depósitos cuyo valor de R fuera mayor de 25% y que no tuvieran depositados los factores de estructura (datos necesarios para reafinar el modelo seleccionado). Las claves (PDB-ID) de las estructuras resultantes se corrieron en *DIPAU*⁹, un programa en lenguaje *Perl*³ que entra a una base de datos PDB en disco y da una tabla de salida de 9 columnas que proporciona la siguiente información: PDB-ID, resolución, valor R, valor R_{free} , número de reflexiones, número de átomos, número de parámetros, nombre, unidad biológica y ligandos. Para disminuir el tamaño de la muestra se descartaron todas las estructuras con una diferencia mayor a 5 unidades porcentuales en $R - R_{free}$.

Posteriormente se reagruparon por palabra de búsqueda, nombre de la proteína y organismo, para realizar una nueva búsqueda, en el mismo PDB, pero ahora por el nombre de cada una de ellas y poniendo particular énfasis en las derivadas de las palabras *induced fit* y *conformational change*, pues en las referencias revisadas fueron las que mostraban cambios de conformación más evidentes.

SOBREPOSICIÓN DE LAS ESTRUCTURAS

Reagrupamos a todas las estructuras por proteína para comparar los grupos espaciales de cada modelo. Sobrepusimos los C_{α} en el programa *O⁷³* para localizar el cambio de conformación, cuantificarlo con desviaciones medias cuadráticas (r.m.s.d.) e identificar las estructuras correspondientes a la forma abierta y cerrada. Ubicamos los residuos alrededor de la zona de unión de ligandos con la intención de comprobar que la modificación respondía a su presencia. Después buscamos los grupos de residuos que cambiaran de manera conjunta, superponiendo nuevamente las estructuras y esperando una menor desviación que la obtenida al superponer la molécula completa en sus C_{α} . Por último, para evitar artificios cristalográficos, reproducimos el cristal en busca de posibles contactos causados por el empaquetamiento que pudieran alterar las estructuras.

³ Escrito en nuestro laboratorio por Paula González-Rubio Garrido.

ANÁLISIS DE VALORES B

Para cada una de las estructuras de cada proteína sacamos el promedio del valor de los valores B de cadena principal y graficamos buscando las regiones de mayor flexibilidad y más variables de una forma a la otra. Regresamos al programa de visualización tratando de encontrar correlaciones entre zonas de modificación estructural, unión de ligandos y zonas de más flexibilidad.

SELECCIÓN DE GRUPOS

Para cada proteína se tomaron por lo menos cuatro criterios distintos de selección de cuerpos pseudo-rígidos: toda la molécula como un solo grupo (a manera de control), los dominios estructurales, los dinámicos resultantes del programa DynDom⁷⁴ y los considerados como pseudo-rígidos tras superponer las estructuras libre y con ligandos en combinación con el análisis de la distribución de valores B de cada conformación:

En lo que respecta a los valores B seleccionamos como grupo a las regiones que más variaban de una conformación a otra, es decir, las que presentaban valores B significativamente menores en la forma cerrada que en la abierta.

Los dominios dinámicos se derivaron del programa DynDom, en el que después de encontrar el mejor ajuste entre dos conformaciones de una estructura, se calculan vectores de rotación para residuos o segmentos cortos de cadena. Con un algoritmo se agrupa a los residuos, de acuerdo a los vectores de rotación, en dominios dinámicos.

El criterio de selección de grupos mediante la sobreposición de las estructuras en O, fue el de considerar como un grupo rígido a aquellos residuos contiguos separados por una desviación media cuadrada (r.m.s.d.) mayor de 0.5 Å entre la conformación abierta y la cerrada.

La utilización de diferentes criterios permitió analizar diferentes grupos y comparar los resultados del afinamiento por tensores TLS.

UNIFICACIÓN DEL AFINAMIENTO DE LOS MODELOS

Debido a que los modelos que llegaron hasta este punto fueron afinados en diferentes programas, cada uno con sus criterios de evaluación, era necesario encontrar una forma de alcanzar un mismo punto de partida antes de correr el afinamiento por tensores TLS. Para ello decidimos realizar entre 20-30 ciclos de minimización energética (dependiendo de la convergencia) y afinamiento de valores B individuales en el programa CNS 1.1⁷⁵. Puesto que en ningún caso encontramos diferencias significativas de los datos estadísticos con los del depósito original, decidimos partir con los datos del PDB.

AFINAMIENTO TLS

Con el objeto de comprobar que los valores de R y R_{free} no variarían de forma considerable por el cambio de programa, antes de hacer el afinamiento por tensores se hizo un afinamiento (*restrained*) isotrópico de 5 ciclos en el programa *refmac5*⁷⁶, dentro de *CCP4*⁷⁷, en sus versiones 4.2.2 y 5.0.2 partiendo del archivo de coordenadas (extensión *.pdb) depositado en el PDB y del obtenido de *CNS*. Los afinamientos se corrieron en las dos versiones de *CCP4* porque encontramos incompatibilidad de resultados entre ambas. Posteriormente encontramos que la incompatibilidad de resultados fue causada por modificaciones realizadas a la versión 5.0.2 del programa y en específico del algoritmo principal de minimización⁷⁸.

Aunque en varias de las referencias consultadas⁷⁹ se hacen 20 ciclos de TLS seguidos del mismo número de ciclos de máxima similitud, decidimos hacer como primer intento cinco ciclos de TLS seguidos de cinco de máxima similitud fijando los valores B a un valor de 20.0 siguiendo el comportamiento de los estadísticos en cada etapa, buscando la convergencia. La convergencia de los estadísticos se alcanzó combinando diferentes números de ciclos de TLS y de máxima similitud, así como variando el valor fijo de los Bs desde fijarlos en el más bajo hasta no fijarlos.

Cabe mencionar que ninguna de las estructuras en la conformación cerrada tenía depositados los factores de estructura, por lo que no pudimos realizar el afinamiento sobre ellas.

ANÁLISIS DE VALORES DE TEMPERATURA ANISOTRÓPICOS

Una vez convergidos los estadísticos, se analizaron los archivos .pdb y .tls de salida de *refmac* con el programa *TLSANL*⁸⁰, seleccionando la opción de valores B totales que calcula Bs anisotrópicos para cada átomo de acuerdo a los valores de los tensores del grupo al que pertenecen y genera un archivo de salida con las coordenadas atómicas.

Los archivos de salida fueron representados en su versión o transformación tridimensional tanto en el programa *O* como en *XtalView*⁸¹. El primero para superponer las estructuras utilizadas como archivos de entrada con los archivos de salida y observar las modificaciones causadas por el afinamiento por tensores en cada uno. En *XtalView* se observaron ejes y elipsoides térmicos buscando relaciones entre la dirección de los ejes mayores con los cambios de conformación.

Por último se representaron los elipsoides anisotrópicos de acuerdo al valor de los valores de temperatura u oscilación, de manera tal que los menores fueran de color azul y los mayores rojos, con las subrutinas *Rastep* y *Render* del programa *Raster 3D*⁸². Los ejes fueron representados con el programa *Molscript*⁸³.

PROGRAMAS

Programa	Referencia	Descripción
DIPAU ³	González-Rubio, P. 2003. Vibraciones atómicas, flexibilidad conformacional y la función biológica de proteínas alostéricas: Estudio a partir de estructuras cristalográficas depositadas en el Protein Data Bank. Tesis de Licenciatura por la Facultad de Ciencias U. N. A. M. México.	Programa en lenguaje <i>Perl</i> ³ que entra a una base de datos PDB (en disco duro) y construye una tabla de salida de 9 columnas que proporciona la siguiente información: PDB-ID, resolución, valor R, valor R_{free} , número de reflexiones, número de átomos, número de parámetros, nombre, unidad biológica y ligandos.
O	Jones, T.A. A graphics model building and refinement system for macromolecules. <i>J. Appl. Cryst.</i> 11 , 268-272(1978).	Programa de construcción y afinamiento de modelos. Se basa en el método de máxima similitud (maximum likelihood).
DynDom	http://www.cmp.uea.ac.uk/dyndom/dyndomMain.do	Determina dominios dinámicos de proteínas. Primero busca el mejor empalme de dos estructuras en diferente conformación. Después calcula vectores de rotación para residuos o segmentos de cadena principal. Por medio de un algoritmo identifica a los residuos que roten de manera conjunta y los clasifica como dominios dinámicos.
CNS 1.1	Crystallography & NMR System. <i>Acta Crystallographica D</i> 54 , 905-921(1998).	Es un programa, resultado de una colaboración internacional entre varios grupos de investigación, diseñado para proporcionar de forma accesible los algoritmos más comunes utilizados para la determinación de estructuras macromoleculares. Incluye búsqueda de átomos pesados, búsqueda experimental de fases, modificación de la densidad electrónica, afinamiento cristalográfico y cálculo de de estructuras de NMR.
refmac5 TLSANL	G. N. Murshudov, A. A. Vagin and E. J. Dodson. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. <i>Acta Cryst.</i> (1997). D53 , 240-255.	Programa de afinamiento macromolecular. Utiliza el método de máxima similitud (<i>maximum likelihood</i>). Programa incluido en la interfase de <i>refmac5</i> . Analiza los tensores TLS Para cada grupo de TLS calcula los valores de T, L y S con respecto a un origen.
<i>XtalView</i>	http://www.sdsc.edu/CCMS/Packages/XTALVIEW/xtalview.html	Es un paquete completo para resolver reemplazos isomórficos de estructuras cristalográficas, incluida la construcción del modelo molecular. Representa elipsoides atómicas.
Raster3D <i>rastep</i> <i>render</i>	http://www.bmsc.washington.edu/raster3d/	Programa que genera imágenes tridimensionales de proteínas y otras moléculas. La subrutina <i>rastep</i> lee un *.pdb con registros anisotrópicos y los representa en tres dimensiones. Es la subrutina de Raster3d que transforma los archivos de coordenadas atómicas en archivos de imágenes *.jpg.
Molscript	http://www.avatar.se/molscript/	Programa para desplegar estructuras moleculares en 3D.

³Escrito en nuestro laboratorio por Paula González-Rubio Garrido.

5 RESULTADOS Y DISCUSIÓN

BÚSQUEDA EN EL PDB

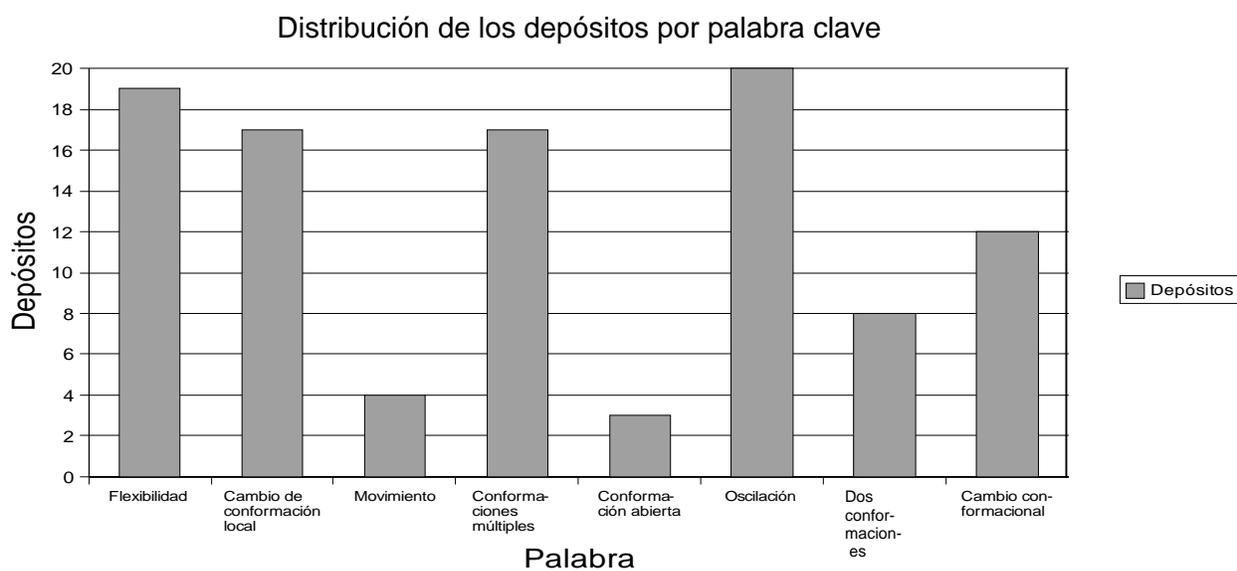
De la búsqueda por palabras clave en el Banco de Datos de Proteínas (PDB) resultaron 1080 proteínas (**Tabla 2**). Para disminuir el tamaño de la muestra y mejorar su calidad, decidimos tomar como primer criterio de exclusión una resolución menor a 2.5 Å. Bajo esta exigencia el tamaño de la muestra se redujo a 634 depósitos un número todavía muy difícil de explorar.

Tabla 2 Resultados de la búsqueda por palabras clave en el PDB

Palabra	Depósitos totales	Resolución $\geq 2.5\text{Å}$
Cambio conformacional	242	139
Flexibilidad	264	84
Cambio de conformación local	84	59
Movimiento	94	62
Conformación abierta	62	39
Oscilación	81	55
Dos conformaciones	191	149
Múltiples conformaciones	62	47
Totales	1080	634

SELECCIÓN DE ESTRUCTURAS

Para reducir más el tamaño de la muestra y trabajar con estructuras de mayor calidad excluimos a aquellas que tuvieran una diferencia mayor de 5 unidades porcentuales entre R y R_{free} y a las que carecieran de factores de estructura depositados, ya que sin estos últimos no es posible realizar ningún tipo de afinamiento. Bajo estos criterios quedaron 160 estructuras pertenecientes a 100 proteínas diferentes.



Gráfica 1 Distribución de los depósitos por palabra clave de búsqueda.

Una exploración de la literatura referida en los depósitos nos mostró que aquellos relacionados con las palabras *conformational change* e *induced fit* eran los que con más frecuencia presentaban los cambios estructurales que deseábamos. Por ello en la etapa de sobreposición nos centramos en las proteínas que incluían dichas palabras en su archivo tipo *.pdb^Ω, aunque sin apartar a las demás.

Del total de proteínas que se superpusieron (42) sólo 15 presentaron más de una conformación indicando cambios conformacionales: DNA-glicosil transferasa, UDP-N-acetilglucosamina enolpiruvato transferasa (EPT), nucleotidasa 5', formato deshidrogenasa, calmodulina, proteína de unión a la caja TATA (TBP), triptofanil tRNA sintetasa, 5-epi-aristoloqueno sintasa, lactoferrina, maltodextrina, toxina de difteria, alcohol deshidrogenasa, UDP-N-acetilmuramoiil-L-alanina:D-glutamato ligasa (MurD), O-acetilserina sulfidrilasa (OAS) y gliceraldehído-3-fosfato deshidrogenasa.

Por distintas razones como presencia de enlaces anormales, en particular puentes disulfuro anómalos, sobreafinamiento, problemas en el formato de los factores de estructura depositados, etc. sólo tres de ellas (EPT, OAS y MurD) pudieron ser afinadas correctamente por el método de tensores TLS en *refmac*. En el caso de MurD nos dimos cuenta de que había un artículo⁸⁴ posterior que demostraba que el cambio de conformación era uno distinto al que estábamos contemplando, o dicho de otro modo, estábamos afinando un artefacto cristalográfico. La estructura que corresponde a la forma abierta se encuentra depositada pero sin factores de estructura, lo que nos obligó a excluirla del trabajo. La molécula SecA resultó de una búsqueda posterior (realizada en junio del 2004) por la palabra *open form* en el PDB.

O-ACETILSERINA SULFIDRILASA

La O-acetilserina sulfidrilasa de *Salmonella typhimurium* (OASS; EC 4.2.99.8) es una enzima homodimérica de 34.4 KDa/subunidad⁸⁵ que cataliza la sustitución de acetato en la cadena lateral de O-acetil-L-serina (OAS) por sulfuro (-SH) para producir L-cisteína⁸⁶ (**Figura 6**). Esta reacción es el último paso en la vía sintética *de novo* de dos pasos para L-cisteína en bacterias entéricas; siendo el primero la reacción de acetilación del β-hidroxilo de la L-serina a partir del donador acetilCoA, catalizada por la transacetilasa de serina⁸⁷.

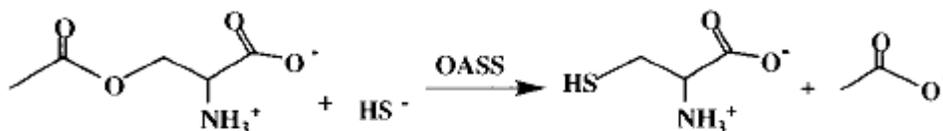


Figura 6 Reacción de sustitución de acetato por sulfuro catalizada por la O-acetilserina sulfidrilasa

^Ω En el encabezado del archivo

Pertenece al grupo de enzimas de remplazamiento β que catalizan la sustitución de un grupo funcional en el carbono β de un aminoácido. Junto a las enzimas que catalizan la eliminación del sustituyente β constituyen un grupo homólogo de proteínas dependientes de piridoxal 5'-fosfato (PLP): la familia β^{88} . Otra enzima del tipo remplazamiento β que ha sido ampliamente estudiada es la triptofano sintetasa (TRPS) también de *S. typhimurium*, la cual ya ha sido afinada por grupos pseudo-rígidos en nuestro grupo⁹. En ella la subunidad β facilita una reacción muy similar a la catalizada por la OASS; el reemplazo del β -hidroxilo de la serina pero en esta ocasión por indol, dando L-triptofano como producto. Aunque sólo tienen un 19% de identidad las secuencias de OASS y TRPS la sobreposición de sus estructuras nos muestra una alta conservación de la topología general.

ESTRUCTURA

La sulfidrilasa de O-acetilserina presenta una estructura de dos dominios. El N-terminal (1-12, 35-145) forma una estructura α/β de cuatro subunidades $\alpha\beta$ ⁸⁹. Esto coloca las α hélices 1, 2 y 3 en una cara de la hoja paralela β de cuatro cadenas (orden 5, 4, 3, 6) y a la α hélice 4 en la otra: la descripción anterior se repite en el C-terminal (14-34, 146-302) en específico en las cadenas 9, 8, 7 y 10 de la hoja β y las hélices 7, 8 y 9 (**Figura 7**), existiendo la posibilidad de un evento de duplicación génica.

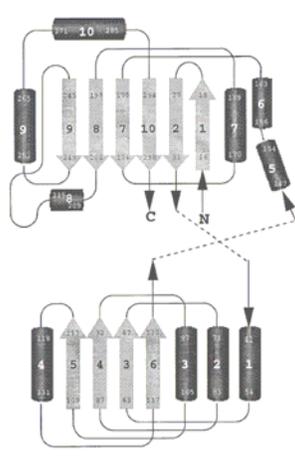


Figura 7 Topología de la enzima O-acetilserina sulfidrilasa. Parte superior C-terminal y parte inferior N-terminal.

La α hélice 1 y las 5 y 6 (separadas por una torcedura en la Pro158) pudieran ser unidades estructurales equivalentes. Los residuos 14-34 que pertenecen al C-terminal, e incluyen un listón β antiparalelo (cadenas 1 y 2), se unen a la hoja β central por medio de la cadena 2 que a su vez está pareada de manera paralela con la cadena 10, por ello es que se forma una hoja de seis cadenas y no de cuatro como en el N-terminal⁹⁰.

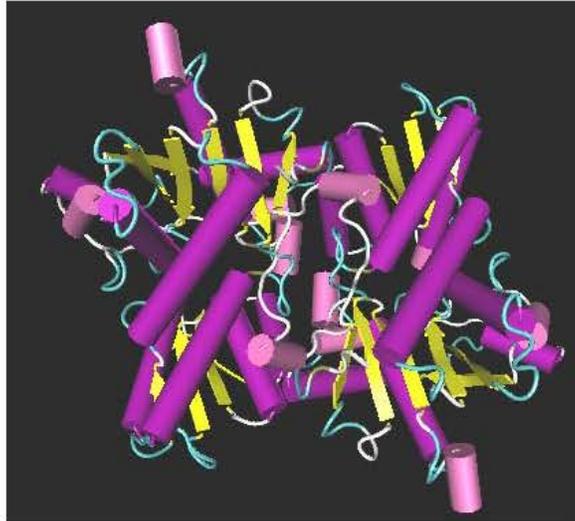


Figura 8 Estructura tridimensional del dímero de O-acetilserina sulfidrilasa.

La interfase dimérica, entre los monómeros, tiene un área de 1350 \AA^2 por cada unidad⁹¹. La Ser1 del N-terminal y el último aminoácido Phe315 pertenecen a ella. El residuo Phe 315 tiene interacciones no polares con las cadenas laterales de Ile97, Arg100, Lys101, y Lys104 de la α hélices 3 de la subunidad adyacente. Lys101, Leu102, Lys104 y Ala105 de la α hélice 3 hacen contactos no polares con Leu306, Phe311 y Leu314 de la otra subunidad. Hay pares iónicos entre Glu 303 y Arg304 de cada subunidad y viceversa. Es decir, en la interfase hay áreas polares y no polares con contribución tanto del amino como del carboxilo terminal involucradas en la estabilización del dímero.

El sitio de unión de piridoxal 5'-fosfato (PLP) se encuentra en la interfase de los dominios N-C terminales de cada monómero. Se forma un enlace covalente tras la reacción mediada por una base de Schiff (o enlace imino) al grupo ϵ -amino de la Lys41, formando la llamada aldimina interna⁹¹. En los alrededores de la misma Lys 41 se encuentra el sitio activo, en la interfase carboxilo-amino terminal.

En la misma interfase dimérica se ha descrito un sitio de unión a Cl^- ⁹¹. El ion Cl^- interacciona con los residuos Asn7 de un monómero y Leu268 del otro. La unión del cloruro en esta región induce cambios conformacionales que define una forma inhibida de la enzima.

ESTRUCTURAS DE OAS EN EL PDB

En el caso de la O-acetilserina sulfidrilasa hay 5 estructuras depositadas en el PDB. Tres de ellas correspondientes a *Salmonella typhimurium*, una a *Homo sapiens* y una a *Thermotoga maritima* (Tabla 6). Como 1jbq y 1o58 son de organismos con un único depósito, en una sola conformación, quedaron descartados.

Tabla 3 Depósitos en el PDB de la molécula O-acetilserina sulfidrilasa.

Clave PDB	Organismo	Grupo espacial	R (%)	R _{free} (%)	Resolución [Å]	Undidad asimétrica	Ligandos
1d6s	<i>S. typhimurium</i>	P 2 ₁ 2 ₁ 2	17.4	22.3	2.3	Dímero	Metionina, PLP*
1fcj	<i>S. typhimurium</i>	P 2 ₁ 2 ₁ 2 ₁	20	24.6	2.0	Tetrámero	SO ₄ , Cl, PLP
1jbq	<i>Homo sapiens</i>	P 3 ₁	25.7	29.6	2.6	Hexámero	Hem* ¹ , PLP
1o58	<i>Thermotoga maritima</i>	P 2 ₁	15.9	19.8	1.8	Tetrámero	PO ₄ * ²
1oas	<i>S. typhimurium</i>	P 2 ₁ 2 ₁ 2 ₁	17.2	20	2.2	Dímero	PLP

PLP*.- Piridoxal-5'-fosfato

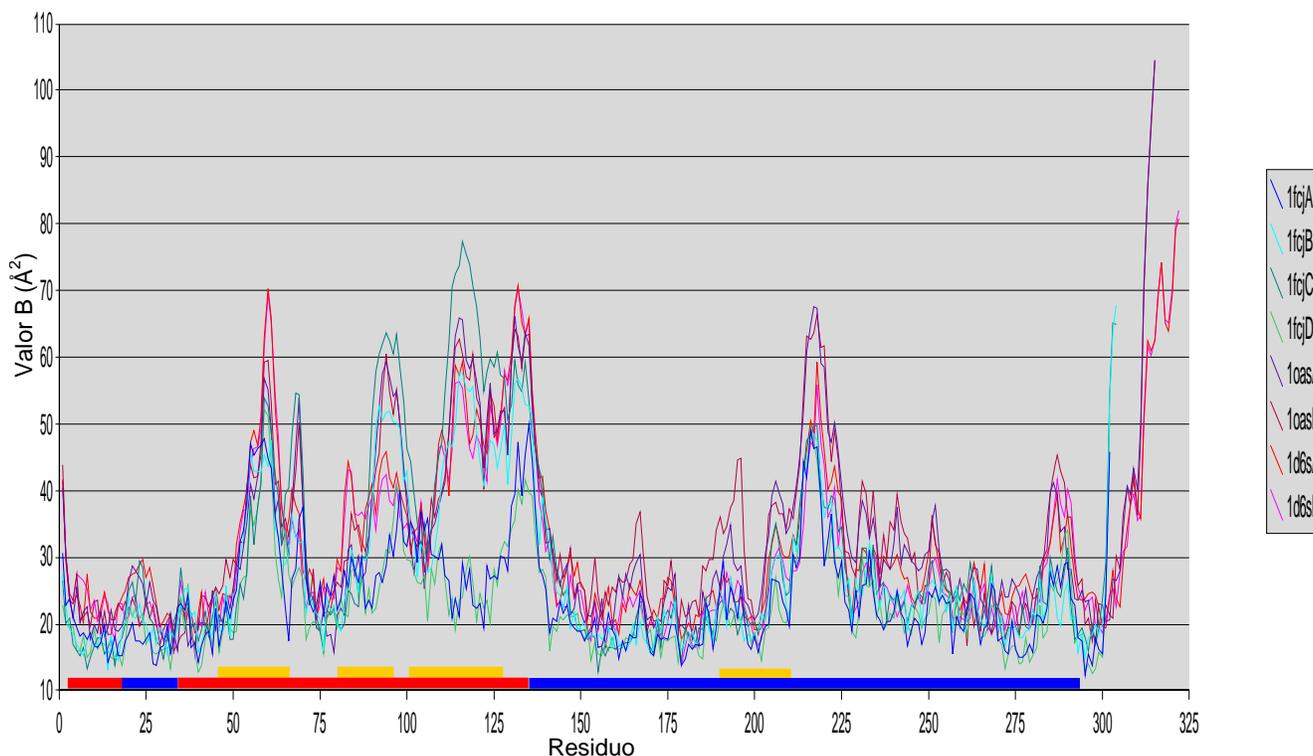
HEM*¹.- Protoporfirina IX-Fe

De las tres restantes la estructura 1d6s es una mutante Lys41Ala en la conformación cerrada, que tiene como ligando en el sitio activo al análogo de sustrato L-metionina. Las estructuras 1fcj y 1oas fueron las seleccionadas para el afinamiento por TLS porque tienen depositados los factores de estructura además de que son reportadas como la forma libre y con un sitio alostérico inhibidor ocupado⁹¹ respectivamente. Las tres tienen unida a la coenzima PLP. Por otra parte, 1fcj presenta un ion cloruro y un sulfato por cada molécula de la unidad asimétrica lo que parece estabilizarla en una conformación distinta a las otras dos. Otra diferencia radica en la unidad asimétrica: la de 1fcj es un tetrámero, mientras que en 1oas se encuentra un dímero.

DISTRIBUCIÓN DE LOS VALORES B

La gráfica de los valores B (**Gráfica 2**), que incluye a cada una de las cadenas del depósito, muestra patrones complicados. Destaca la zona 116-136, en la que los valores de 1fcjA y 1fcjD están por debajo ($B < 30 \text{ \AA}^2$) de los de las demás cadenas ($B > 45 \text{ \AA}^2$). Sin embargo, las cadenas A de 1fcj y D de 1fcj no corresponden a la forma cerrada (1d6s) sino a la conformación con el sitio alostérico ocupado. La diferencia entre cadenas nos hace suponer que esas regiones se estabilizan por la unión del ligando al sitio alostérico o por contactos entre cadenas y/o entre unidades asimétricas que reducen el valor de los factores de temperatura al estabilizar los residuos incluidos en dichos intervalos de la secuencia. A excepción de la región ubicada alrededor del residuo 90, las dos cadenas correspondientes a la forma cerrada (1d6sA y 1d6sB) tienen un patrón que, prácticamente, se superpone con el de las cadenas 1fcjB, 1fcjC, así como con las cadenas A y B del depósito 1oas.

Distribución de valores de temperatura



Gráfica 2 Distribución de valores B en las estructuras 1d6s, 1fcj y 1oas. Las líneas de color rojo en la base representan los residuos del N-terminal y en azul a los del C-terminal. Las franjas anaranjadas indican las regiones de mayor flexibilidad.

Si observamos las zonas con valores B mayores de 50 en la gráfica (marcadas en rojo en el eje X) y en la representación tridimensional (**Figura 9 a**) podemos ver claramente que tres de las cuatro regiones (residuos 50-70, 87-103, 115-140) están en la región N-terminal, que es la zona del cambio conformacional (ver **Cambio conformacional**). La región de los residuos 212-226 es un asa del C-terminal que se encuentra libre de contactos con el núcleo de la estructura, es decir en la periferia. Los residuos 1-49, todos ellos con valores menores de 30 Å², se encuentran en la interfase del dímero que constituye la unidad biológica.

La **Figura 9 b**) muestra que la región 115-140 de la cadena D, en específico la zona de los residuos 115-119 está a distancia de puentes de hidrógeno con los residuos 1-6 de la cadena C (ambas en rojo), lo que podría estabilizar la zona que en las demás cadenas es de alta flexibilidad y por lo tanto explicar la reducción en los Bs de dicha región. Si observamos con detalle, esos contactos mantienen unidos a los dos dímeros (A-C y D-B) que forman el tetrámero de la unidad asimétrica. Para explicar los valores bajos en la misma región de la cadena A tenemos dos posibilidades. Una es reproducir el cristal en el programa O, mediante la aplicación de los operadores simétricos (grupo espacial), buscar contactos cristalinos entre moléculas vecinas, pues en el tetrámero parece que la región está libre de contactos. Y la otra es la propuesta por Burkhard,

et al.⁹¹: suponer que sólo la cadena A está estabilizada por el inhibidor alostérico en una conformación inactiva que entre otras características ha disminuido las oscilaciones en las regiones implicadas en el cambio de conformación. Respecto a la primera, el residuos 115 y 116 de la cadena A están a 4 Å de los residuos 165 y 166 de la cadena D del vecino cristalográfico (**Figura 9 c**), lo que complica el que la estabilización de esa parte del dominio sea sólo a causa de contactos cristalográficos.

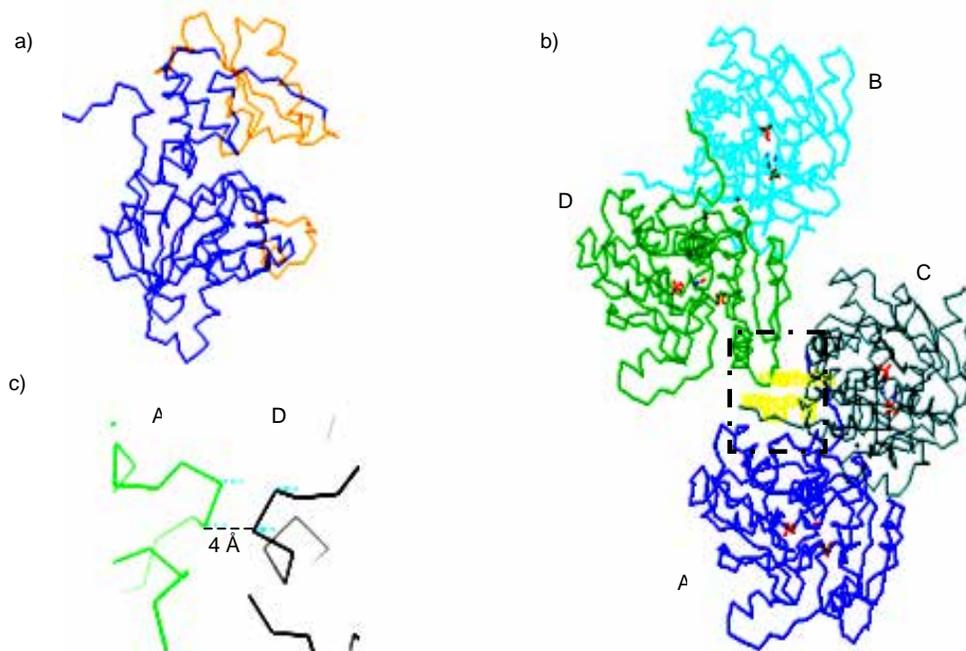


Figura 9 a) Zonas de mayor flexibilidad (anaranjado) de acuerdo a la distribución de valores B. b) Contactos entre las cadenas A y D del tetrámero de la unidad asimétrica. c) Vecino (negro) cristalográfico de la cadena A (verde).

SOBREPOSICIÓN DE ESTRUCTURAS

La sobreposición de las estructuras 1fcj-1d6s dio como resultado una desviación cuadrática media (r.m.s.d) de 0.31 Å en el carboxilo terminal (residuos 13-34, 146-302) y de 1.43 Å en el dominio N-terminal (residuos 1-12, 35-145). La sobreposición de 1oas-1d6s resultó en una desviación de 0.42 Å y 1.26 Å en el C-terminal y N-terminal respectivamente. Para observar las diferencias entre la conformación libre de ligandos y la de sitio alostérico ocupado superpusimos las cadenas A de 1oas-1fcj. Obtuvimos una r.m.s.d. =0.27 Å en la porción carboxilo terminal y r.m.s.d.=1.21 Å en la amino terminal, menores a las anteriores.

La diferencia entre 1fcj-1oas explicaría que la ocupación del sitio alostérico por un ion cloruro estabiliza a la molécula en una conformación intermedia entre la conformación libre (1oas) y la cerrada (1d6s). Sin embargo, esta diferencia ocurre sólo con la cadena A del tetrámero en 1fcj. La sobreposición de la cadena A con las cadenas B, C y D de 1fcj da como resultado en el N-terminal A-

B= 1.17 Å, A-C= 1.11 Å, A-D= 1.11 Å mientras que entre las demás cadenas: B-C= 0.39 Å, B-D= 0.31 Å y C-D= 0.45 Å. Es decir la cadena A está en una conformación distinta a B, C y D quienes entre sí son más parecidas. De igual forma la sobreposición de la cadena A de 1fcj con las cadenas A y B de 1oas resulta en una desviación media cuadrada de 0.25 Å en el C-terminal y de 1.08 Å en el N-terminal. La sobreposición de las cadenas B, C y D sobre 1oas resultó en valores alrededor de 0.33 Å (B-oasA= 0.33 Å, C-oasA= 0.36 Å y D-oas= 0.26 Å) en el C-terminal y menores que 0.5 Å (B-oasA= 0.38 Å, C-oasA= 0.49 Å y D-oas= 0.30 Å) en el N-terminal, apoyando, junto con la ausencia de contactos cristalográficos, la idea de que sólo la cadena A se encuentra en una conformación distinta. La zona de mayor diferencia en la sobreposición de las estructuras coincide con la que le habíamos atribuido contactos cristalinos con la cadena D, confirmando nuestra hipótesis de que la reducción en los Bs de esa región es producto de los contactos entre los monómeros del tetrámero. Por lo demás, las superposiciones de las cadenas B, C y D son congruentes (menores a 0.5 Å) con la conformación de 1oas por lo que concluimos que ambas estructuras están en la forma abierta o libre de ligandos y que la presencia de iones Cl^- o SO_4^{2-} no modifican la conformación.

CAMBIO CONFORMACIONAL

El cambio conformacional ha sido descrito⁹² a partir de la estructura 1d6s, en la que el análogo de sustrato L-metionina imita el papel de la O-acetilserina durante la catálisis. La presencia de la L-metionina causa la formación de una aldimina externa (interna en ausencia del análogo) y la rotación de 13° del cofactor PLP. Este movimiento desplaza a la unión aldimina hacia el exterior, en específico hacia la entrada del sitio activo. El grupo α -carboxilato de la L-metionina forma parte de una red de puentes de hidrógeno con los residuos Pro67 - Thr68 - Asn69 -Gly70 del asa de asparagina (α -hélice 2 del N-terminal). De particular interés es la interacción electrostática entre el nitrógeno de la amida del Asn69 y el carboxilato de la metionina y con el oxígeno O3' del cofactor. En esta posición, el nitrógeno está a 7 Å de su conformación libre.

La unión del α -carboxilato al asa de asparagina y el reacomodo de Asn69 tienen una gran influencia sobre la topología general de la proteína. El residuo Met95, adyacente a Asn69 en la conformación abierta, sigue el movimiento de la cadena lateral de Asn69, desplazamiento que coloca al azufre a 6.5 Å de su ubicación anterior. El subdominio N-terminal (87-131) se mueve de manera conjunta como un cuerpo rígido: rota unos 13° hasta crear una red de puentes de hidrógeno con el grupo α -carboxilato de la L-metionina. Esta rotación no altera la topología general del dominio N-terminal. A excepción del cambio en el asa de asparagina, la reestructuración es un movimiento del subdominio 87-131 del N-terminal sin cambios localizados en ninguno de los dos dominios. El arreglo del subdominio acerca a varios residuos apartados en la conformación abierta. Se forman nuevos puentes de hidrógeno y se establecen interacciones hidrofóbicas entre el C-terminal y el subdominio reposicionado. La cadena lateral de la L-metionina forma interacciones no polares con sus

contrapartes de Asn69 y Met119. Ahora, el sustrato se encuentra oculto al interior de la proteína y el sitio activo se vuelve inaccesible para moléculas grandes. Queda sólo un pequeño canal por el que pueden pasar moléculas pequeñas como sulfato (el segundo sustrato) o acetato (el primer producto de la reacción).

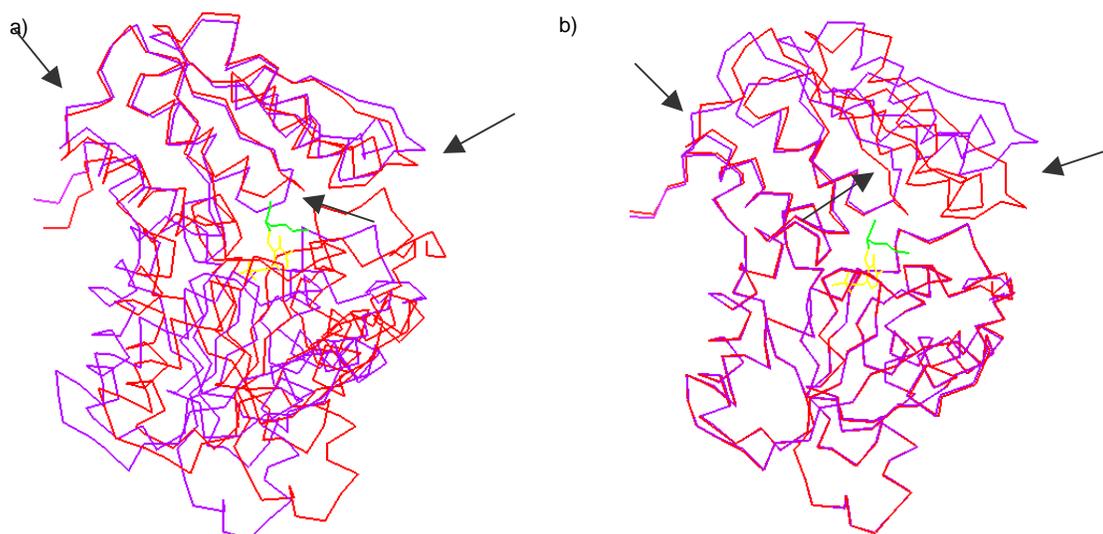


Figura 10 Sobreposición de las estructuras 1oas-1d6s en las regiones a) N-terminal y b) C-terminal. Las flechas indican las zonas de mayor modificación. Parte superior izquierda: bisagra del N-terminal y a la derecha las asas del mismo dominio que circundan el sitio catalítico. En verde la cisteína y en amarillo el piridoxal fosfato.

SELECCIÓN DE GRUPOS PSEUDO-RÍGIDOS PARA EL AFINAMIENTO TLS

Para ambas estructuras tomamos como control el considerar a todas las cadenas (A, B, C y D de 1fcj y A y B de 1oas) como un solo grupo, selección a la que llamamos **1G**. Otra selección consistió en considerar a cada cadena como un grupo pseudo-rígido; **4G** en el caso de 1fcj y **2G** en el caso de 1oas. También tomamos los dominios estructurales N (residuos 1-14, 34-156) y C-terminal (residuos 15-33, 157-302) de cada cadena como un cuerpo pseudo-rígido, lo que en 1fcj produjo 8 grupos y en 1oas 4 grupos, selección que recibió el nombre de $\frac{1}{2}$ (**Figura 11 a**). Las dos selecciones restantes de pseudo-rígidos fueron las resultantes del programa *Dyndom* para las estructuras 1oas y 1fcj. En el caso de 1oas el dominio fijo se compone de los residuos (1-57, 74-77, 142-302), las bisagras por los residuos (58-66, 68-73, 79-89, 139-141) y el dominio móvil por los residuos (67, 90-138). Esta selección recibió el nombre de **DDoas** (**Figura 11 b**). Siguiendo el mismo orden los dominios del depósito 1fcj se compusieron por los residuos (1-58, 144-302), (59-64, 67-88, 142-143) y (65-66, 89-141), selección identificada como **DDfcj** (**Figura 11 c**).

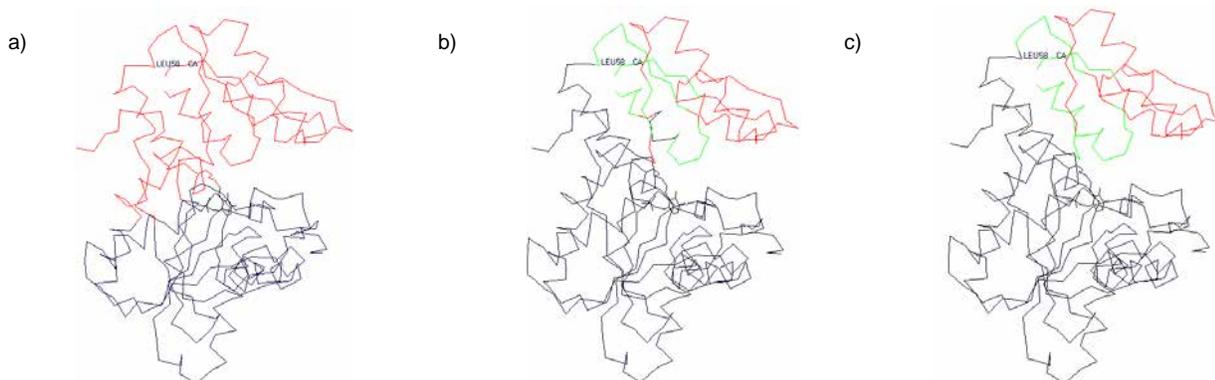


Figura 11 Selección de grupos realizada para la O-acetilserina sulfidrilasa. a) ½, b) DDoas y c) DDfcj.

AFINAMIENTO TLS

Lo primero que observamos de los resultados de los afinamientos de 5 ciclos (**Tablas 7a) y 7b)** es que se reduce más el error en la estructura 1oas que en la 1fcj. La estructura depositada con clave 1oas reporta valores de $R = 17.2\%$ y $R_{free} = 20.0\%$ que en la mejor de las selecciones (DD) se reduce hasta poco más de 3 puntos porcentuales tanto en R como R_{free} . En cambio para la depositada con clave 1fcj $R = 20\%$ y $R_{free} = 24.6\%$ el mejor resultado (8G 1/2) muestra una reducción de poco más de 1 unidad porcentual en ambos valores. El mejor resultado es 8G ½ porque aunque tiene una R ligeramente mayor que DD, en R_{free} y FOM tiene mejores porcentajes. En los dos casos tras el afinamiento isotrópico se reducen los valores originales. En 1fcj la diferencia en R entre el isotrópico y la mejor selección es de apenas 0.69 unidades porcentuales, mientras que en 1oas alcanza las 2.07 unidades porcentuales. Para R_{free} ocurre lo mismo: en 1fcj la diferencia es menor de 1 unidad porcentual y en 1oas de poco más de 2 unidades porcentuales.

Tablas Afinamientos isotrópico, TLS y anisotrópico; 5 ciclos de TLS por 5 de máxima similitud.

a) Estructura 1fcj

Grupos	R (%)	R_{free} (%)	FOM (%)
Isotrópico	19.63	24.28	82.51
1G	19.62	24.22	82.46
4G	19.03	23.27	83.50
DDfcj	18.92	23.34	83.50
1/2	18.94	23.19	83.58

b) Estructura 1oas

Grupos	R (%)	R_{free} (%)	FOM (%)
Isotrópico	16.21	19.21	83.23
1G	14.24	17.17	85.74
2G	14.25	17.15	85.87
DDoas	14.14	16.91	86.06
1/2	14.22	17.01	85.90

La convergencia del afinamiento se alcanzó a los 10 ciclos (**Tablas 8a) y 8b)**. En la estructura 1fcj la selección en la que tomamos como grupos a cada dominio N y C-terminal de cada cadena tiene mejores valores de R que DD, pero no de R_{free} ni de FOM. Por eso ahora consideramos que la mejor representante es DD. En 1oas las tendencias se mantienen: un grupo aparte del isotrópico con valores considerablemente menores en las R s y considerablemente mayores en FOM, con DD como la mejor selección con los mejores porcentajes en los tres estadísticos.

Tablas 5 Afinamientos isotrópico, TLS y anisotrópico; 10 ciclos de TLS por 10 de máxima similitud.

a) Estructura 1fcj

Grupos	R (%)	R _{free} (%)	FOM(%)
Isotrópico	19.52	24.37	82.44
1G	19.39	24.00	82.70
4G	18.80	23.15	83.69
DD	18.72	23.06	83.80
1/2	18.67	23.15	83.78

b) Estructura 1oas

Grupos	R (%)	R _{free} (%)	FOM(%)
Isotrópico	16.17	19.37	82.94
1G	14.17	17.06	85.90
2G	14.14	17.11	85.98
DD	13.98	16.90	86.28
4G 1/2	14.10	16.97	86.03

Con la intención de analizar los resultados de los afinamientos y encontrar algunas diferencias significativas en las estructuras representamos en tres dimensiones cada modelo con sus elipsoides térmicos (**Figura 12**).

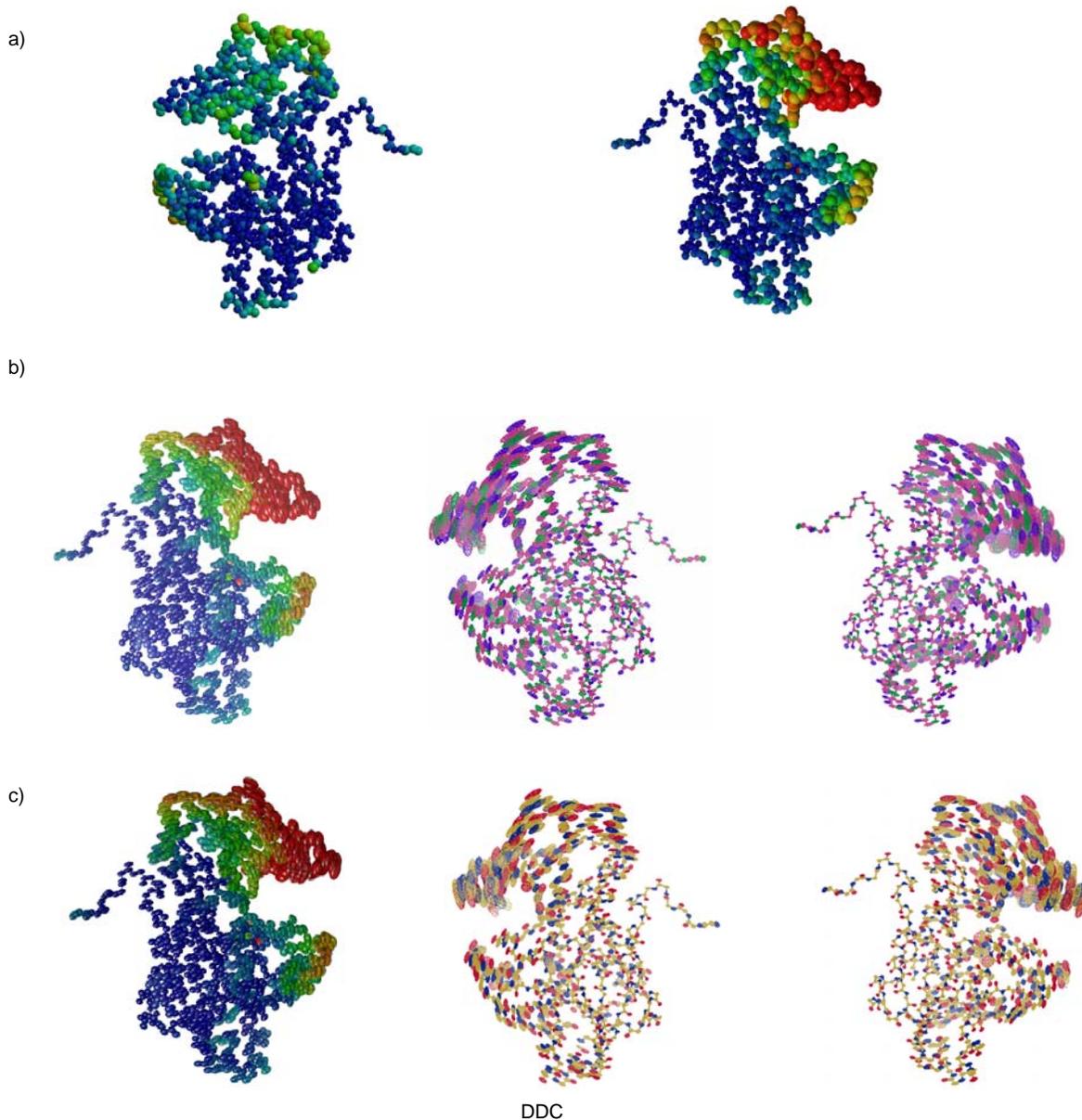


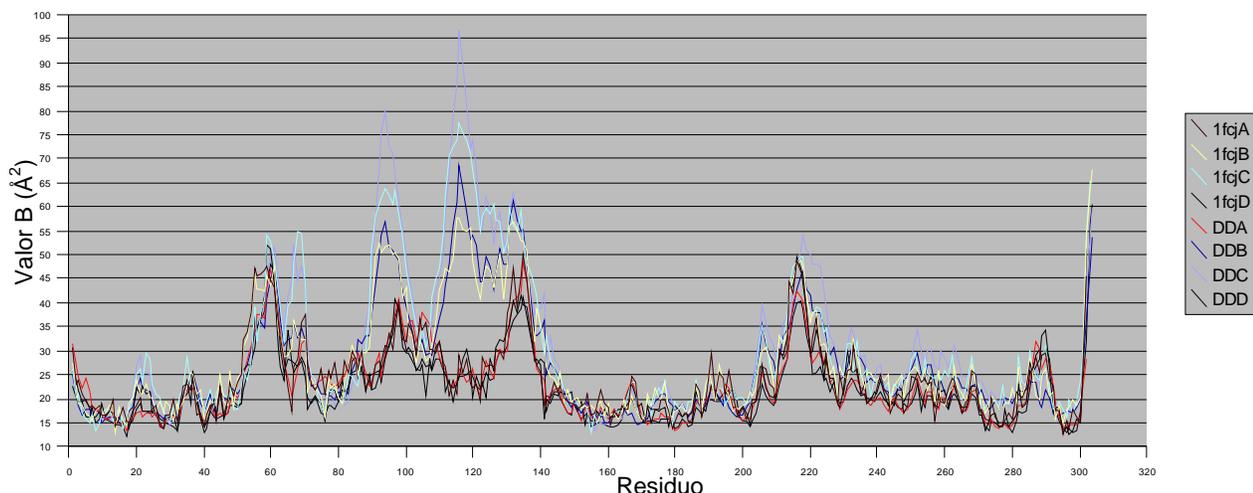
Figura 12 Representación de los elipsoides anisotrópicos producto del afinamiento por tensores TLS, en tres selecciones de grupos distintas del depósito 1fcj. A) Afinamiento isotrópico, b) afinamiento TLS con la selección $\frac{1}{2}$ y c) afinamiento TLS con la selección DD.

La representación de los elipsoides es consecuente con la distribución de los valores B. Es evidente que la zona de más flexibilidad se encuentra en el N-terminal y que las cadenas B y C presentan Bs mayores que las cadenas A y D en la misma región. Además los ejes mayores de los elipsoides en color rojo y amarillo del N-terminal parecen desplazarse conjuntamente hacia el C-terminal, movimiento que tiende a juntar los dominios cerrando la interfase.

En las cadenas A y D (Figuras no mostradas) de la estructura 1fcj el movimiento es menor debido a la acción del inhibidor en la primera y a los contactos con el N-terminal de la cadena C.

Cuando graficamos el patrón de distribución de Bs del depósito original junto a los de la mejor selección (DD) (**Gráfica 3**) observamos un ligero aumento en los valores de B en los residuos 92-97 y 114-120 de las cadenas B (3-5 Å² promedio; 6 en el 94 Å² y 11 Å² en el residuo 116) y C (10 Å² en promedio; 16 Å² en el residuo 94 y 19 Å² en el residuo 116). Esto quiere decir que la contribución de los grupos en que se encuentran esos residuos es mayor en esas zonas que en las demás, en las que prácticamente no hay cambio. O lo que es lo mismo, la componente oscilatoria se concentra en estas regiones.

Distribución de valores B tras el afinamiento por tensores TLS (1fcj)



Gráfica 3 Distribución de los Bs en el depósito original y tras el afinamiento TLS con la selección DD.

En el caso del depósito 1oas se mantienen los mismos patrones de oscilación. La interfase dimérica, junto con la zona interna de la proteína, es la región que presenta los valores térmicos más bajos. Las zonas de mayor vibración, en la representación tridimensional de los elipsoides (**Figura 13**) es la de los residuos 115-140 y 214-222, junto con la región 52-68 topológicamente adyacente a la primera. También se mantiene la direccionalidad de los ejes mayores de los elipsoides que va desde uno de los dominios estructurales hacia el otro. Esa dirección de los elipsoides tiende a aproximar a los dominios N y C-terminal y cerrar la interfase, posicionando a los aminoácidos involucrados en la unión de ligandos y en la catálisis. Las flechas en la **Figura 13** muestran la

dirección del movimiento de los ejes mayores de los elipsoides en la cadena A de la selección de pseudo-rígidos DD.

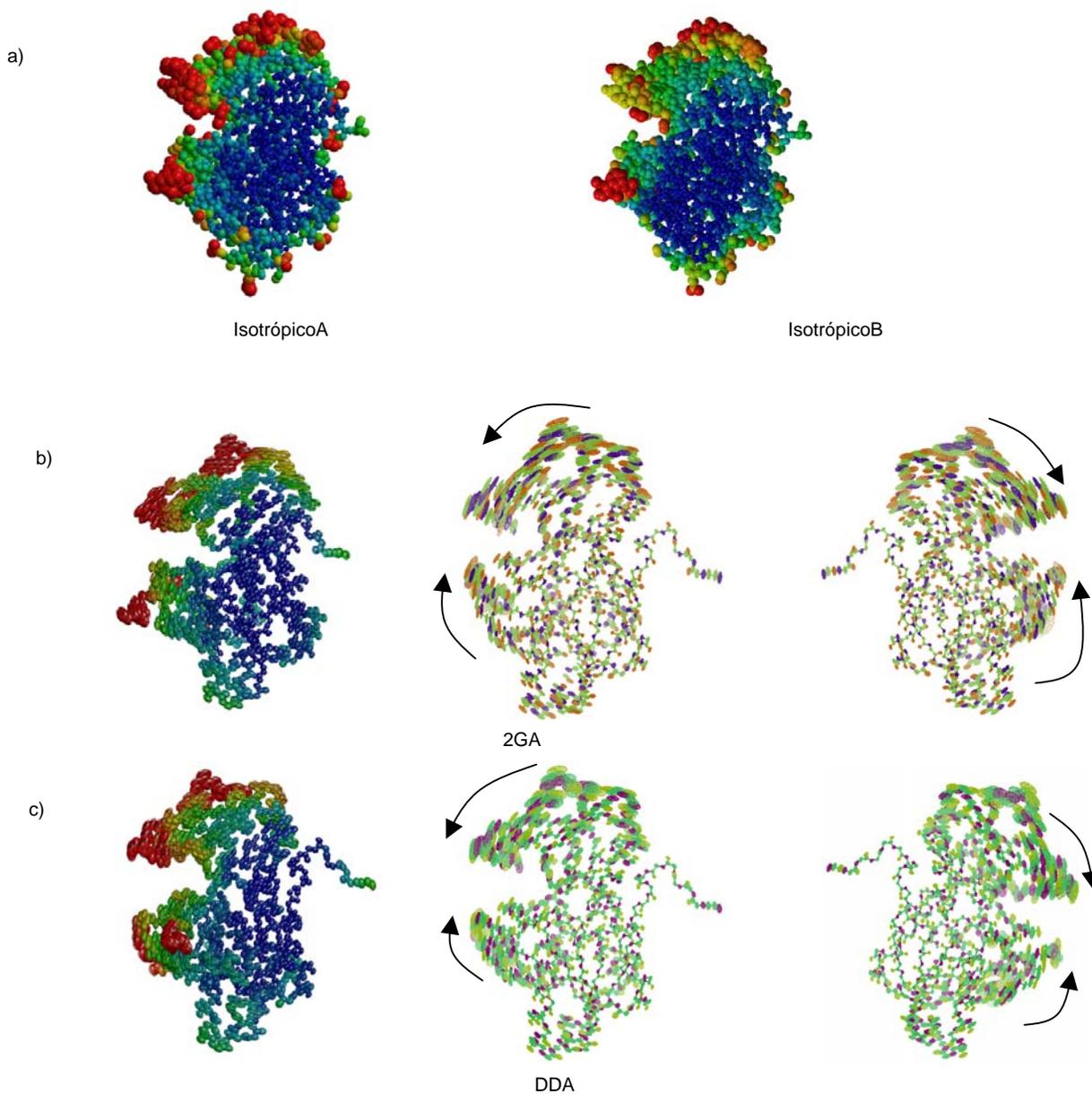
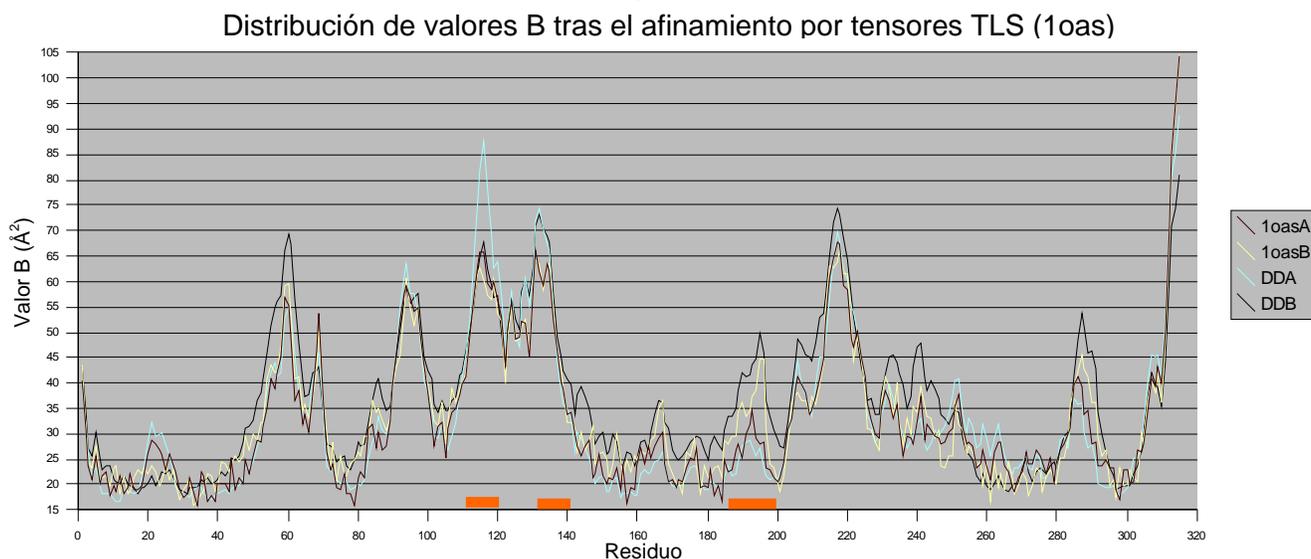


Figura 13 Representación de los elipsoides anisotrópicos producto del afinamiento por tensores TLS del depósito 1oas: a) afinamiento isotrópico, b) afinamiento TLS de la selección $\frac{1}{2}$ y c) afinamiento TLS de la selección DD.

La contribución de los cuerpos pseudo-rígidos se manifiesta en aumentos significativos de los valores de temperatura en algunas regiones de la proteína. En especial la región 113-120 de la cadena A muestra un aumento de 10 \AA^2 en promedio y de 22 \AA^2 en el residuo 116. En el residuo 60 también hay un incremento de 10 \AA^2 , mientras que en el 192 de la cadena B es de 8 \AA^2 . En los residuos 136-139 el máximo cambio se observa en el aminoácido 134 con un valor de B 13 \AA^2 mayor

al original (**Gráfica 4**). Excepto el que ocurre en un asa del C-terminal, residuos 186-201, todos los cambios ocurren en la zona del cambio conformacional.



Gráfica 4 Distribución de los Bs en el depósito original y tras el afinamiento TLS con la selección DD. En color naranja, sobre el eje x, las zonas con mayor diferencia.

También en este modelo podemos decir que los datos cristalográficos experimentales contienen información oscilatoria relacionada con el cambio conformacional y en consecuencia con la actividad enzimática.

La representación de los elipsoides muestra que el dominio N-terminal de las estructuras libres oscila en dirección del cambio conformacional, tendiendo a reducir la cavidad interdominios. La comparación con la estructura cerrada hace suponer que la presencia de la L-metionina establece interacciones con residuos del sitio activo; en específico una red de puentes de hidrógeno con los residuos Pro67-Thr68-Asn69-Gly70 del asa de asparagina que estabilizan la conformación cerrada. El subdominio N-terminal (87-131) se mueve de manera conjunta como un cuerpo rígido: rota unos 13° hasta crear una red de puentes de hidrógeno con el grupo α -carboxilato de la L-metionina. En concreto la nueva posición del subdominio N-terminal (87-131) ha rotado como un cuerpo rígido cerca de 13° en comparación con estructura nativa. La importancia del reacomodo radica en que el nuevo arreglo del subdominio acerca a residuos involucrados en la catálisis y unión de ligandos que en la conformación abierta estarían muy apartados. Por último se establecen interacciones hidrofóbicas (el inhibidor L-metionina con Asn69 y Met119) entre el C-terminal y el N-terminal que cierran la cavidad a la entrada de moléculas grandes, quedando sólo un pequeño canal por el que pueden pasar moléculas pequeñas como sulfato (el segundo sustrato) o acetato (el primer producto de la reacción).

TRANSPORTADOR SECA

Cerca del 30 por ciento del proteoma de cualquier organismo está destinado a membrana y polipéptidos secretados. Hasta la fecha han sido descritas más de siete rutas distintas para la exportación de proteínas tan sólo en el dominio bacteriano. De ellas, la vía Sec (por secreción) es esencial, ubicua y responsable de la biogénesis tanto de membrana como de pared celular, secreción de polipéptidos (p. ej. enzimas hidrolíticas), toxinas, moléculas señalizadoras y apéndices de anclaje y de movimiento^x. La exportación de proteínas secretorias es un proceso postraducciona de tres etapas. Comienza con la unión de la pre-proteína a la chaperona de exportación SecB⁹³, posteriormente SecA actúa como receptor reconociendo los dominios líder y maduro de la preproteína⁹⁴ y epítopes de la chaperona SecB⁹⁵. El complejo formado se une a su vez a la translocasa, cuyos principales componentes son las proteínas transmembranales SecY, SecE y SecG quienes forman el canal conductor o poro SecY, así como con la chaperona SecB y la ATP_{asa} SecA⁹⁶. La unión (segunda etapa) activa la función de ATPasa de SecA⁹⁷. La energía de unión ATP-SecA conduce la translocación de un segmento corto de la pre-proteína al interior del canal mientras que la hidrólisis de ATP libera a SecA del complejo, dando paso a la fase de translocación mediada por la fuerza protón-motriz $\Delta\mu_{H^+}$ membranar. La repetición de este ciclo concluye el proceso⁹⁸. La última etapa involucra la remoción del péptido señal por peptidasas líder y el plegamiento adecuado consecuente de la cadena en el exterior de la membrana.

Debido a que la estructura del canal SecY⁹⁹ de *E. coli* sugiere que no puede unirse a SecA, las explicaciones alternativas, como la participación de otras proteínas, habían tomado fuerza hasta que distintas conformaciones de la ATPasa combinados con datos experimentales han apoyado la formación del complejo. Lo que queda claro ahora, es que SecA sufre varios cambios de conformación acoplados a la interacción con sus ligandos y que el movimiento de dominios conducido por la actividad de ATP_{asa} y la interacción con ligandos son necesarios para la exportación del polipéptido al periplasma¹⁰⁰.

ESTRUCTURA

El núcleo de SecA está formado por un par de plegamientos de unión a nucleótidos (NBF, del inglés *nucleotide binding fold*) del tipo F1 nombrados NBF1 (14-223, 353-396) y NBF2 (397-572). Entre las dos subunidades NBF, en la interfase, se encuentra el sitio de alta afinidad para ATP. La función que desempeña el ATP en esta molécula parece ser el adecuado empaquetamiento o posicionamiento de los dominios NBF1 y NBF2. Uno con respecto al otro tienen el mismo plegamiento a excepción de las sustituciones de un meandro en la subunidad NBF2 por el α hélice que conecta las cadenas 4 y 5 de la hoja β central y de la cadena 7 paralela de la misma hoja β por una antiparalela. Las láminas centrales, en ambos plegamientos, tienen el patrón característico de una interdigitación de cadenas β común en ATPasas (por ejemplo F1 y RecA). En NBF1 la secuencia

consenso Walker A¹⁰¹ forma la estructura asa P-hélice que une los fosfatos α y β del ATP, mientras que la secuencia Walker B forma una cadena β hidrofóbica que concluye en Asp207 que hace contactos con el cofactor Mg^{2+} . El residuo Glu208 es el principal candidato a ser el residuo catalítico en la hidrólisis del fosfato γ del ATP. Los equivalentes topológicos a las secuencias Walker A y B en NBFII contactan con la ribosa y el fosfato γ del ATP.

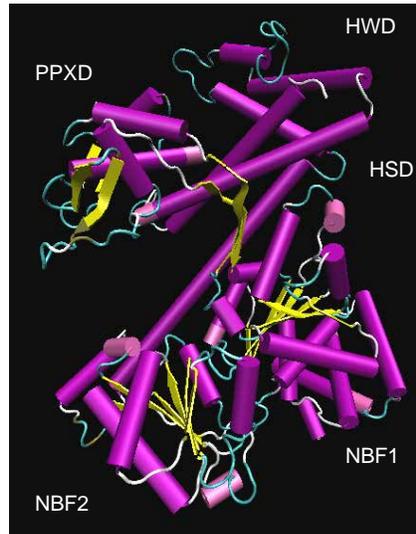


Figura 14 Estructura del transportador SecA y sus cinco dominios señalados.

En el otro extremo topológico hay dos dominios que participan en el entrecruzamiento de SecA con las apoproteínas. Han sido denominados residuos PPXDC y PPXDN (del inglés *preprotein cross-linking domain*, la C y N finales corresponden respectivamente a los carboxilo y amino terminales). Ambos pertenecen al segmento lineal insertado entre las cadenas cinco y seis de la hoja β central de NBF1. El carboxilo terminal se estructura en dos dominios α helicoidales (*alpha helical*). Uno de ellos formado por tres α hélices antiparalelas, con parecido a algunos *coiled coils* α helicoidales, recibe el nombre de HSD (*alpha helical scaffold domain*). Entre las hélices 1 y 2 y unido débilmente a HSD se inserta la subunidad de plegamiento HWD (*alpha helical wing domain*), juntas forman un fragmento de unos 30 KDa resistente a proteasas sólo cuando forman complejo con SecYEG¹⁰². Los últimos 22 residuos del carboxilo terminal, rico en cisteína, son un dominio de unión a Zn que interactúa tanto con fosfolípidos como con la chaperona SecB¹⁰³. En las interfases PPXD-HSD/HWD y NBF1-NBFII-HSD se encuentran dos sitios potenciales de unión a fosfolípidos y pre-proteína.

ESTRUCTURAS DE SECÁ DEPOSITADAS EN EL PDB

Hay 6 estructuras completas del transportador SecA[®] en el PDB. Destaca la de clave 1tf5 por estar depositada como la forma abierta de SecA e incluir factores de estructura y por supuesto por la alta resolución en comparación con las otras estructuras. El depósito 1m74 está en la misma conformación que 1m6n. Seleccionamos esta última porque tiene mayor resolución y valores similares de R y R_{free}. Lo mismo ocurre con 1tf2 que se encuentra en la misma conformación que 1tf5 pero, aunque con R menor, la diferencia entre R y R_{free} es de 6.4 puntos porcentuales, mayor a los 5 que establecimos como criterio de exclusión. Por ello utilizamos a 1tf5 para el afinamiento posterior. Es importante mencionar que los depósitos 1m6n y 1tf5 corresponden a conformaciones distintas y ambas tienen depositados los factores de estructura. La molécula 1tf5 está depositada como forma abierta, por lo que debemos suponer que la conformación de 1m6n es la equivalente a la forma cerrada. Sin embargo la única diferencia entre 1tf5 y 1m6n son los iones sulfato en la segunda. La presencia del ADP no parece estabilizar ninguna de las dos conformaciones, pues 1m74 es muy parecida a 1m6n (r.m.s.d= 0.5 Å) y 1tf2 a 1tf5 (r.m.s.d= 0.28 Å). En el caso de los depósitos de *M. tuberculosis*, además de ser la misma conformación ninguna de las dos moléculas (1nkt y 1nl3) tienen depositados factores de estructura por lo que quedaron descartadas.

Tabla 6 Depósitos en el PDB del Transportador SecA.

Clave PDB	Organismo	Grupo espacial	R (%)	R free (%)	Resolución [Å]	Undidad asimétrica	Ligandos
1m6n	<i>Bacillus subtilis</i>	P 3 ₁ 1 2	22.0	30.1	2.7	Monómero	SO ₄
1m74	<i>Bacillus subtilis</i>	P 3 ₁ 1 2	21.7	29.1	3.0	Monómero	SO ₄ , Mg, ADP
1nkt	<i>Mycobacterium tuberculosis</i>	P 6 ₂ 2 2	21.6	26.5	2.6	Dímero	Mg, ADP
1nl3	<i>Mycobacterium tuberculosis</i>	P 6 ₂ 2 2	19.6	25.9	2.8	Dímero	-----
1tf2	<i>Bacillus subtilis</i>	C 1 2 1	22.8	29.2	2.9	Monómero	Mg, ADP
1tf5	<i>Bacillus subtilis</i>	C 1 2 1	24.1	27.3	2.18	Monómero	-----

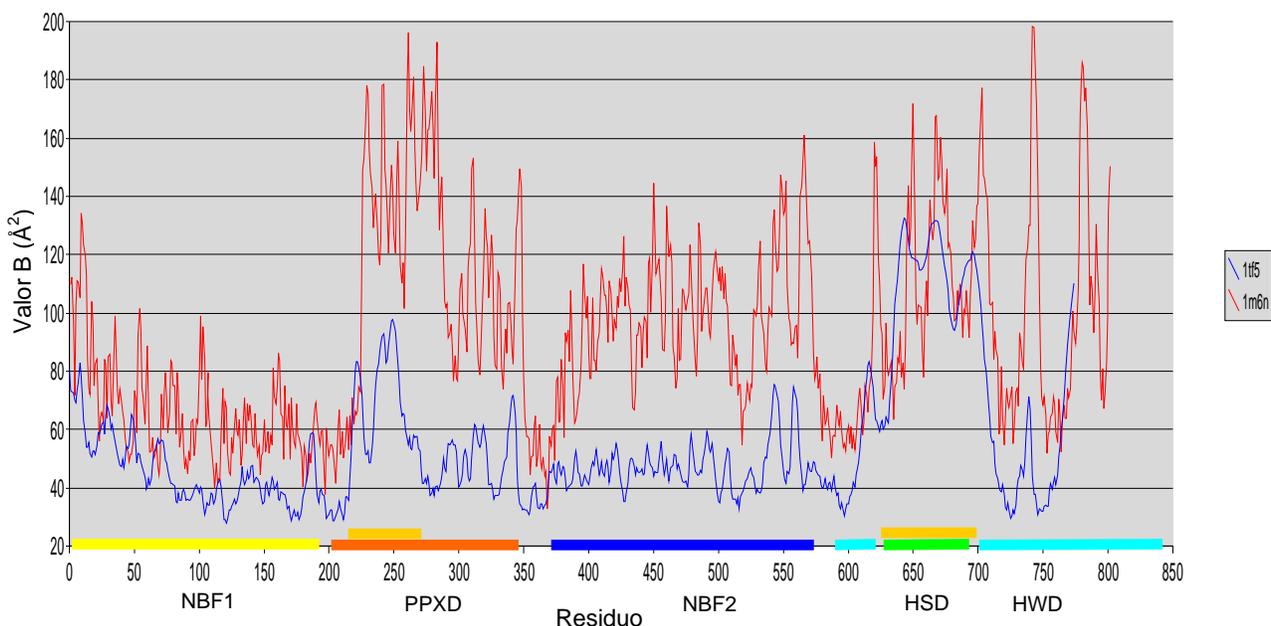
DISTRIBUCIÓN DE LOS VALORES B

La primera observación que surge de la gráfica de distribución de valores térmicos de los depósitos de SecA es que la estructura 1m6n tiene valores de B mucho mayores que los de 1tf5, la reportada como forma abierta. Por ello no podemos considerar a 1m6n como la forma cerrada de la estructura, sino que nuestra teoría indicaría lo contrario: la molécula que más oscila (1m6n) es la forma abierta, mientras que la que tiene valores menores de B es la cerrada. Las zonas de mayor flexibilidad en 1m6n se encuentran en cuatro regiones: 235-290, 300-349, 370-576 y 610-736, mientras que en 1tf5 en dos: una que se superpone parcialmente con la primera de 1m6n (235-266) y la otra que coincide completamente con la zona C-terminal más flexible de la misma estructura. En 1tf5 la región 370-576, que coincide con el dominio estructural NBF2, parece haberse estabilizado si

[®] Otros cuatro depósitos son fragmentos

la comparamos con el patrón de 1m6n. Puede ocurrir lo mismo en la zona 275-349 en la que los valores de B rondan los 50 \AA^2 , valor considerablemente menor a los presentes en las zonas más flexibles, en las que se alcanzan valores mayores de 80 \AA^2 que, en 1m6n, se mantienen por encima de 100 \AA^2 (**Gráfica 5**).

Distribución de factores B en las dos conformaciones de SecA



Gráfica 5 Distribución de valores B en las dos conformaciones de SecA. Las marcas naranjas sobre el eje X indican las zonas de mayor flexibilidad. Los colores amarillo, azul verde y azul claro representan los dominios que se representan en el plegamiento en la figura 15 b).

Dadas las diferencias en resolución de la estructura 1tf5 (2.18) y 1m6n (2.7) y la diferencia de 8.1 puntos porcentuales entre los valores de R y R_{free} del segundo es posible que la diferencia en los valores térmicos sea un problema estadístico. Esto se confirma cuando revisamos la proporción datos/parámetros en cada archivo. El modelo de 1m6n tiene una proporción datos/parámetros de 1.02; con 32,620 reflexiones por 32,000 parámetros (800 residuos). Los datos experimentales de 1tf5 dan una proporción de 1.72; resultado de 53,135 reflexiones y 30960 parámetros. Esa diferencia puede explicar las diferencias tanto en los valores de R y R_{free} como en la distribución de los valores de temperatura.

SOBREPOSICIÓN DE ESTRUCTURAS

La sobreposición global (de los $C\alpha$) de la conformación abierta (1tf5) con 1m6n (780 átomos) dio como resultado r.m.s.d.= 10.014 \AA . Se superpuso a cada uno de los dominios estructurales para observar su comportamiento tras el cambio de conformación. Los dos dominios de unión a nucleótidos fueron los que presentaron un cambio menor de una conformación a otra. El NBF1 (14-223, 353-396) superpone con una desviación de 1.33 \AA cuando se superpone todo el dominio, un

valor considerablemente alto. Al excluir a los residuos 14-17, 207-223 y 353-358 se obtiene una desviación de 0.56 Å. Si sólo se superpone la zona 17-206, 353-375 y 385-396 la desviación es de 0.49 Å. El valor más bajo se obtiene al comparar la región continua 17-206: r.m.s.d.= 0.37 Å. La sobreposición del dominio NBF2 (397-572) presenta una diferencia de 1.39 Å. Si excluimos a los residuos amino terminales 543-572 el valor se reduce a 0.64 Å. La unión de los dos dominios de unión a nucleótidos (17-206, 360-396 y 397-542) significa una diferencia considerablemente baja (r.m.s.d.= 0.62 Å), si consideramos las diferencias particulares entre cada dominio (**Figura 15 a**).

La hélice que atraviesa la molécula de extremo a extremo sugiere dos secciones: una que se mantiene sin o con pocos cambios (573-597, r.m.s.d.= 0.38 Å) y otra, la más lejana a los NBFs (598-618), que se modifica considerablemente de una conformación a la otra (r.m.s.d.=0.95 Å).

Los dominios PPXD y HSD-HDW son los que más cambian de una conformación a la otra. La sobreposición de PPXD (224-352) da como resultado 2.33 Å, mientras que la de HSD-HWD (619-780) resulta en 2.37 Å. En ninguna de las anteriores se obtienen mejoras significativas cuando se excluyen regiones del dominio.

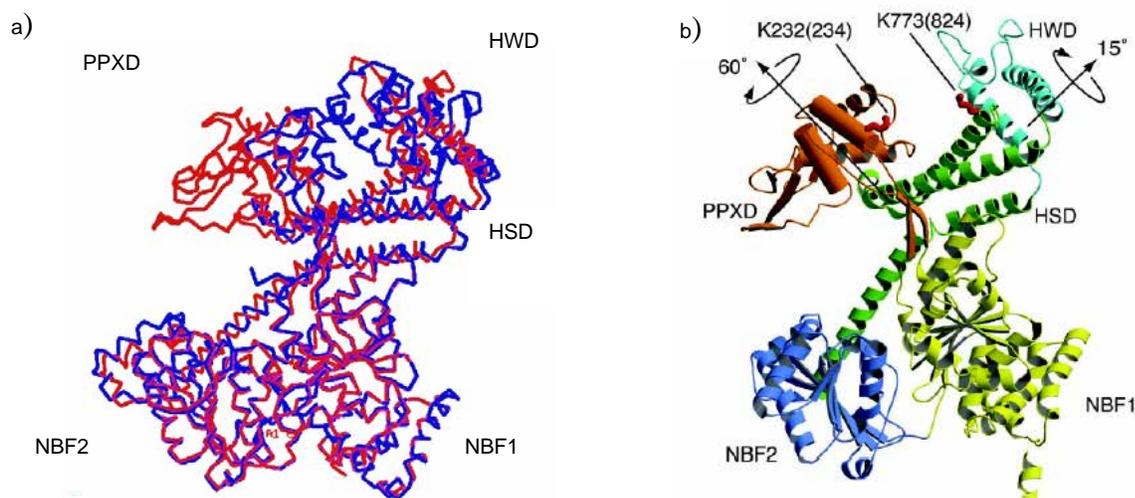


Figura 15 a) Sobreposición de las estructuras 1m6n (azul) y 1tf5 (rojo) en los dominios NBF1 y NBF2. b) Ejes y ángulos de rotación de los dominios PPXD y HWD reportados por Osborne et al.

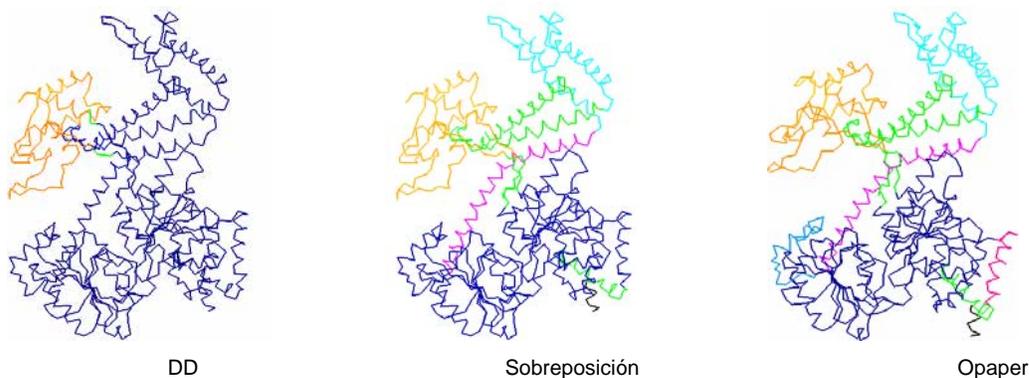
CAMBIO CONFORMACIONAL

La conclusión que surge de superponer las dos conformaciones (1tf5 y 1m6n) es que los dominios NBF1 y NBF2 se mantienen sin alteraciones globales significativas, son esencialmente idénticos en los dos modelos. En contraste, los otros dominios son secciones de cambios significativos. El más notable de ellos es la rotación de cuerpo rígido de $\sim 60^\circ$ del PPXD¹⁰⁰. La bisagra sobre la que gira se encuentra alrededor de los residuos 226 y 348, en las dos cadenas que conectan a PPXD y NBF1. La rotación de PPXD trae como consecuencia la ruptura de la interfase porosa que formaba junto

con HSD y HWD y la posterior formación de un gran surco entre los tres dominios. Juntos, los dominios HSD y HDW, rotan 15° en relación a su posición en 1m6n. Aunado a su rotación, HSD se desplaza una distancia de casi 2 Å en dirección de NBF2. Puesto que los cambios conformacionales no ocurren en los dominios NBF se puede concluir que no son reacomodos dependientes de nucleótidos.

SELECCIÓN DE GRUPOS PSEUDO-RÍGIDOS PARA EL AFINAMIENTO TLS

La selección de grupos pseudo-rígidos se hizo de la misma manera que en los casos anteriores. Toda la molécula como un solo grupo constituyó la selección **1G**. El programa *Dyndom* dio como resultado a los residuos [0-7[□], 14-230, 350-780] formando el dominio rígido, una bisagra [231-232 348-349] y un dominio móvil [233-347], todos dentro del conjunto nombrado **DD**. El conjunto de los grupos que resultaron de la sobreposición de las estructuras en el programa *O* recibió el nombre de **Sobreposición**, dividido en los siguientes ocho grupos: [0-7], [14-37], [38-216, 357-571], [226-350], [217-225, 351-356], [572-620], [621-706] y [707-780]. Basados en la sobreposición de dominios que hacen Osborne *et al.*⁹³ se construyó la selección **O_{paper}** que consiste en los grupos [0-7], [14-37], [38-56], [57-217, 357-540], [218-225, 351-356], [226-350], [541-571], [572-620], [621-706] y [707-780]. Partiendo del anterior, juntamos la hélice que atraviesa toda la molécula a lo largo con las dos hélices del dominio estructural HSD para formar el conjunto **O_{jointedhelix}**, idéntico al anterior excepto por la fusión de [572-620] y [707-780] en el grupo [572-620,707-780]. La última división de grupos fue la denominada **c₁₀₀**, que recibe ese nombre porque la molécula se dividió en un grupo cada 100 aminoácidos ([0-7, 14-100] [101-201] [202-302] [303-403] [404-504] [505-605] [606-706] [707-780]). La **Figura 16** muestra todas las selecciones de cuerpos pseudo-rígidos.



[□] El residuo 0 corresponde a una histidina, parte de la cola que se utiliza para purificar proteína y que no fue proteolizada.

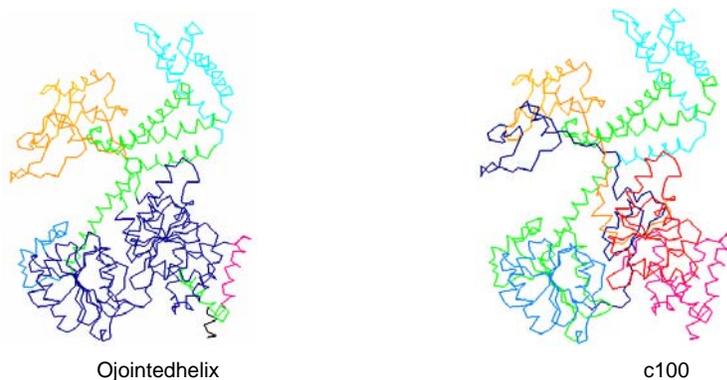


Figura 16 Selección de grupos realizada para el transportador SecA.

AFINAMIENTO TLS EN REFMAC5

Al igual que en los casos anteriores, comparando el depósito original con el afinamiento TLS los valores de R y R_{free} (24.1 y 27.3 respectivamente) se reducen en 2 puntos porcentuales (**Tabla 7**). Excepto la selección **1G**, todas las demás tienen valores de R y R_{free} menores de 1 punto porcentual si los comparamos con el isotrópico realizado en *refmac*. Después de los 5 ciclos de afinamiento la selección de la que se deriva el mejor modelo es **O_{paper}** pues, aunque con poca diferencia, tiene los valores más bajos en los residuales y el más alto en la figura de mérito. Le sigue de cerca la selección sobreposición, seguida de cerca por las selecciones **C₁₀₀** y **O_{jointedhelixes}**. El afinamiento de toda la molécula como un solo grupo resulta en una diferencia de 0.68 puntos porcentuales en R, 0.79 en R_{free} y 1.21 en FOM, valores significativamente mayores que los obtenidos para las demás selecciones.

Tabla 7 Afinamientos isotrópico, TLS y anisotrópico; 5 ciclos de TLS por 5 de máxima similitud.

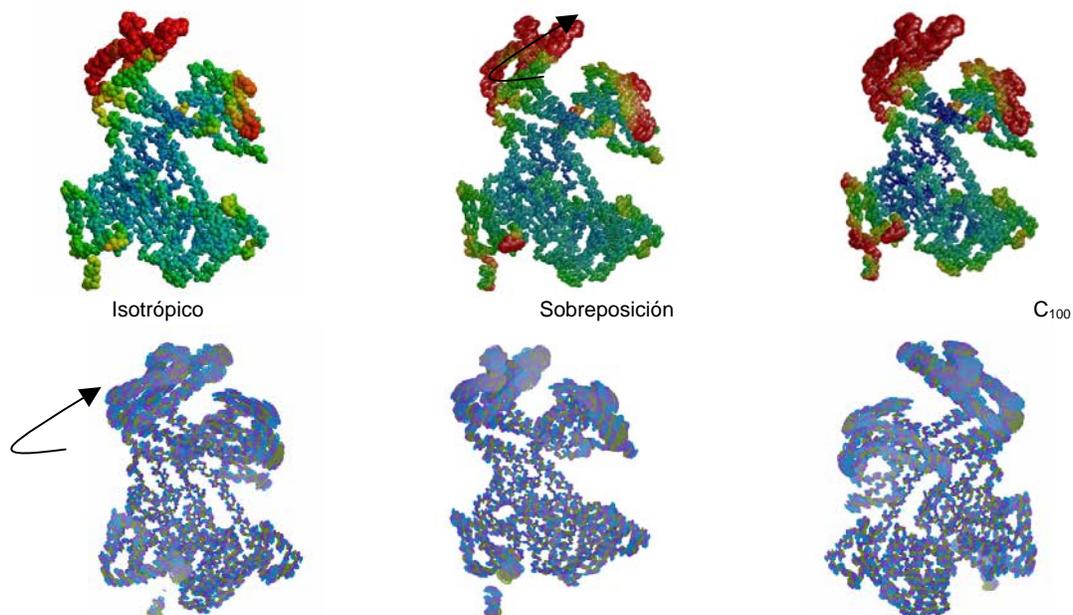
Grupo	R (%)	R_{free} (%)	FOM (%)
Isotrópico	22.08	25.95	81.01
1G	21.26	25.21	82.37
Sobreposición	20.72	24.54	83.38
DD	20.90	24.82	82.80
O _{jointedhelixes}	20.76	24.56	83.26
O _{paper}	20.58	24.40	83.58
C ₁₀₀	20.77	24.58	83.12

Después de los 10 ciclos alcanzamos la convergencia de los estadísticos y encontramos resultados diferentes de los anteriores (**Tabla 8**). Aunque el valor de R del isotrópico disminuye 0.10 puntos porcentuales, aumenta el de R_{free} en 0.6 y disminuye la FOM en 0.67. Una divergencia similar ocurre con el control **1G**; disminuye 0.32 puntos porcentuales en R, pero aumentan tanto R_{free} (0.25 puntos porcentuales) como FOM (0.13 puntos porcentuales). Los valores siguen reduciéndose en las demás selecciones de pseudo-rígidos de forma tal que ahora la mejor selección es **Sobreposición**, seguida de **C₁₀₀**, **O_{paper}**, **O_{jointedhelixes}** y **DD**.

Tabla 8 Afinamientos isotrópico, TLS y anisotrópico; 10 ciclos de TLS por 10 de máxima similitud.

Grupo	R(%)	R _{free} (%)	FOM (%)
Isotrópico	21.98	26.55	80.44
1G	20.94	25.46	82.24
Sobreposición	20.37	24.49	83.51
DD	20.68	24.83	82.81
O _{jointedhelixes}	20.60	24.81	83.11
O _{paper}	20.50	24.42	83.40
C ₁₀₀	20.39	24.52	83.25

Consecuente con la distribución de valores térmicos del archivo 1tf5, el dominio HWD es el que presenta los valores más altos. La representación tridimensional de los elipsoides, en la selección de grupos **Sobreposición**, nos muestra que los ejes mayores coinciden con un desplazamiento del dominio con dirección tal que tendería a cerrar el surco formado con la interfase del dominio HWD. La otra zona de alta flexibilidad se encuentra en el dominio PPXD. También en esta región los ejes mayores de los elipsoides en rojo de la parte inferior parecen desplazarse en dirección de NBF2. Mientras que la dirección de rotación de HWD se conserva en todas las selecciones de pseudo-rígidos y coincide con el cambio de conformación descrito en la literatura¹⁰⁰, la situación se complica en PPXD, pues la dirección de desplazamiento no tiene nada que ver con el cambio conformacional y desaparece en la selección **C₁₀₀**.



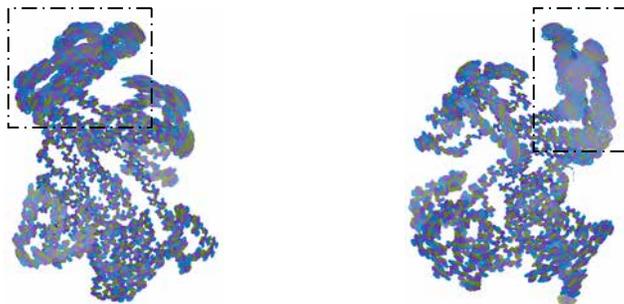


Figura 17 Representación de los elipsoides anisotrópicos de acuerdo al valor de B en la molécula 1tf5 después del afinamiento por tensores TLS de la selección Sobreposición.

Cuando analizamos la **Gráfica 6** observamos que la distribución de Bs en el dominio PPXD (224-352) no parece la de un cuerpo pseudo-rígido que oscila de manera conjunta. Más bien se distinguen dos zonas: una de alta (234-266) y otra de baja (275-349) vibración. Los elipsoides de esa región no muestran una dirección clara (cuadro punteado **Figura 17**). Lo mismo ocurre con el dominio HWD. Aunque en este último el color rojo de los elipsoides del dominio indica que el valor de los Bs y por lo tanto de las oscilaciones es mayor. La diferencia en los valores de B en la porción 275-349 podrían explicarse por una estabilización causada por contactos cristalinos entre los residuos 322-325 de una molécula y 158-160 de su vecino cristalográfico (**Figura 18**).

Como no es clara la dirección de las oscilaciones decidimos analizar los valores de los tensores para cada grupo en particular buscando más información sobre las oscilaciones en SecA. Por los bajos valores del tensor T (en promedio menores que una décima de 1 Å) en cada uno de los grupos podemos decir que en general los movimientos se restringen a libraciones (datos no mostrados o Apéndice). Los valores más altos del tensor de libración (L), en las selecciones **Sobreposición** y **C₁₀₀**, se encuentran en los residuos 226-350 y 621-706 que se superponen con los dominios estructurales PPXD y HWD respectivamente. Lo mismo encontramos en la selección **C₁₀₀**, en la que los valores más altos del tensor de rotación corresponden a las regiones 202-302, 404-505 y 607-707. Topológicamente sobreponen con PPXD, NBF2 y HWD.

La representación gráfica de los ejes de libración, con longitud proporcional a la magnitud, permite observar las direcciones predominantes. Así pues para poderlos localizar corrimos de nuevo el análisis con TLSANL sólo que ahora con la opción de ejes representables en Molscript⁸³. En la **Figura 18** vemos que la dirección predominante de la libración en el dominio HWD coincide con el esperado para el cambio de conformación. En tanto que la del dominio PPXD el eje principal de libración se encuentra perpendicular al que esperaríamos resultase en el cambio conformacional que hemos descrito anteriormente. Sin embargo hay otro eje de libración, de magnitud considerable, que corresponde a la bisagra que conecta a PPXD con NBF1. Una libración de esa hoja β sobre ese eje afectaría directamente la posición de PPXD.

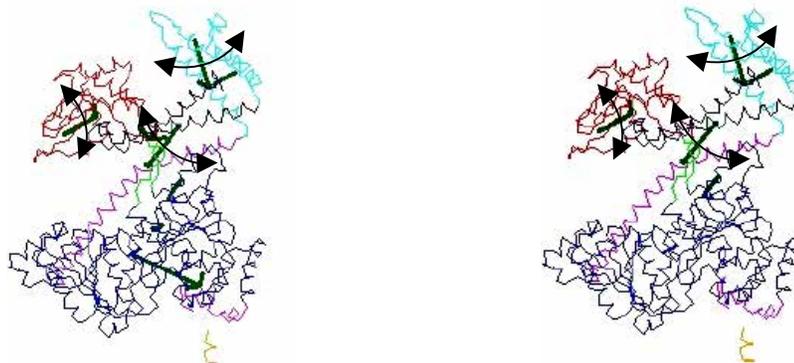


Figura 18. Ejes de libración del afinamiento TLS de la selección sobreposición.

Como mencionamos en el apartado de cambio conformacional, los dominios NBF1 y NBF2 se mantienen sin cambios significativos entre la estructura abierta y cerrada. La rotación del dominio PPXD, que rota $\sim 60^\circ$ de una conformación a la otra, es ocasionada primordialmente por el giro de la bisagra 217-225 y 351-356 posicionándolo más cerca de HWD. Dicha rotación rompe la interfase porosa que formaba con los dominios HSD y HWD, formando un gran surco entre los tres dominios. El dominio HDW rota 15° con respecto de su posición en 1m6n. Esos movimientos podrían ser los responsables de la exportación de los pre-péptidos hacia el periplasma.

UDP-N-ACETILGLUCOSAMINA ENOLPIRUVIL TRANSFERASA (EPT)

El paso inicial en la síntesis de mureína, el componente más abundante de la pared celular bacteriana, es la transferencia del grupo enolpiruvil, con el fosfoenolpiruvato (PEP) como donador, al 3-hidroxilo (3-OH) de la UDP-N-acetilglucosamina (UDP-GlcNAc). La reacción es catalizada por la enzima UDP-N-acetilglucosamina enolpiruviltransferasa (EPT; E C 2.5.1.7). Como producto de la actividad enzimática se obtiene 3-enolpiruvil éter de enolpiruvil-UDPGlcNAc (EP-UDP-GlcNAc). La formación del azúcar mureína se completa tras la reducción de este producto al 3-lactiléter UDP-MurNAc catalizada por la flavoenzima MurB¹⁰⁴. El grupo lactil carboxilo libre de UDP-MurNAc sirve como sitio de unión para la porción peptídica de la pared celular.

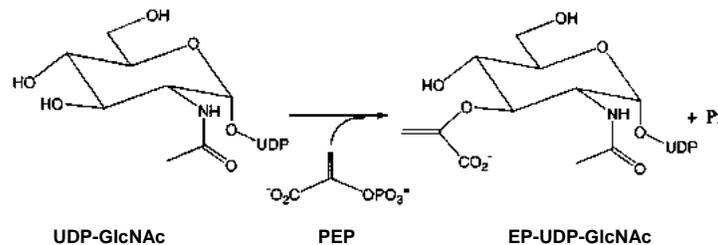


Figura 19 Primera reacción para la síntesis de mureína: transferencia del enolpiruvil a la UDP-GlcNAc con la consecuente formación de EP-UDP-GlcNAc.

La reacción catalizada por la EPT es mecánicamente inusual pues involucra el ataque en la posición electrofílica C₂ del PEP causando la ruptura del enlace C-O, siendo lo más común en las enzimas dependientes de PEP la transferencia o donación de un grupo fosforilo¹⁰⁵.

ESTRUCTURA

La EPT es una proteína de 419 residuos de aminoácidos plegada en dos dominios globulares casi esféricos, con radio aproximado de 20 Å, constituidos por seis α -hélices y tres hojas β de cuatro cadenas, unidos por una bisagra de doble cadena (**Figura 20**). El 33% de los residuos de aminoácidos se encuentra en las hélices mientras que en las hojas β se distribuye el 26% restante. Tanto el C como el N terminal se encuentran en el dominio I (residuos 1-17, 232-419), dominio formado en su mayoría por residuos de la porción C-terminal. El dominio II, o porción N-terminal, comprende los residuos 22-227. La unión de ambos dominios está mediada por los segmentos 18-21 y 228-231 de cuatro aminoácidos cada una.

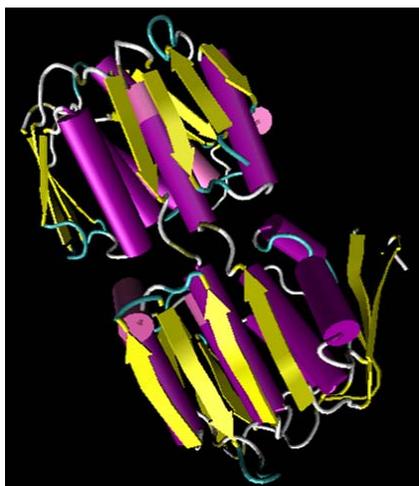


Figura 20 Estructura tridimensional de la unidad biológica de UDP-N-Acetil-glucosamina-enolpiruvato transferasa.

La característica topológica más interesante de esta estructura es que consiste en la repetición de seis subdominios o unidades de plegamiento. Cada unidad consiste de una hoja β de cuatro cadenas, dos paralelas (β_1 y β_2) y dos antiparalelas (β_3 y β_4) y dos α -hélices paralelas (α_1 y α_2) (**Figura 21**). Tres de los subdominios se forman por segmentos contiguos (I_c , II_a y II_c), mientras que en los tres restantes una cadena de la hoja β no es adyacente en la estructura primaria de la subunidad (corresponden a los residuos 3-8 β_4 de I_a =antiparalela, 13-17 β_1 de I_b =paralela y 224-227 β_1 de II_b =paralela). Por lo tanto cinco de seis cadenas antiparalelas están unidas por *hairpin loops*. La excepción es la unidad II_a en la que las antiparalelas se unen a través de un asa de 5 residuos (207-211), la conexión correspondiente en las unidades II_c (140-141), I_c (349-350), I_b (276-278) y II_b (67-69), son *hairpin loops* de dos y tres residuos respectivamente.

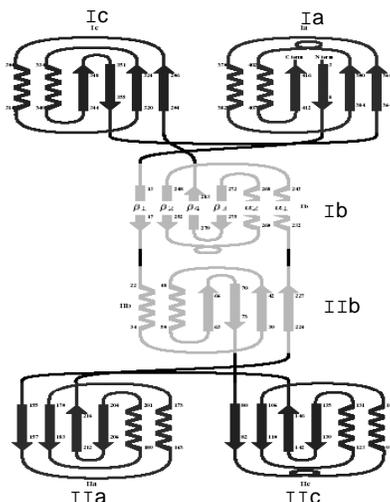


Figura 21 Topología de EPT. Se muestra la repetición de las 6 unidades de plegamiento que comprenden los elementos β_1 , α_1 , β_2 , α_2 , β_3 y β_4 . Las unidades son nombradas Ia-Ic y Iia-Iic de acuerdo a su ubicación en el dominio respectivo. Las loops con giro de las unidades Ia, Ib y Iic simbolizan α -hélices de un solo giro. El número de aminoácido indica el inicio o fin de los elementos de estructura secundaria.

Los núcleos de ambos dominios comprenden α hélices rodeadas por 3 α hélices con caras accesibles al solvente, y las tres hojas β de 4 cadenas. Por tanto se puede caracterizar el plegamiento como un barril α/β dentro-fuera.

Su centro de masa reside en la interfase de los interdominios. El eje z se puede considerar como un eje de simetría 3 relacionando las subunidades dentro de cada dominio. El eje y relaciona a ambos dominios por una díada aproximada.

A pesar de la similitud en la estructura secundaria y plegamiento general, el único elemento repetitivo en la estructura primaria es Leu-X₃-Gly(Ala) presente en el asa que conecta la α hélice 2 con la hoja β_3 , y es, aparentemente el responsable de la unión entre las subunidades.; el grupo carbonilo de cada cuarto residuo de la secuencia repetida hace puentes de hidrógeno con una amida de cadena principal de la subunidad adyacente.

ESTRUCTURAS DE EPT DEPOSITADAS EN EL PDB

La búsqueda en el PDB arrojó ocho estructuras depositadas, resueltas todas ellas por difracción de rayos X, correspondientes a la EPT: seis de *Enterobacter cloacae* y dos de *Escherichia coli* (Tabla 9).

Tabla 9 Depósitos de UDP-N-acetilglucosamina enolpiruvato transferasa.

Clave PDB	Organismo	Grupo espacial	R (%)	R _{free} (%)	Resolución [Å]	Unidad asimétrica	Ligandos
1a2n	<i>E. coli</i>	P 3 2 1	17.5	---	2.8	Monómero	TET* ¹
1dlg	<i>E. cloacae</i>	C 1 2 1	17.3	19.7	1.9	Dímero	PO ₄ , AMC* ²
1ejc	<i>E. cloacae</i>	P 2 ₁ 2 ₁ 2 ₁	16.7	19.2	1.8	Monómero	PO ₄ , GOL* ³
1ejd	<i>E. cloacae</i>	C 1 2 1	18.5	20.7	1.55	Dímero	PO ₄ , HAI* ⁴
1eyn	<i>E. cloacae</i>	P 2 ₁ 2 ₁ 2 ₁	18.0	21	1.7	Monómero	2AN* ⁵ , GOL
1naw	<i>E. cloacae</i>	C 1 2 1	19.7	27	2.0	Dímero	HAI
1q3g	<i>E. cloacae</i>	P 1 2 ₁ 1	21.7	25.3	2.65	Decahexámero	UDA* ⁶ , EGL* ⁷
1uae	<i>E. coli</i>	P 3 2 1	18.5	---	1.8	Monómero	FCN* ⁸ , UD1* ⁹

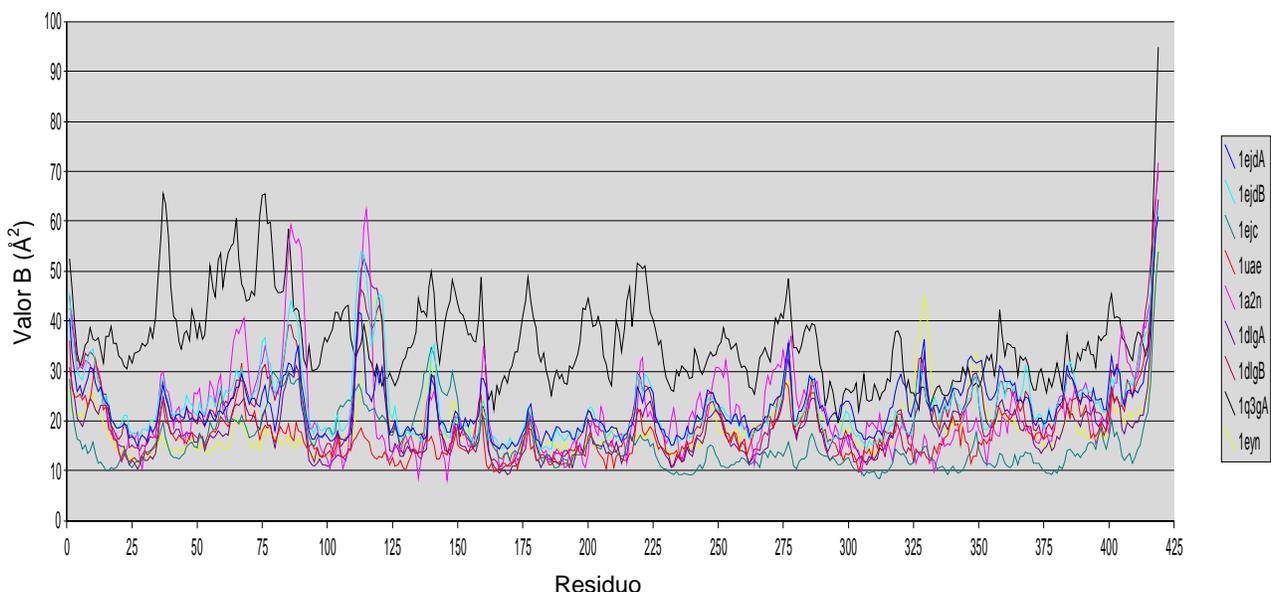
*¹TET.- Análogo fluorinado de UDP-N-acetilglucosamina. *²AMC.- aminometilciclohexano. *³GOL.- glicerol. *⁴HAI.- ion hidroximetilamonio. *⁵2AN.-1-anilino-8-naftaleno-sulfonato. *⁶UDA.- N-acetilglucosamina difosfato. *⁷EGL.- etilenglicol. *⁸FCN.- fosfomicina, *⁹UD1.- UDP-N-acetilglucosamina.

A pesar de que las dos estructuras de *E. coli* tienen buena resolución y valores aceptables de R, sólo existen en la forma con ligandos además de que no presentan factores de estructura, esta última característica impide el afinamiento de cualquiera de ellas. En cambio, las estructuras 1ejc y 1ejd de *E. cloacae* corresponden a moléculas libres de ligandos y contienen depositados los factores de estructura. Aunque sin factores de estructura depositados, la molécula de clave 1q3g corresponde a la forma cerrada de *E. cloacae* ofreciéndonos un punto de comparación para 1ejc y 1ejd. La unión de N-acetilglucosamina uridina bifosfato (UDA) estabiliza a 1q3g en la conformación cerrada.

DISTRIBUCIÓN DE VALORES B

Los valores B de la estructura 1q3g (**Gráfica 6**) presentan un patrón muy diferente y valores más altos que las demás. A primera vista las diferencias pueden ser explicadas por su baja resolución (2.65 Å): al revisar el archivo .pdb notamos que la relación datos/parámetros era muy baja (0.69) para un afinamiento isotrópico, pues reportan 185,835 reflexiones para una estructura con 16 cadenas en la unidad asimétrica de 419 residuos de aminoácidos cada una, lo que como mínimo requeriría de 268,160 observaciones necesarios (número aproximado de átomos en la unidad asimétrica por 4 parámetros; tres de posición y uno térmico).

Distribución de valores B

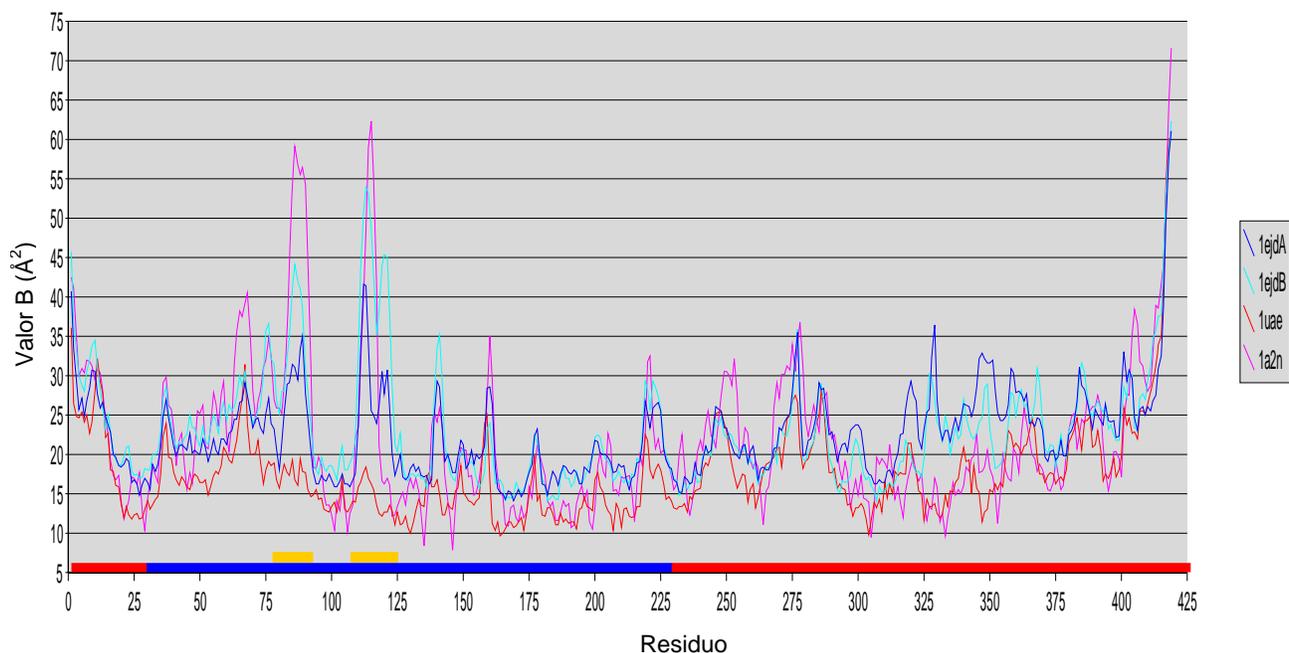


Gráfica 6 Distribución de los valores B en las estructuras depositadas en el PDB de la UDP-N-acetilglucosamina enolpiruvato transferasa.

Los patrones de distribución de las estructuras restantes son bastante parecidos. La estructura 1ejd alcanza sus valores más altos de B en las regiones 74-91 y 109-122. Aunque también está depositada como libre de ligandos, la estructura con clave 1ejc tiene valores más bajos que las

demás en el dominio C terminal. En lo que respecta al N terminal su comportamiento se parece bastante al de 1ejd, pero con valores menores, en particular en la región 109-122, correspondiente al asa catalítica. Aunque pertenece a otro organismo, podríamos considerar al depósito 1uae como la forma cerrada de la estructura pues su patrón de valores térmicos es muy similar a los demás, a excepción de la región 110-121: la zona del cambio conformacional. Además, su estructura tridimensional (ver Sobreposición) y mecanismo de acción¹⁰⁶ son prácticamente los mismos.

Comparación de los valores B en las estructuras abiertas y cerrada



Gráfica 7 Distribución de los valores B en las estructuras abierta (1ejd) y cerradas (1uae y 1a2n).

SOBREPOSICIÓN DE ESTRUCTURAS

La sobreposición de toda la cadena A de la estructura libre (1ejd) con la misma cadena de la estructura con ligando (1q3g) resulta en una desviación media cuadrada (r.m.s.d.) de 2.12 Å. Un valor muy cercano se obtiene cuando se superpone el dominio I (residuos 20-229) de la estructura 1ejc (r.m.s.d.= 2.02 Å.). Al observar la sobreposición en el programa 'O' notamos que la mayor parte del dominio I superpone muy bien, a excepción de las regiones 110-121 y 137-144. La primera corresponde al asa catalítica en la que se encuentra el residuo Cys115 responsable de la unión de PEP y fosfomicina, mientras que 137-144 es un asa más corta adyacente a la catalítica en la estructura tridimensional y que no participa directamente en la unión de ligandos ni aporta cadenas laterales al sitio catalítico. Al excluir estas regiones de la sobreposición obtuvimos una r.m.s.d.= 0.390 Å (**Figura 22 a**), mucho menor que las anteriores. De igual forma se realizó la sobreposición del dominio II (1-19, 230-419). Considerando a todo el dominio obtuvimos una r.m.s.d.= 0.584 Å

(Figura 22 b)). Es importante recordar que en este dominio se encuentran las dos regiones, amino y carboxilo, terminales; las zonas con mayor desorden y vibración en la mayoría de las estructuras. Observando la sobreposición excluimos del siguiente intento a los residuos 1-3 y 415-419 obteniendo una r.m.s.d.= 0.427 Å. La sobreposición de las cadenas A y B es perfecta (r.m.s.d= 0.00). Cuando superponemos alguna de las cadenas de 1ejd con 1ejc obtenemos una desviación de 1.22 Å, que se reduce a menos de 0.5 Å (0.41 Å) si calculamos sólo en el dominio C-terminal o en el N-terminal excluyendo 110-121 y 136-144.

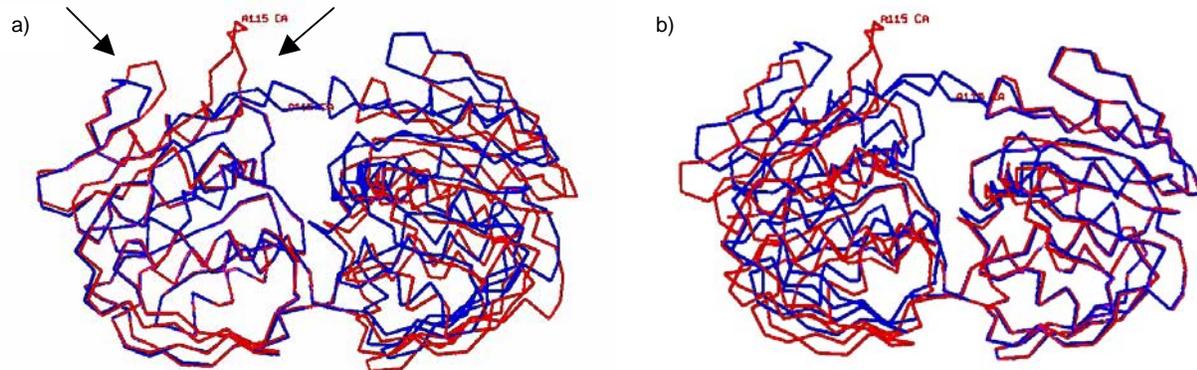


Figura 22 Sobreposición de las conformaciones libre (rojo) y con ligandos (azul) de la UDP-N-acetilglucosamina enolpiruvato transferasa en a) su dominio amino terminal y b) en la porción carboxilo terminal. Las flechas en a) señalan los sitios que más se modifican tras la unión del ligando. La flecha de la izquierda señala el asa 137-144 y la de la derecha el asa 111-121.

Por la similitud en los patrones de distribución de las estructuras de *E. coli* y *E. cloacae* decidimos superponer las estructuras abiertas y cerrada de una con otra. La comparación de las formas cerradas (1q3g y 1uae) da una sobreposición de 0.78 Å para todos los residuos. La mayor diferencia está en los residuos 111-115 del asa catalítica y la región 134-144 (figura no mostrada). Cuando fueron excluidos de la sobreposición obtuvimos una desviación de 0.47 Å. De la sobreposición 1uae-1ejd las formas abiertas obtuvimos 0.47 Å en el dominio N-terminal (22-109, 122-133, 144-228) y 0.43 Å en el dominio C-terminal (229-415). Y para 1uae-1ejc 0.5 Å (22-109, 122-133, 144-228) y 0.67 Å (229-415).

A pesar de la poca similitud entre las secuencias (21%) de *E. coli* y *E. cloacae* su sobreposición resultó muy parecidas a las realizadas entre los modelos de una misma *E. cloacae*. O incluso mejor si tomamos en cuenta que superpusimos un mayor número de átomos con la de *E. coli*.

CAMBIO CONFORMACIONAL

La unión de ligandos provoca cambios estructurales en ambos dominios. El dominio C-terminal rota sobre la bisagra en dirección del dominio adyacente. En el dominio N-terminal el asa catalítica (Pro112-Gly-Gly-Cys-Ala-Ile-Gly-Ala-Arg-Pro121), se acerca al dominio C-terminal, cerrando la interfase por la parte superior (**Figura 22 a**)). La conformación cerrada se estabiliza por múltiples

puentes de hidrógeno entre los residuos 113-119 del asa catalítica y los residuos 329-330 del N-terminal. En la segunda mitad del asa se encuentran las mayores diferencias¹⁰⁷: en la forma abierta, cadenas de solvente de 11-14 Å mantienen unidos a los residuos Asp123, Gly141 y Tyr142, unión que se deshace, en la forma cerrada, por la interacción con UDP-N-acetilglucosamina. Una interacción más que se modifica es la hidrofóbica entre los residuos Leu111 y Val122 (que funcionan como ancla), indicando que su reacomodo debe ser importante y necesario para los cambios estructurales del asa. Como consecuencia del reacomodo de las cadenas laterales de Asp123 y Val122, el anillo de la Pro121 queda libre para rotar cerca de 180°, alterando la conformación de los residuos que suceden. También el asa que forman a los residuos 137-144 sufre un cambio de conformación, pero sin aparente relación directa con la actividad enzimática. En suma, los cambios producen una molécula más compacta, con menor área de interfase y quizá con menor acceso al solvente, aislando al sustrato del ambiente exterior.

SELECCIÓN DE GRUPOS PSEUDO-RÍGIDOS PARA EL AFINAMIENTO TLS

A manera de control seleccionamos las cadenas A y B como un solo grupo (**1Grupo**) que incluye a los residuos 1-419A y 1-419B. El grupo denotado como **2** considera a cada cadena como un solo grupo, es decir, los residuos 1-419 de la cadena A y 1-419 de la B por separado. El grupo denotado como **2Grupos** incluye cuatro pseudo-rígidos (dos por molécula); uno que equivale al dominio estructural I y el otro al dominio estructural II (**Figura 23a**). La selección **3Grupos** es una modificación de la anterior: incluye al dominio II [1-21, 230-419] completo pero considera al asa catalítica [111-122] como un cuerpo independiente y a la sección restante del dominio I [22-110, 123-229] como otro (**Figura 23b**). El grupo **DD** fue sugerido por los resultados del programa DynDom. Partiendo de las estructuras 1ejd y 1q3g obtuvimos tres dominios dinámicos: uno fijo [231-419], uno móvil [22-228] y la bisagra que los conecta [19-20, 229-230] (**Figura 23c**). La selección **Topología** toma como un grupo a cada una de las seis repeticiones de los subdominios estructurales ([1-8, 364-419] [294-355] [80-146] [155-216] [13-17, 232-283] [22-73, 224-227]) y a las secciones que los conectan ([9-12] [74-79] [147-154] [217-223] [18-21, 228-231] [284-293] [353-363]) (**Figura 23d**). Por último la selección **Sobreposición** toma como cuerpos rígidos a los residuos excluidos de la sobreposición por el programa o, [1-3] [4-19, 230-362, 370-389, 399-414] [20-109, 122-136, 144-229] [110-121] [137-143] [363-369] [390-398] y [415-419] (**Figura 23 e**)).

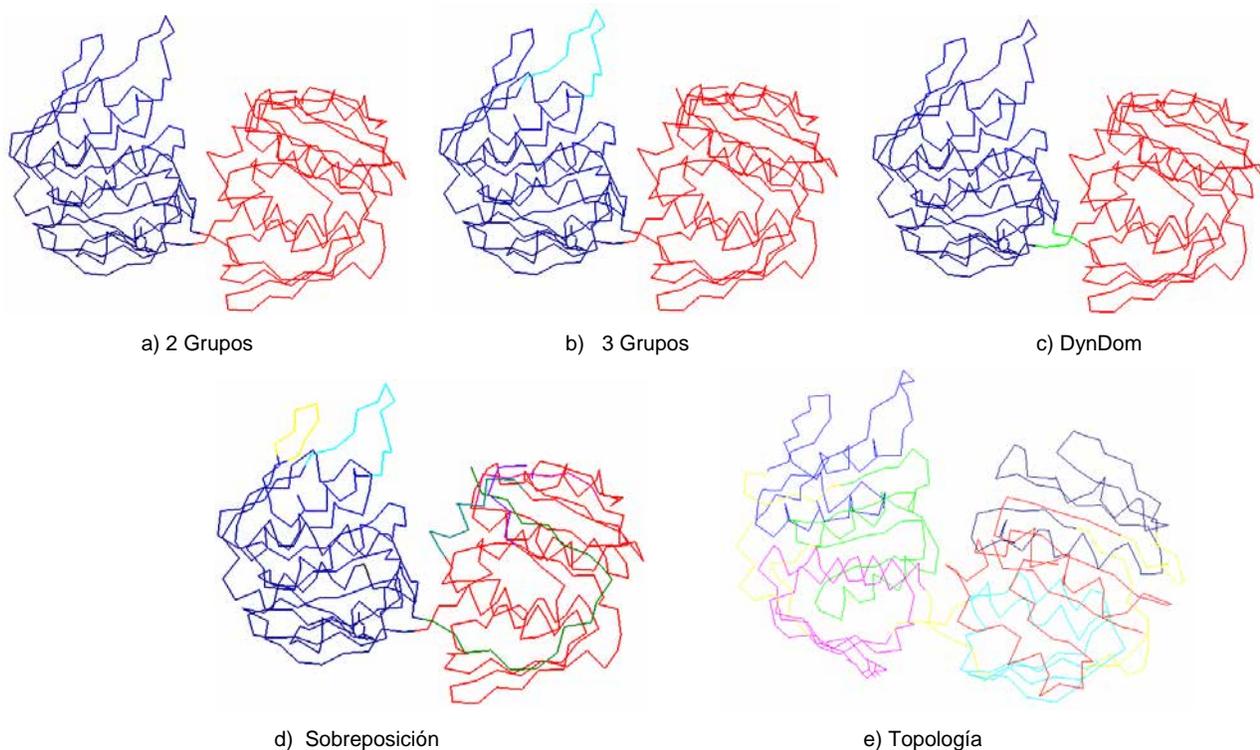


Figura 23. Grupos seleccionados como cuerpos pseudo-rígidos para el afinamiento por tensores de la enzima EPT.

AFINAMIENTO TLS

El afinamiento por tensores se realizó para las dos estructuras reportadas como forma libre (1ejc y 1ejd). En el caso de 1ejc todos los intentos resultaron en sobreafinamiento reportado por el programa como *'Problems in make U Positiv'*. Lo cual quiere decir que se está forzando demasiado el afinamiento del modelo, es decir, se obliga tanto a los átomos (electrones) para ajustar la densidad electrónica que se rompen las restricciones geométricas. En cambio, los realizados con la estructura 1ejd resultaron todos exitosos, tal como muestra la **Tabla 9**.

Tabla 9 Afinamientos isotrópico, TLS y anisotrópico; 5 ciclos de TLS por 5 de máxima similitud.

Grupos	R (%)	R_{free} (%)	FOM (%)
Isotrópico	18.70	20.73	85.99
1 Grupo	18.65	20.73	86.07
2 Grupos	17.67	19.70	88.12
3 Grupos	17.36	19.42	88.43
Topología	17.18	19.23	88.67
Sobreposición	17.23	19.24	88.59
Anisotrópico	17.55	19.75	87.72

El afinamiento isotrópico se mantiene casi idéntico a los valores R y R_{free} depositados (18.5 y 20.7). Valores muy parecidos se obtienen tras el afinamiento, ya por tensores TLS, de 1 Grupo. A partir de los 2 Grupos los valores empiezan a descender poco más de un punto porcentual en los dos residuales. Lo mismo ocurre con la figura de mérito que es muy cercana a 86% en el isotrópico y

1 Grupo, mientras que en los demás está alrededor de 88 %. El mejor resultado lo obtuvimos con la selección **Topología**, aunque con muy poco margen sobre **Sobreposición** y **3 Grupos**. Es de notar que estas tres selecciones dan mejor resultado que el mismo afinamiento anisotrópico. Después de analizar paso a paso cada ciclo del afinamiento notamos que no se había alcanzado aún la convergencia por lo que decidimos aumentar el número de ciclos paulatinamente. La convergencia se alcanzó con diez ciclos de TLS por la misma cantidad de máxima similitud (**Tabla 10**).

Tabla 10 Afinamiento TLS; 10 ciclos de TLS por 10 de máxima similitud

Grupos	R (%)	R_{free} (%)	FOM (%)
Isotrópico	18.70	20.80	86.14
1grupo	18.52	20.54	86.14
2Grupos (por cadena)	17.32	19.38	88.47
3Grupos	17.18	19.30	88.63
Topología	16.99	19.06	88.85
Sobreposición	17.10	19.15	88.72
anisotrópico	17.49	19.74	87.91

Tras los diez ciclos observamos ligeras reducciones en los evaluadores de la calidad del modelo así como el mismo comportamiento general: los valores más altos son los del afinamiento isotrópico que había convergido desde los cinco ciclos, demostrado por el aumento del valor de R_{free} , aunque con una ligera mejoría en la FOM. Para la selección **1 Grupo** mejoraron ligeramente pero se mantienen mucho más cerca de los valores depositados que de las demás selecciones. **Topología** alcanza menos de 17% en R, poco más de 19% en R_{free} y 88.85% en FOM. Es decir, una reducción de 1.5 y 1.6 puntos porcentuales en R y R_{free} en comparación de los valores depositados. Si comparamos con los valores del afinamiento isotrópico tenemos mejores resultados aún (1.71 y 1.74 puntos porcentuales). Se mantienen cercanos a Topología los valores de 3 Grupos y Sobreposición; los tres mejores que el anisotrópico.

Después de la convergencia la observación más importante es que las selecciones de pseudo-rígidos siguen dando mejores resultados que el afinamiento anisotrópico. Nuestra explicación original era que la resolución del modelo (1.55 Å) se encuentra dentro de la zona gris, en la que se duda si la cantidad de datos es suficiente para aumentar los parámetros a los requeridos por el afinamiento anisotrópico. Sin embargo la relación datos/parámetros (considerando los 9 parámetros por átomo necesarios para un afinamiento anisotrópico) es mayor que 2. Entonces, la zona gris no es un argumento del todo convincente. Quizá lo que tengamos sea un modelo mejor representado por los grupos pseudo-rígidos que por los anisotrópicos individuales; algo polémico y difícil de argumentar.

Aunque las diferencias en los estadísticos parecen mínimas, decidimos representar las elipsoides en el programa `Raster 3D` para buscar diferencias cualitativas entre cada una de las selecciones. En la **Figura 24** podemos observar que las zonas de mayor vibración (hacia los colores amarillo y rojo) se mantienen constantes de una selección a otra. La selección de los grupos no modifica las zonas más flexibles de la molécula, pero sí aumenta el valor de los valores B. Podemos observar

que en la zona del asa catalítica (la parte superior derecha de la cadena B) la zona bajo el color rojo tiene mayor cantidad de amarillo en las selecciones 3G y topología que en isotrópico y anisotrópico, estas últimas muy similares entre sí. También es significativo el aumento de los Bs en la zona correspondiente a las dos regiones terminales de las selecciones que involucraron afinamiento por tensores TLS. Además del aumento de los valores de B en la región del asa se puede observar claramente, sobre todo en la selección 3G, que el eje mayor de los elipsoides del dominio N terminal tienen una dirección común, marcada por la flecha, que tiende a cerrar la cavidad interdominios. Esta tendencia se mantiene en la selección topología. La dirección de los ejes mayores descrita es la misma que la que sigue el cambio de conformación causado por la unión de ligantes.

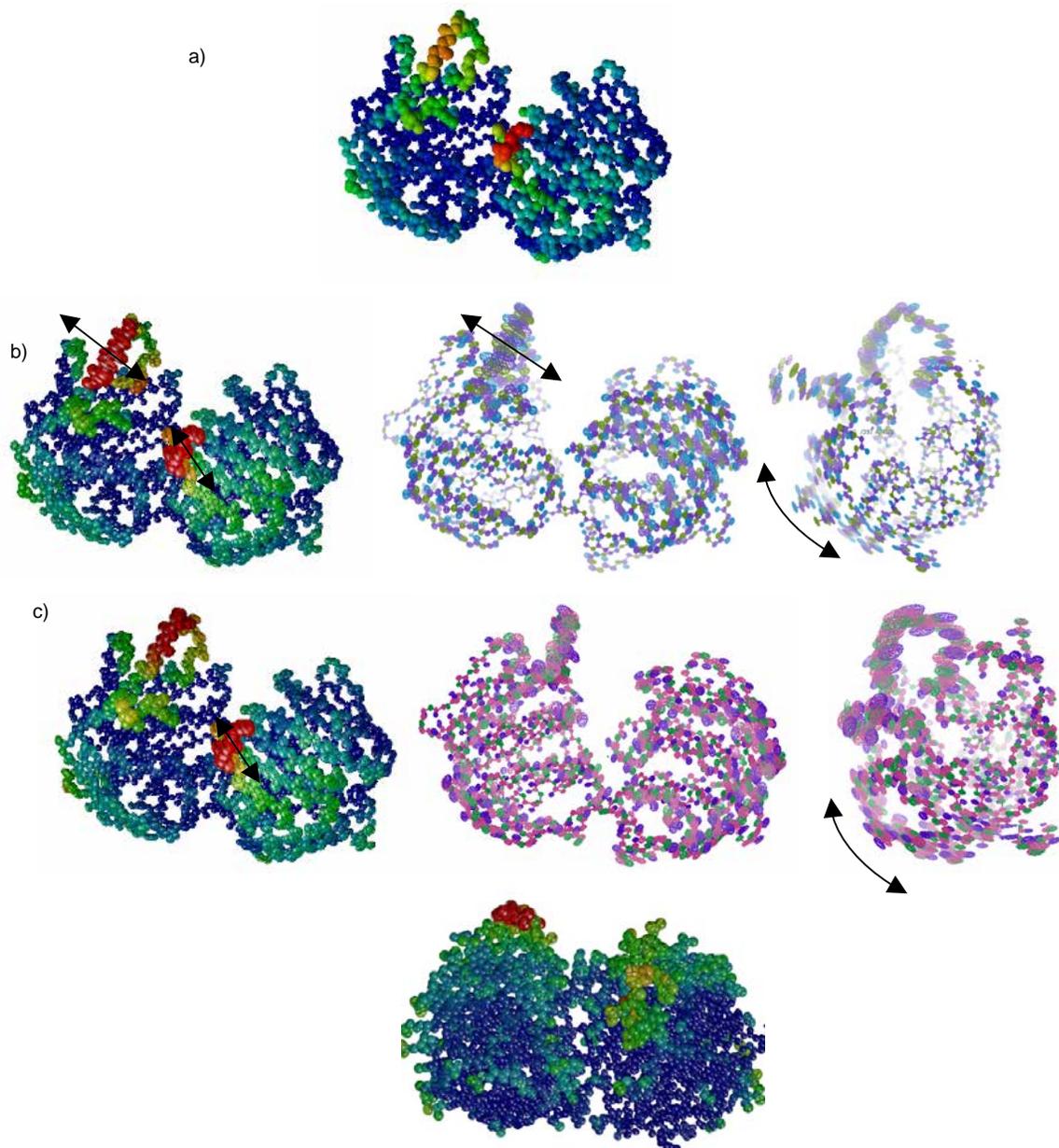
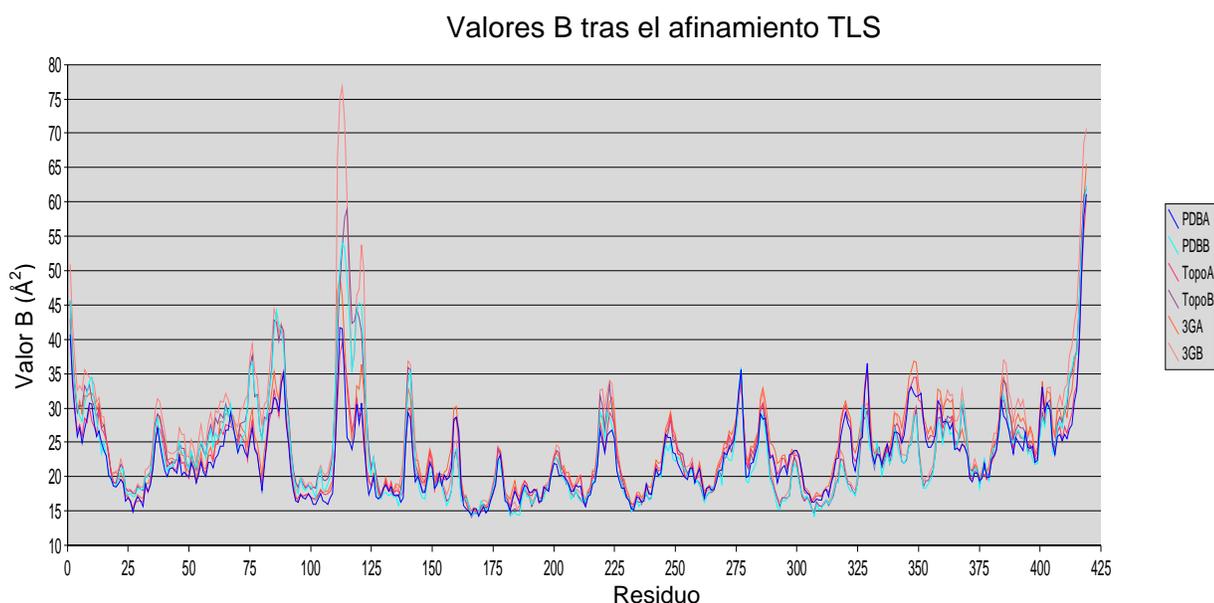


Figura 24 Elipsoides térmicas isotrópicas y anisotrópicas para los diferentes grupos seleccionados: a) isotrópico, b) 3G, C) topología y D) anisotrópico.

La otra región en rojo corresponde a las regiones terminales. Aunque lo común es excluir a las regiones terminales, en esta ocasión podrían tener algo que ver en la rotación del dominio C-terminal pues los ejes mayores parecen coincidir en dirección a la parte superior (donde se encuentra el asa catalítica) del dominio adyacente.

Para demostrar que el aumento de los Bs no es puramente cualitativo, la graficamos nuevamente residuo vs. valor B del depósito original (1ejd) y de los afinamientos por tensores que mejor comportamiento mostraron (**Gráfica 8**). Excluimos al isotrópico y anisotrópico porque prácticamente se superponían con los del depósito original. Es evidente que tras el afinamiento por tensores TLS los valores de B aumentan en la región del asa catalítica (110-120), siendo la selección **3G** en la que más se incrementan, alcanzando valores mayores de 75 en comparación del 55, el valor mayor en el PDB original. Esto si comparamos solamente las cadenas B. En el caso de las cadenas A el aumento es menos evidente; se incrementa tan sólo 6 unidades (de 42 a 48) en los residuos con Bs mayores, alrededor del 114.



El análisis de los elipsoides coincide con la descripción del cambio conformacional. El asa catalítica (residuos 111-122, en particular los primeros 5 residuos) oscila en dirección al dominio carboxilo terminal. Aparentemente los dos extremos, amino y carboxilo terminales, tienen un movimiento que tiende a acercar al dominio C-terminal al asa catalítica. Con la estabilización del asa tras la unión de los sustratos, se establecen interacciones.

En conjunto las modificaciones que ocurrirán tras la unión de ligandos pueden deducirse del análisis de los elipsoides. El asa catalítica tiende a acercarse al dominio C-terminal, quien a su vez oscila en dirección del N-terminal, al menos en la región de los extremos terminales. La unión de los sustratos provocaría la estabilización de la estructura cerrada, en cuyo sitio activo se encuentra ahora

posicionado el residuo catalítico Cys 115 del asa catalítica. La estabilización se da por la formación de múltiples puentes de hidrógeno entre los residuos 113-119 del asa catalítica y los residuos 329-330 del N-terminal. Los cambios producidos dan lugar a una molécula con un sitio activo cerrado (menor acceso al solvente) que aísla a los sustratos durante la catálisis. La liberación de los productos sería la vuelta a la forma abierta.

A manera de conclusión sobre los ejemplos anteriores podemos afirmar que, a pesar de ser tres sistemas distintos, los movimientos al interior de estas proteínas están relacionados con el cambio conformacional e indican que, aun en estado cristalino, los desplazamientos oscilan entre las conformaciones abierta y cerrada. Incluso en el cristal, las moléculas exploran un cierto número de sub-estados conformacionales, desviados del estado de mayor estabilidad o de menor energía. Oscilaciones o exploraciones de sub-estados que son parte importante de los procesos catalíticos. Ejemplo de ello es el trabajo de Schultz-Heienbrok *et al.*¹¹⁰, en el que han mostrado que la rotación de 96° del C terminal (cambio de conformación abierta a cerrada) de la 5'nucleotidasa de *E. coli* es necesaria para la actividad catalítica. Por medio de puentes disulfuro introducidos por ingeniería genética produjeron dos variantes: una que se mantiene en la forma abierta y otra que puede adoptar tanto la conformación cerrada como una semiabierta. Al realizar análisis cinéticos sobre las variantes observan que la atrapada en la conformación abierta es inactiva hasta que se reducen los puentes disulfuro que la estabilizan. La otra variante presenta la mitad de actividad, aún en la forma oxidada, aunque si presenta una mayor inhibición por sustrato, ambas con respecto a la silvestre. Esto puede explicarse por la flexibilidad que mantiene aun en presencia de los puentes disulfuro y que le permite moverse entre la forma semiabierta y cerrada. La estabilización de la forma cerrada es favorecida por la unión de ligandos que modifican las interacciones interatómicas tanto en los alrededores del sitio de unión como en regiones más alejadas¹¹¹. Esta estabilización debe ocurrir durante un periodo corto de tiempo para que la molécula pueda regresar a su conformación libre y repetir el ciclo catalítico. Este es un buen ejemplo de como se relaciona la flexibilidad estructural con la actividad biológica (enzimática en este caso particular) y establece la relevancia del obtener esa información a partir de los experimentos de difracción de rayos X.

6 CONCLUSIONES

1. Existe y es posible extraer información oscilatoria de estructuras cristalográficas de proteínas.

Podemos afirmar que aún en estado cristalino las oscilaciones atómicas se mantienen y es posible extraerlas si se les da un tratamiento adecuado a los datos experimentales de la difracción, es decir, si la resolución es suficiente y los estadísticos (en nuestro caso R, R_{free}, FOM) son congruentes con un buen modelo. La extracción de dicha información expande el nivel de descripción con el que se puede detallar una estructura obtenida por cristalografía de rayos X.

2. El PDB contiene un alto número de depósitos con baja calidad y/o incompletos, es decir, tienen una diferencia de $R - R_{free} > 5\%$, carecen de valor de R_{free} y/o de factores de estructuras; elementos que permiten un análisis más detallado de las estructuras.

Parte esencial del proceso es el análisis de la calidad de los modelos. Gran parte de los depósitos derivados de las búsquedas en el PDB fueron excluidos porque no contaban con los requisitos mínimos de calidad: valores R, R_{free}, diferencia menor del 5 % entre ellos y/o factores de estructura. Sin los factores de estructura es imposible hacer el afinamiento TLS, mientras que valores elevados de los estadísticos generalmente derivan en modelos menos exactos.

3. Así como es indispensable contar con modelos de buena calidad es necesario tratar a cada estructura como sistema independiente.

No hay una metodología estándar para llegar directamente al mejor modelo posible, sino que hay que tratar a cada proteína como un sistema independiente.

4. En algunos casos, del afinamiento por método de TLS pueden resultar mejores modelos, incluso, que los obtenidos por afinamiento anisotrópico.

Los desplazamientos atómicos que se pueden observar por medio de la cristalografía están típicamente dominados por modos de frecuencia bajos correlacionados sobre grupos de átomos; de forma tal que el afinamiento anisotrópico individual puede ser, en muchos casos, una sobre parametrización del problema¹⁰⁸.

Un ejemplo de ello es la última estructura (UDP-N-acetilglucosamina enolpiruvatotransferasa) en la que el afinamiento TLS resulta ligeramente mejor que el anisotrópico. El método de tensores TLS reduce los parámetros al tratar grupos de átomos como cuerpos pseudo-rígidos sobre los que se permiten traslaciones y libraciones desde un punto fijo (o centro de reacción). Durante el afinamiento se ajusta los desplazamientos de los rígidos con los datos de difracción, dando como resultado tres tensores T, L y S que describen los desplazamientos anisotrópicos y conjuntos de los átomos del grupo. Es importante señalar que los tensores no describen modos normales individuales, sólo

aproximan los efectos de una colección de modos de baja frecuencia. La mejor descripción de los modelos se ve reflejada en la reducción, en todos los casos, de los valores de los estadísticos R y R_{free} así como un aumento de la FOM.

5. Las magnitudes obtenidas de los tensores TLS son una aproximación a la descripción de los modos o movimientos predominantes.

Si bien es cierto que las magnitudes de los tensores obtenidas no son equivalentes a las que esperaríamos encontrar en solución, sí son buenas aproximaciones a la dirección de los movimientos predominantes, en esta ocasión particular, los cambios de conformación.

6. La información oscilatoria es similar (se conserva) en estructuras homólogas.

La conservación de los patrones oscilatorios en estructuras homólogas como lo son la UDP-acetil enolpiruvatotransferasa de *E. coli* y de *E. cloacae* (con 21 % de identidad, aunque muy buena sobreposición) fortalece la hipótesis y es una evidencia más de la relación entre oscilaciones y función. Además de establecer una cuestión sobre una posible relación entre propiedades dinámicas y procesos evolutivos.

8 PERSPECTIVAS

Siendo esta la segunda tesis de Licenciatura que se hace en el laboratorio sobre este tema y con los mismos resultados, las perspectivas se proyectan hacia soportar o sustentar nuestra tesis con otros métodos experimentales o teóricos como lo son las dinámicas moleculares.

Quizá un último punto por hacer sería comparar el resultado de los afinamientos por TLS de las estructuras libres con los afinamientos de las estructuras con ligandos. Mostrando que la estabilización de la estructura cerrada (la reducción de los valores B en regiones móviles) corresponde con menores variaciones en los valores de R y R_{free} demostraríamos, de manera más contundente, que la inclusión de información oscilatoria mejora el modelo. Este punto no se realizó por la falta de factores de estructura depositados para las estructuras en conformación cerrada.

Otro posible campo de exploración para los TLS son estructuras derivadas de cristales de otras biomoléculas, tipo fragmentos de DNA (promotores, sitios de reconocimiento para factores de transcripción, etc.). Junto a ello la exploración de los factores transcripcionales o proteínas de unión a dichos segmentos tratando de correlacionar los modos oscilatorios y obtener información biológica sobre procesos tan importantes a partir de la difracción de los cristales.

Puesto que el afinamiento por tensores TLS es un proceso que requiere de poco tiempo de CPU (siempre que se seleccionen modelos adecuados) no suena una idea descabellada el que en un futuro las estructuras proteicas incluyan un análisis de las oscilaciones derivadas de los TLS. Ello

permitiría pasar de un modelo estático sobre el que se pueden hacer inferencias funcionales hacia otro en el que los desplazamientos explícitos las sustenten, ampliando el nivel descriptivo y aclarando los mecanismos moleculares.

9 REFERENCIAS

1. Frauenfelder, H., McMahon, B.H., Austin, R.H., Chu, K. and Groves, J.T. The role of structure, energy landscapes, dynamics, and allostery in the enzymatic function of myoglobin. *Proceedings of the National Academy of Science USA* **98**; 2370–2374 (2001).
2. Karplus, M. Aspects of Protein Reaction Dynamics: Deviations from Simple Behavior. *Journal of Physical Chemistry B* **104**; 11-27 (2000).
3. Schnell, J.R., Dyson, H.J., Wright, P.E. Structure, Dynamics, and Catalytic Function of Dihydrofolate Reductase. *Annual Review of Biophysics and Biomolecular Structure* **33**; 119-140 (2004).
4. Halle, B. Flexibility and packing in proteins. *Proceedings of the National Academy of Science USA*. **99**; 1274-1279 (2002).
5. Saenger, W. Structure and dynamics of water surrounding biomolecules. *Annual Review Biophysics and Biophysical Chemistry* **16**; 93-114 (1987).
6. Margolin, A.L. Novel crystalline catalysts. *Trends in Biotechnology* **14**; 223–230 (1996).
7. Milton, R. F., Dewan, J.C., Petsko, G.A. Effects of Temperature on Protein Structure and Dynamics: X-ray Crystallography Studies of the Protein Ribonuclease-A at Nine Different Temperatures from 98 to 320 K. *Biochemistry* **31**; 2469-2481 (1992).
8. Rudiño, E. 2001. Estudios estructurales sobre el mecanismo de activación alostérica de la Glucosmina-6-fosfato desaminasa. Tesis Doctoral por la Universidad Nacional Autónoma de México y el Instituto de Biotecnología. México.
9. González-Rubio, P. 2003. Vibraciones atómicas, flexibilidad conformacional y la función biológica de proteínas alostéricas: Estudio a partir de estructuras cristalográficas depositadas en el Protein Data Bank. Tesis de Licenciatura por la Universidad Nacional Autónoma de México y la Facultad de Ciencias. México.
10. Berman, H.M., Westbrook, J., Feng, Z., Gilliard, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., Bourne, P.E. The Protein Data Bank. *Nucleic Acids Research* **28**; 235–242 (2000).
11. Shomaker, V. and Trueblood, H.K. On the Rigid-Body Motion of Molecules in Crystals. *Acta Crystallographica B* **24**; 63-76 (1968).
12. Murshudov, G. N., Vagin, A. A. and Dodson, E. J. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallographica D* **53**, 240 (1997).
13. Rouseau, J.J. (1998). **Basic crystallography**. John Wiley & Sons. New York, USA. Pp. 426.
14. *Íbidem*.
15. <http://www.bookrags.com/biography/max-von-laue>
16. Blundell, T.L. and L.N. Johnson. 1976. **Protein Crystallography**. Academic Press, New York, USA.
17. Vöet, D and Vöet, J. 2004. Biochemistry. Volume 1 Biomolecules, Mechanisms of Action, and Metabolism 3rd Edition. John Wiley & Sons, U.S.A. p. 350.
18. Bergfors, T. 1999. Protein Crystallization. Techniques, Strategies, and Tips. Internacional University Line. USA pp 299.
19. Perutz, M. Early days of protein crystallography. pp.3-18. En *Methods in enzymology* Volume 114, Pages 3-588 (1985) Diffraction Methods for Biological Macromolecules Part A. Edited by: Harold W. Wyckoff, C.H.W. Hirs and Serge N. Academic Press.
20. Perutz, M. Early days of protein crystallography. p 3-18. En *Methods in Enzymology* Volume **114**, Pages 3-588 (1985). Diffraction Methods for Biological Macromolecules Part A. Edited by: Wyckoff, H. W., Hirs, C.H.W. and Timasheff, S. N.
21. agni.phys.iit.edu/~howard/itp/itpmacro.html
22. Kendrew, J.C., Dintzis, H.M., Parrish, R.G., Wyckoff, H., Phillips, D.C. A three-dimensional model of the myoglobin molecule obtained by X-ray analysis. *Nature* **181**; 662–666 (1958).
23. Perutz, M.F., Rossman, M.G., Cullis, A.F., Muirhead, H., Will, G., and North, A.C.T. Structure of haemoglobin. A three-dimensional Fourier synthesis at 5.5Å resolution, obtained by X-ray analysis. *Nature* **185**; 416–422 (1960).
24. agni.phys.iit.edu/~howard/itp/itpmacro.html
25. Rhodes, G. 2000. Crystallography; Made Cristal Clear. 2nd Edition. Academic Press U.S.A. p. 43.
26. <http://www.rcsb.org/pdb/holdings.html>
27. Kleywegt, G.J., and Jones, T.A. Where freedom is given, liberties are taken. *Structure* **3**, 535–540 (1995).
28. Rhodes, G. 2000. Crystallography; Made Cristal Clear. 2nd Edition. Academic Press U.S.A. p. 61.
29. <http://www-structure.llnl.gov/Xray/tutorial/spcgrps.htm>
30. International Tables for Crystallography. Volume A Space-Group Symmetry. 4th revised edition. Edited by Hahn, T. Kluwer Academic Press. Dordrecht, The Netherlands. 1996. pp 878.
31. http://www.chem.qmw.ac.uk/surfaces/scc/scat1_1b.htm

32. <http://hyperphysics.phy-astr.gsu.edu/hbase/quantum/bragg.html>
33. Rudiño, E. Comunicación personal.
34. Blundell, T.L. and L.N. Johnson. 1976. **Protein Crystallography**. Academic Press, New York, USA.
35. <http://www.ndted.org/EducationResources/CommunityCollege/Radiography/Physics/radmatinteraction.htm>
36. Rhodes, G. 2000. *Crystallography; Made Cristal Clear*. 2nd Edition. Academic Press U.S.A. p. 166.
37. Blundell, T.L. and L.N. Johnson. 1976. **Protein Crystallography**. Academic Press, New York, USA. p 1.
38. Merrit, E. Expanding the model: anisotropic displacement parameters in protein structure refinement. *Acta Crystallographica D* **55**; 1109-1117 (1999).
39. Kleywegt, G.J. Validation of protein crystal structures. *Acta Crystallographica D* **56**, 249–265 (2000).
40. Brünger, A.T. Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355**, 472-475 (1992).
41. Kleywegt, G.J., and Brünger, A.T. (1996). Checking your imagination: applications of the free R value. *Structure* **4**, 897–904.
42. Zwart, P. H. and Lamzin V. S. The influence of positional errors on the Debye effects. *Acta Crystallographica D* **60**, 220-226 (2004).
43. Kleywegt, G.J., and Jones, T.A. Phi/psi-chology: Ramachandran revisited. *Structure* **4**, 1395–1400 (1996).
44. Eric W. Weisstein et al. "Tensor" From-- <http://mathworld.wolfram.com/Tensor.html> -- A Wolfram Web
45. Winn, M.D., Isupov, M.N., Murshudov, G.N. Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallographica D* **57**; 122-133 (2001).
46. Sali, A., Veerapandian, B., Cooper, J. B., Moss, D. S., Hofmann, T. & Blundell, T. L. Domain flexibility in aspartic proteinases. *Proteins*, **12**, 158-170 (1992).
47. Shen, Y., Kong, Y., and Ma, J. Intrinsic flexibility and gating mechanism of the potassium channel KcsA. *Proceedings of the National Academy of Science U S A*. **99**; 1949-1953 (2002).
48. Brooks, B. and Karplus, M. (1985). Normal modes for specific motions of macromolecules: application to the hinge-bending mode of lysozyme. *Proc. Natl Acad. Sci. USA*, **82**, 4995-4999.
49. Cruickshank, D.W.J. The Analysis of the Anisotropic Thermal Motions of Molecules in Crystals. *Acta Crystallographica* **9**; 754-756.
50. Shomaker, V. and Trueblood, H.K. On the Rigid-Body Motion of Molecules in Crystals. *Acta Crystallographica B* **24**; 63-76 (1968).
51. Eric W. Weisstein. "Euler's Theorem." From <http://mathworld.wolfram.com--A Wolfram Web Resource>. <http://mathworld.wolfram.com/EulersTheorem.html>
52. Winn, M.D., Murshudov, G.N., Papiz, M.Z. Macromolecular TLS Refinement in REFMAC at Moderate Resolutions. Pp. 300-321. En *Methods in Enzymology* Volume **374**, pp. 1-681 (2003). Macromolecular Crystallography, Part D. Edited by: Carter, C. W., and Sweet, R. M.
53. Holbrook, S.R., Dickerson, R.E., Kim, S-H. Anisotropic Thermal-Parameter Refinement of the DNA Dodecamer CGCGAATTCGCG by the Segmented Rigid-Body Method. *Acta Crystallographica B* **41**; 255-262 (1985).
54. Schultz-Heienbrok, R., Maier, T. and Sträter, N. Trapping a 96 Degree Domain Rotation in Two Distinct Conformations by Engineered Disulfide Bridges. *Protein Science* **13**, 1811-1822 (2004).
55. Harris, G.W., Pickersgill, R.W., Howlin, B., Moss, D.S. The Segmented Anisotropic Refinement of Monoclinic Papain by the Application of the Rigid-body TLS Model and Comparison to Bovine Ribonuclease A. *Acta Crystallographica B* **48**; 67-75 (1992).
56. Wilson, M.A., Brünger, A.T. Domain flexibility in the 1.75Å resolution structure of Pb²⁺-calmodulin. *Acta Crystallographica D* **59**; 1782-1792 (2001).
57. Howlin, B., Butler, S.A., Moss, D.S. Harris, G.W., Driessen, H.P.C. TLSANL: TLS parameter-analysis program for segmented anisotropic refinement of macromolecular structures. *Journal of Applied Crystallography* **26**, 622-624 (1993).
58. Kleywegt, G.J. Use of Non-crystallographic Symmetry in Protein Structure Refinement. *Acta Crystallographica D* **52**, 842-857 (1996).
59. Kleywegt, G.J. and Read, R. J. *Structure* **5**, 1557-1569 (1997).
60. Rice, L.M. and Brünger, A. *Proteins Structure Function and Genetics* **19**, 277-290 (1994).
61. http://www.structmed.cimr.cam.ac.uk/Course/Basic_refinement/Refinement.html
62. Ayala M., Horjales E., Pickard M. A., Vazquez-Duhalt R. Cross-linked crystals of chloroperoxidase. *Biochemical and Biophysical Research Communication* **295**, 828-831 (2002).
63. Zaccai, G. How Soft is a Protein? A Protein Dynamics Force Constant Measured by Neutron Scattering. *Science* **288**; 1604-1607 (2000).
64. Creighton, T.E. Stability of folded conformations. *Current Opinion in Structural Biology* **1**; 5 (1991).
65. Smith, J.L., Hendrickson, W. A., Honzatko, R.B., and Sheriff, S. Structural Heterogeneity in Protein Crystals. *Biochemistry* **25**; 5018-5027 (1986).

66. Daniel, R.M., Dunn, R.V., Finney, J. L., and Smith, J.C. The Role of Dynamics in Enzyme Activity. *Annual Reviews on Biophysical and Biomolecular Structure* **32**; 69-92 (2003).
67. Levitt, M., Sander, C., and Stern P.S. Protein normal-mode dynamics: trypsin inhibitor, crambin, ribonuclease and lysozyme. *Journal of Molecular Biology* **181**; 923-947 (1985).
68. Weiss, M.A., Sauer, R.T., Patel, D.J., Karplus, M. Amino-Terminal Arm of the I Repressor: A 1H NMR Study. *Biochemistry* **23** ; 5090-5095 (1984).
69. Englander, S.W., and Kallabach, N.R. Hydrogen exchange and structural dynamics of proteins and nucleic acids. *Quarterly Reviews of Biophysics* **16**; 521-655.
70. Vöet, D and Vöet, J. 2004. Biochemistry. Volume 1 Biomolecules, Mechanisms of Action, and Metabolism 3rd Edition. John Wiley & Sons, U.S.A. p. 350.
71. Koshland, D.E. Application of a theory of enzyme specificity to protein synthesis. *Proceedings of the National Academy of Science USA* **44**; 98-105 (1958).
72. Frauenfelder, H., Petsko, G., & Tsernoglou, D. Temperature-dependent X-ray diffraction as a probe of protein structural dynamics. *Nature* **280**; 558-563 (1979).
73. T.A. Jones. A graphics model building and refinement system for macromolecules. *Journal of Applied Crystallography* **11**, 268-272 (1978).
74. Hayward, S. Berendsen, H.J.C. Systematic analysis of domain motions in proteins from conformational change: new results on citrate synthase and T4 lysozyme. *Proteins, Structure, Function and Genetics*, **30**, 144-154 (1998).
75. Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallographica D* **54**: 905-21 (1998).
76. G. N. Murshudov, A. A. Vagin and E. J. Dodson. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Cryst.* (1997). **D53**, 240-255.
77. Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallographica D*. **50**, 760-763 (1994).
78. Wynn, M. Comunicación personal.
79. Wilson M. A., Brunger A.T. The 1.0 angstrom crystal structure of Ca²⁺-bound calmodulin: an analysis of disorder and implications for functionally relevant plasticity. *Journal of Molecular Biology* **301**; 1237-1256 (2000).
80. Howlin, B., Butler, S. A., Moss, D. S., Harris, G. W. and Driessen, H. P. C.. *TLSANL*: TLS parameter-analysis program for segmented anisotropic refinement of macromolecular structures. *Journal of Applied Crystallography* **26**, 622-624 (1993).
81. McRee, D.E. XtalView/Xfit--A versatile program for manipulating atomic coordinates and electron density. *Journal of Structural Biology* **125**: 156-165 (1999).
82. Merritt, Ethan A. and Bacon, David J. *Raster3D*: Photorealistic Molecular Graphics. 505-524. *Methods in Enzymology* Volume **277**, Pages 3-664 (1997). *Macromolecular Crystallography Part B* Edited by: Charles W. Carter Jr. and Robert M. Sweet.
83. Kraulis, P.J. MOLSCRIPT: A Program to Produce Both Detailed and Schematic Plots of Protein Structures. *Journal of Applied Crystallography* **24**, 946-950 (1991).
84. Bertrand, J. A., Fanchon, E., Martin, L., Chantalat, L. Auger, G., Blanot, D., van Heijenoort, J. and Dideberg, O. "Open" Structures of MurD: Domain Movements and Structural Similarities with Folylpolyglutamate Synthetase. *Journal of Molecular Biology* **301**, 1257-1266 (2000).
85. Byrne CR, Monroe RS, Ward KA, Kredich NM. DNA sequences of the cysK regions of *Salmonella typhimurium* and *Escherichia coli* and linkage of the cysK regions to ptsH. *Journal of Bacteriology* **170**; 3150-3157 (1988).
86. Becker MA, Kredich NM, Tomkins GM. The purification and characterization of O-acetylserine sulfhydrylase-A from *Salmonella typhimurium*. *Journal of Biological Chemistry* **244**: 2418-27 (1969).
87. Kredich NM, Tomkins GM. The enzymic synthesis of L-cysteine in *Escherichia coli* and *Salmonella typhimurium*. *Journal of Biological Chemistry* **241**; 4955-65 (1966).
88. Alexander FW, Sandmeier E, Mehta PK, Christen P. Evolutionary relationships among pyridoxal-5'-phosphate-dependent enzymes. Regio-specific alpha, beta and gamma families. *European Journal of Biochemistry* **219**; 953-960 (1994).
89. Richardson JS. Handedness of crossover connections in beta sheets. *Proceedings of the National Academy of Science USA*. **73**; 2619-2623 (1976).
90. Burkhard P, Rao GS, Hohenester E, Schnackerz KD, Cook PF, Jansonius JN. Three-dimensional structure of O-acetylserine sulfhydrylase from *Salmonella typhimurium*. *Journal of Molecular Biology*. **283**; 121-133 (1998).

91. Burkhard, P., Tai, C-H., Jansonius, J. N. and Cook, P. F. Identification of an Allosteric Anion-binding Site on O-Acetylserine Sulfhydrylase: Structure of the Enzyme with Chloride Bound. *Journal of Molecular Biology* **303**, 279-286 (2000).
92. Burkhard, P., Tai, C-H., Ristroph, C. M., Cook, P. F. and Jansonius, J. N. Ligand Binding Induces a Large conformational Change in O-Acetylserine Sulfhydrylase from *Salmonella typhimurium*. *Journal of Molecular Biology* **291**, 941-953 (1999).
93. Wickner W, Driessen AJ, Hartl FU. The enzymology of protein translocation across the *Escherichia coli* plasma membrane. *Annu Rev Biochem.* **60**; 101-24 (1991).
94. Lill R, Dowhan W, Wickner W. The ATPase activity of SecA is regulated by acidic phospholipids, SecY, and the leader and mature domains of precursor proteins. *Cell* **60**; 271-280 (1990).
95. Hartl FU, Lecker S, Schiebel E, Hendrick JP, Wickner W. The binding cascade of SecB to SecA to SecY/E mediates preprotein targeting to the *E. coli* plasma membrane. *Cell* **63**; 269-279 (1990).
96. Benach, J., and John F. Hunt Shape-shifting protein channel. *Nature* **427**; 24-26 (2004).
97. Lill R, Cunningham K, Brundage LA, Ito K, Oliver D, Wickner W. SecA protein hydrolyzes ATP and is an essential component of the protein translocation ATPase of *Escherichia coli*. *EMBO J.* **8**; 961-966 (1989).
98. Schiebel E, Driessen AJ, Hartl FU, Wickner W. Delta mu H⁺ and ATP function at different steps of the catalytic cycle of preprotein translocase. *Cell* **64**; 927-939 (1991).
99. Van den Berg B, Clemons WM Jr, Collinson I, Modis Y, Hartmann E, Harrison SC, Rapoport TA. X-ray structure of a protein-conducting channel. *Nature* **427**; 36-44 (2004).
100. Osborne AR, Clemons WM Jr, Rapoport TA. A large conformational change of the translocation ATPase SecA. *Proceedings of the National Academy of Science USA* **101**; 10937-10942 (2004).
101. Walker, J.E., Saraste, M., Runswick, M.J. and Gay, N.J. Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO Journal* **1**, 945-951 (1982).
102. Jerry Eichler and William Wickne. Both an N-terminal 65-kDa domain and a C-terminal 30-kDa domain of SecA cycle into the membrane at SecYEG during translocation. *Proc Natl Acad Sci USA* **94**, 5574-5579 (1997).
103. Fekkes, P., de Wit, J.G., Boorsma, A., Friesen, R.H.E and Driessen, A.J.M. Zinc Stabilizes the SecB Binding Site of SecA. *Biochemistry* **38**, 5111-5116 (1999).
104. Schönbrunn, E. Sack, S., Eschenburg, S., Perrakis, A., Krekel, F., Amrhein, N., and Mandelkow, E. Crystal structure of UDP-N-acetylglucosamine enolpyruvyltransferase, the target of the antibiotic fosfomycin. *Structure* **4**, 1065-1075 (1996).
105. Walsh CT, Benson TE, Kim DH, Lees WJ. The versatility of phosphoenolpyruvate and its vinyl ether products in biosynthesis. *Chem Biol.* **3**; 83-91 (1996).
106. Eschenburg, S., Kabsch, W., Healy, M.L., and Schönbrunn, E.. A New View of the Mechanisms of UDP-N-Acetylglucosamine Enolpyruvyl Transferase (MurA) and 5-Enolpyruvylshikimate- 3-phosphate Synthase (AroA) Derived from X-ray Structures of Their Tetrahedral Reaction Intermediate States. *Journal of Biological Chemistry* **278**; 49215-49222 (2003).
107. Schönbrunn, E., Eschenburg, S., Krekel, F., Luger, K., and Amrhein, N. Role of the Loop Containing Residue 115 in the Induced-Fit Mechanism of the Bacterial Cell Wall Biosynthetic Enzyme MurA. *Biochemistry* **39**, 2164-2173 (2000).
108. Chaudhry, C., Horwich, A. L., Brunger, A. T. and Adams, P. D. Exploring the Structural Dynamics of the E. Coli Chaperonin GroEL Using Translation-libration-screw Crystallographic Refinement of Intermediate States *Journal of Molecular Biology* **342**, 229-245 (2004).
109. Rossmann, M. The Molecular Replacement Method. *Acta Crystallographica A* **46**, 73-82 (1990).
110. Schultz-Heienbrok, R., Maier, T. and Sträter, N. A Large Hinge Bending Domain Rotation Is Necessary for the Catalytic Function of *Escherichia coli* 5'-Nucleotidase. *Biochemistry* **44**, 2244-2252 (2005).
111. Freire, E. The propagation of binding interactions to remote sites in proteins: Analysis of the binding of the monoclonal antibody D1.3 to lysozyme. *Proceedings of the National Academy of Sciences USA* **96**, 10118-10122 (1999).

10 ÍNDICE DE FIGURAS

Figura 1: www.pilleux.cl/mt771/apuntes/02OndasyDifraccion/Cap02OndasyDifraccion.PDF

Figura 2: Merrit, E. Expanding the model: anisotropic displacement parameters in protein structure refinement. *Acta Crystallographica D* **55**; 1109-1117 (1999).

Figura 3: <http://www-structure.llnl.gov/Xray/resolution.html>

Figura 4: Shomaker, V. and Trueblood, H.K. On the Rigid-Body Motion of Molecules in Crystals. *Acta Crystallographica B* **24**; 63-76 (1968).

Figura 5: Winn, M.D., Murshudov, G.N., Papiz, M.Z. Macromolecular TLS Refinement in REFMAC at Moderate Resolutions. *Methods in Enzymology* **374** 300-321 (2003)p 304.

Figura 6: Chia-Hui Tai, Peter Burkhard, David GanjiThierry Jenn, Corey Johnson and Paul F. Cook, Characterization of the Allosteric Anion-Binding Site of O-Acetylserine Sulfhydrylase. *Biochemistry* **40**, 7446-7452 (2001).

Figura 7: Burkhard P, Rao GS, Hohenester E, Schnackerz KD, Cook PF, Jansonius JN. Three-dimensional structure of O-acetylserine sulfhydrylase from *Salmonella typhimurium*. *Journal of Molecular Biology*. **283**; 121-133 (1998).

Figura 8: Visual and Molecular Dynamics (VMD)

Figura 9: Swiss-PDBViewer v3.7

Figura 10: Swiss-PDBViewer v3.7

Figura 11: Swiss-PDBViewer v3.7

Figura 12: raster 3D

Figura 13: raster 3D

Figura 14: Visual and Molecular Dynamics (VMD)

Figura 15: a) Swiss-PDBViewer v3.7 b) Osborne AR, Clemons WM Jr, Rapoport TA. A large conformational change of the translocation ATPase SecA. *Proceedings of the National Academy of Science USA* **101**; 10937-10942 (2004).

Figura 16: Swiss-PDBViewer v3.7

Figura 17: Raster 3D

Figura 18: Molscript

Figura 19: Du, W., Brown, J. R., Sylvester, D. R., Huang, J., Chalker, A. F., So, C. Y., Holmes, D. J., Payne, D. J. and Wallis, N. G. Two Active Forms of UDP-N-Acetylglucosamine Enolpyruvyl Transferase in Gram-Positive Bacteria. *Journal of Bacteriology* **182**, 4146-4152 (2000).

Figura 20: Visual and Molecular Dynamics (VMD)

Figura 21: Schönbrunn, E. Sack, S., Eschenburg, S., Perrakis, A., Krekel, F., Amrhein, N., and Mandelkow, E. Crystal structure of UDP-N-acetylglucosamine enolpyruvyltransferase, the target of the antibiotic fosfomicin. *Structure* **4**; 1065-1075 (1996).

Figura 22: Swiss-PDBViewer v3.7

Figura 23: Swiss-PDBViewer v3.7

Figura 24: Raster 3D