Difracción de Rayos X de Proteínas

- Aspectos experimentales: equipos de Introducción difracción y procesado de datos. **5**. 1. 2. Simetripoducción de rayos X. Cristales y su obtención. 3. Dimoc Generadores de Kayos X: 4. Aspectos experimentales, equipos de difracción y procesado de datos. 5. El Problema de la Fase: métodos de Detectores de rayos X. resolución estructural. 6.
- 7. Refinamiente, experieded y validación estructural.
- 8. Aptica diánes a da de datos de difeta de informa de la companya de la companya

FUENTES DE RAYOS X



Sincrotrón.

- Policromático.
- Sintonización de la λ .
- Brillo altísimo.



Ánodo rotatorio.

- Monocromático.
- λ = 1.5418 Å
- Brillo moderado.





FUENTES DE RAYOS X



Producción de los Rayos X.





Distribución de longitudes de onda de los rayos X que se producen en tubos convencionales de rayos X con ánodo de cobre (Cu), o molibdeno (Mo).Sobre espectro contínuo aparecen las líneas características K-alpha y K-beta.

FUENTES DE RAYOS X



Anodo Rotatorio.





- Monocromático (λ = 1.5418 Å).
- Intensidad moderada.

SINCROTRÓN



Electrons emitted by an electron gun are first accelerated in a linear accelerator (linac) and then transmitted to a circular accelerator (booster synchrotron) where they are accelerated to reach an energy level of 6 billion electronvolts (6 GeV).

These high-energy electrons are then injected into a large storage ring --844 metres in circumference --where they circulate in a vacuum environment, at a constant energy, for many hours.

SINCROTRÓN















CMBE

LA LINEA DE RADIACIÓN SINCROTRON



Cabaña Óptica







Cabaña Experimental









Cabaña Experimental

Montaje automático de cristales en Sincrotrón.



SSRL



BM30A-ESRF











Cabina de Control



DETECTORES DE RAYOS X







CCD

- Área de detección pequeña.
- Tiempos de impresión y lectura cortos.
- goniómetro con varios grados de libertad.

Image Plate

- Área de detección grande.
- Tiempos de impresión y lectura largos.
- goniómetro con un único grado de libertad.

SISTEMAS DE DETECCION DE RAYOS X





CCD

ImagePlate



Axiom

MONTAJE EXPERIMENTAL DE CRISTALES DE PROTEINAS







Método oscilatorio



Las etapas fundamentales del análisis de datos

1. Indexación.

Inspección visual de las imágenes de difracción.

Auto indexado.

Determinación de los parámetros del cristal: celda unitaria, grupo de Laue probable, orientación del cristal, estimación de la mosaicidad.

Refinamiento de la geometría de la difracción.

2. Integración.

Integración de los picos de difracción, refinando simultáneamente la orientación del cristal, su mosaicidad y los parámetros del detector.

3. Ensamblado y escalamiento.

Conversión de los datos a una escala común.

Determinación de la simetría y ensamblado de las reflexiones relacionadas por simetría.

Sumatoria estadística y estimación de los errores.



Tres etapas





Programas para el procesamiento de datos

- 1 y 2. Indexación e integración
- En dos dimensiones: MOSFLM (paquete CCP4) http://www.mrclmb.cam.ac.uk/harry/mosflm/ Denzo (paquete HKL) http://www.hkl-xray.com/
- En tres dimensiones: d*TREK http://www.rigaku.com/software/dtrek.html XDS http://www.mpimf-heidelberg.mpg.de/~kabsch/xds/
- 3. Escalamiento
- Scala (paquete CCP4) http://www.ccp4.ac.uk/html/scala.html Scalepack (paquete HKL)



Select item			
Main menu	Edits allowed	y Select item	• [
	Processing params a. 0.00	Main menu	Min 1 Max 700 Cursor position
Read image	b : 0.00 c : 0.00	Read image	Colour VBlack on white Mag V x4 PS Zoom
Find spots	beta : 0.00 gamma : 0.00	Find spots Edit spots	
Edit spots	PsiX : 0.00 PsiY : 0.00 PsiZ : 0.00	Clear spots	
Clear spots	Mosaic : 0.000 Divh : 0.000	Select images	
Select images	Lambda : 1.542 Distance: 120.00	Estimate mosaicity	
Autoindex	Beam X : 90.00 Y : 90.00 CCOMEGA : 0.000	Predict Clear prediction	
Estimate mosaicity	ROFF 0.00 TOFF 0.00 YEON 1.0000	Adjust	
Predict	Pick area: X: 11 Y: 11	Refine cell Integrate	
Clear prediction	Int threshold: 20 Vector scale 1 Two theta 0.00	Strategy	
Adjust	Resolution 0.00 *SPOT SEARCH*	Find hkl	
Refine cell	Rmin 9.00 Rmax 81.00	Pick	
Integrate	X offset 0.00 Y offset 0.00 Min X size 0.50	Circles	Waiting for input
Strategy	Max X size 2.00 Min Y size 0.50 Max V size 2.00	Beam / backstop	
Keyword input	Min no of pix 6 X splitting 0.30	Save/Exit	iteres estimates -
Find hkl	AUTOINDEXING* Min I/siq(I): 20		
Pick	Prompts On	Pixel X, Y 1801 0 XC, YC mm 180.0 0.0	
Measure cell	After refinement No After integration No	Resolution 0.00 Indices 0 0 0 F Phi 0.00 width 0.00	
Circles	Timeout mode Off	Intensity O Sigma O	
Beam / backston		Spacing A 0.000 Average 0.0	
Dean / Dackstop		Rms 0.0 Number 0 Zoomfactor 0	
Save/Exit		Circle resolution A 0.0 0.0 0.0 0.0 Phi 0.00 0.75	
		Missets ThetaX, Y, Z 0.00 0.00 0.00	Blue: fulls, Yellow: partials, Red: overlaps Green: too wide in phi



Input reply				
List of possible Laue groups, sorted on penalty index. The lower the PENALTY, the better				
Only solutions with PENALTY less than 200 are listed, a complete list is given in the terminal window				
20 190 oC 61.80 265.83 61.83 87.3 101.9 103.3 c222, c2221				
19 190 mC 265.83 61.80 61.83 101.9 92.7 76.7 C2 18 190 oC 61.80 265.83 61.83 87.3 101.9 103.3 C222,C2221				
17 189 mC 265.83 61.80 61.83 101.9 92.7 76.7 C2 16 189 mC 61.80 265.83 61.83 92.7 101.9 76.7 C2				
15 189 mC 61.83 265.90 61.80 92.7 101.9 76.6 C2				
13 94 hP 61.80 61.83 129.33 90.0 90.1 101.9 P3, P31, P32, P312, P312, P3121, P3212, P3221				
P6, P61, P65, P62, P64, P63, P622, P6122, P6522, P6222, P6422, P6322 12 93 mC 61.80 126.31 129.33 90.0 90.1 73.3 C2				
11 93 mC 126.31 61.80 129.33 90.1 90.0 73.3 C2 10 34 tP 61.80 61.83 129.33 90.0 90.1 101.9 P4,P41,P42,P43,P422,P4212,P4122,P41212,P4222,P42212 P4322,P43212				
9 34 oP 61.80 61.83 129.33 90.0 90.1 101.9 P222, P2221, P21212, P212121 8 33 mP 61.80 61.83 129.33 90.0 90.1 101.9 P2.P21				
7 33 mP 61.80 61.83 129.33 90.0 90.1 101.9 P2,P21				
5 1 oC 77.87 96.03 129.33 89.9 90.0 90.0 c222,c2221				
4 1 mC 77.87 96.03 129.33 89.9 90.0 90.0 C2 3 1 mP 61.80 129.33 61.83 90.0 101.9 90.1 P2,P21				
2 U aP 61.80 61.83 129.33 90.0 89.9 78.1 P1 1 0 aP 61.80 61.83 129.33 90.0 90.1 101.9 P1				
Select a solution AND a spacegroup from list above (eg 3 p42) or 0 to abandon or T to change min I/sig(I):_				





1	Min 1Max 750Overlay () on	Cursor position
1 Partie	Colour 🛛 Black on white	Mag ⊽ x4 PS Zoom
		• •





1. Montaje y Transporte de cristales de proteína

2. Cargador de muestras:







LN2 Tank and

- Holds up to 50 samples stored in 5 baskets of 10 samples each
- MUST be used in conjunction with SPINE sample holder
- Integrated data matrix reader for sample tracking

SAMPLE CHANGER 4 baskets prototype

2. Cargador de muestras: Interface gráfica



3. Visualización de la muestra y centraje















- Assisted "3 Clicks" : Rapid and accurate!
- □ Automated crystal centering (~80% successful)



En un futuro muy muy cercano....

The Missing Link : ISPyB Pocket Sample

Links the Crystallization information system to the data collection information system





Biocristalografía: El experimento en el futuro?



Difracción de Rayos X de Proteínas

- 16. Introducióna de la Fase: métodos de
- 2. Simeasalución estructural.
- 3. stEllesoylemade en faten.
- Difracción de Rayos X. La función de Patterson.
 Aspectos experimentales: equipos de difracción y procesado de datos. Reemplazamiento multiple isomorfo (MIR).
 El Problema de la Fase: métodos de
 - Euclétodos MAD. tural.
- 7. Refinamiento, modelado y validación
- 8. Aplicación: relaciones estructura-función.



EL PROBLEMA DE LA FASE....





 $\rho (x y z) = 1/V \Sigma |F(h k I)| \exp \{-2\pi i (hx+ky+Iz)+ i \Phi (h k I)\}$
IMPORTANCIA DE LA FASES....





Efecto de las amplitudes de difracción

• La película muestra el efecto (casi imperceptible) de calcular un mapa de densidad electrónica con amplitudes (intensidades de de difracción) erróneas, es decir, muy mal medidas.

Las imágenes de esta película representan el cambio gradual de las amplitudes (desde valores razonables de desacuerdo con las teóricas, R=10%, hasta valores aleatorios, R=75%).Es interesante darse cuenta de que si las fases son correctas (tal como es el caso) el mapa apenas cambia hasta llegar a valores de R=30% y que es todavía interpretable hasta valores próximos a R=50%.



Efecto de las fases

La película de la derecha muestra el efecto del cálculo de un mapa de densidad electrónica con fases mal estimadas. La "figura de mérito" de las fases (el coseno del error de la fase) se muestra como "m". Obsérvese la fuerte dependencia existente entre la bondad de las fases y la apariencia del mapa para poder reconocer la localización de los átomos.



EL PROBLEMA DE LA FASE.... ¿¿¿CÓMO RESOLVERLO???



Métodos de Faseado

Métodos directosρ≥0, átomos discretosRemplazamiento MolecularModelo homólogoRemplazamiento IsomorfoSubestructura de átomos pesadosDispersión AnómalaSubestructura de átomos anómalos

Modificación de Densidad (mejora de las fases) Aplanamiento de solvente *Histogram matching* Promediado simetría no-cristalográfica Estructura Parcial Extensión de fases



Arthur L. Patterson(1902-1966)

LA FUNCIÓN DE PATTERSON

Históricamente hablando, la primera solución al problema de las fases vino de la mano de Arthur Lindo Patterson. Basándose en la imposibilidad de resolver de un modo directo la función de la densidad electrónica, y tras su aprendizaje sobre convolución de transformadas de Fourier con el matemático Norbert Wiener, en 1934 Patterson introdujo una nueva función P(uvw). Esta nueva función, que define en un nuevo espacio (espacio de Patterson), puede considerarse sin exageración como el desarrollo singular más importante para la Cristalografía, tras el propio descubrimiento de los rayos X por Röntgen en 1895. Su elegante fórmula, conocida como la función de Patterson, supone una simplificación de la información contenida en la función de densidad electrónica, ya que suprime la información de las fases, y los módulos de los factores de estructura se sustituyen por sus cuadrados. Es, pues, una función que puede calcularse de inmediato a partir de la información experimental de que se dispone (las intensidades, que a su vez se derivan de los módulos de los factores de estructura)

 $\rho(xyz) = (1/V) \Sigma\Sigma\Sigma [F(hkl)] \cos 2\pi (hx + ky + lz - \Phi(hkl))$ electrónica

Función de densidad

 $P(uvw) = (1/V) \Sigma\Sigma\Sigma [F(hkl)]^2 \cos 2\pi (hu + kv + lw))$

Función de Patterson

LA FUNCIÓN DE PATTERSON





LA FUNCIÓN DE PATTERSON-PROPIEDADES





LA FUNCIÓN DE PATTERSON PROPIEDADES







 $Z_i Z_j \longrightarrow$ facilidad para detectar átomos pesados

LA FUNCIÓN DE PATTERSON PROPIEDADES





La simplificación es consecuencia de la pérdida de información que ocurre al pasar de:



uvw + uvw

REMPLAZAMIENTO ISOMORFO





- |F_P| Amplitudes cristal nativo
- |F_{PH}| Amplitudes cristal con derivado
- |F_H| Amplitudes átomo pesado



REMPLAZAMIENTO ISOMORFO, SIR





La fase α_{P} tiene una ambigüedad

REMPLAZAMIENTO ISOMORFO, MIR





La adición de un segundo derivado elimina la ambigüedad

EL REMPLAZAMIENTO ISOMORFO



Incluir átomos pesados en el cristal nativo que queden en posiciones fijas de la molécula sin deformar ni la celdilla ni la conformación de la proteína.



ii LAS DIFERENCIAS EN INTENSIDADES DEBEN SER EXCLUSIVAMENTE DEBIDAS A LOS ATOMOS PESADOS !!

MÉTODO:

- 1. Preparación de, al menos, un derivado (Hg, Au, Pt, U, Sm...Xe,Kr...Br,I).
- 2. Toma de datos de difracción para nativa y derivados.
- 3. Aplicación de función de Patterson y obtención coordenadas del metal.
- 4. Refinamiento parámetros átomo pesado y cálculo ángulos de fase.
- 5. Cálculo de mapas de densidad electrónica de la proteína.

EL REMPLAZAMIENTO ISOMORFO





INCONVENIENTES:

- **1.** Errores en $|F_{PH}|$ y $|F_{P}|$.
- **2.** No isomorfismo (4% < d_{min}).
- 3. Desorden en sitios minoritarios.
- 4. Escalado.



DISPERSIÓN ANÓMALA

VENTAJAS:

- 1. No hay problemas de isomorfismo.
- 2. Da mejores resultados a alta resolución

DISPERSIÓN ANÓMALA:



Espectro de fluorescencia de un cristal de proteína con Se-Met.





DISPERSIÓN ANÓMALA



MÉTODO:

1. Inclusión de dispersores anómalos en la estructura.



- 2. Medida del espectro de absorción.
- 3.Toma de datos de difracción de rayos X a diferentes longitudes de onda.
- 4. Medida de los pares de Friedel { (h k l), (-h, -k -l)}

Categoría	Dispersor Anómalo	
Metaloproteínas		
metales de transición	Fe. Cu. Zn. Mn	
otros metales	Ca. Mo	
Remplazamiento de metales		
Ca2+. Mg2+ por Lantánidos	Tb.Ho.Yb	
Zn por Mercurio	На	
Complejos con átomos pesados	5	
derivados comunes	Pt.Au.Ha.Pb.W.U	
Compuestos con "cluster"	Ta.W	
Proteínas modificadas		
Selenometionina o selenocisteír	na Se	
Telurometionina	Те	
nucleótidos Bromados o Iodado	s Brl	
Crioprotectorescon haluros		
Selenometionina o selenocisteír	na Brl	

REEMPLAZAMIENTO MOLECULAR





Hemoglobin (ascaris suum)



Myoglobin (sperm whale)





Hemoglobin (*lucina pectinata*)



Leghemoglobin (lupinus luteus)



Hemoglobin (glycera dibranchiata) Hemoglobin (urechis caupo)

Familia Estructural



Percentage identity matrix

1flp	100
1ithb	22.8 100
2gdm	21.2 18.0 100
1mbc	18.5 15.1 17.0 100
2hbg	24.4 23.1 19.9 22.1 100
1ash	13.3 10.1 15.8 15.9 14.6





- Completitud y calidad de los datos.
- Homología entre el modelo molecular y las moléculas reales que constituyen el cristal (> 30%).
- Tamaño del modelo molecular respecto al contenido de la celdilla.

REEMPLAZAMIENTO MOLECULAR





Función de Rotación

 $\mathsf{R}(\alpha,\beta,\gamma) = \bigcap_{r=1}^{n} \varPhi_{1}(u) \ge \mathsf{P}_{2}(u_{r}) \, \mathrm{d}u$

Obtención del modelo inicial











ρ (x y z) = f (I_{exp} , ϕ)





2. d. sen θ = n. λ









 ρ (x y z) = f (I_{calc}, ϕ_{calc})









siempre

 ρ (x y z) = f (I_{calc}, ϕ _{calc})





siempre

 ρ (x y z) = f (I_{calc}, ϕ_{calc})



CE CON

siempre

ρ (x y z) = f (I_{calc}, $\phi 1_{calc}$)

ρ (x y z) = f (I_{exp}, ϕ)

El problema de las fases



 ρ (x y z) = f (I_{exp}, ϕ)



siempre

ρ (x y z) = f (I_{calc}, $\phi 2_{calc}$)



$\rho (x y z) = f (I_{exp}, \phi)$



ρ (x y z) = f (I1_{exp}, ϕ)



 ρ (x y z) = f (I1_{exp}-I2_{exp}, ϕ)





siempre

ρ (x y z) = f (I_{calc}, ϕ_{calc})

ρ (x y z) = f (I_{exp}, ϕ)

El problema de las fases


siempre

ρ (x y z) = f (I_{calc}, ϕ_{calc})

 ρ (x y z) = f (I_{exp}, ϕ)