

Capítol 2

Objectius

Tots els estudis previs que s'han fet sobre l'estructura dels connectors eren a partir de microscòpia electrònica o de força atòmica. Aquestes proporcionen una idea general de l'envolupant de la proteïna, però no permet conèixer-ne l'estructura i les interaccions a nivell atòmic que són indispensables per al coneixement del seu mecanisme.

Els objectius plantejats en començar aquesta tesi van ser:

1. Obtenció de l'estructura tridimensional a alta resolució del connector del fag $\Phi 29$ mitjançant les tècniques de la cristal·lografia de proteïnes.

Com que l'únic que es coneixia eren models de baixa resolució obtinguts a partir de microscòpia electrònica, això obligava a:

2. Elaborar un mètode de reemplaçament molecular i extensió de fases per a la resolució d'estructures d'alta simetria a partir de models de molt baixa resolució.

Paral·lelament s'intentaria resoldre l'estructura per tècniques convencionals:

3. Resoldre l'estructura a partir de cristalls derivats fent servir les tècniques del MIR i del MAD.

Per últim, l'estructura a nivell atòmic permetrà:

4. Proposar un mecanisme que expliqui l'empaquetament del DNA a l'interior de la càpsida dels virus bacteriòfags.

Part II

Materials i Mètodes

Capítol 3

La Cristal·lografia de proteïnes

3.1 Definicions

3.1.1 Els cristalls

Un cristall és un ordenament periòdic de molècules. La unitat que es repeteix, per translació, és la cel·la unitat. La manera d'ordenar-se les molècules en una cel·la unitat es diu empaquetament. A causa de la diversitat d'empaquetaments possibles, una cel·la unitat té simetries (rotacions, translacions, inversions, etc.) caracteritzades en 230 grups espacials. Aquestes simetries divideixen la cel·la unitat en unitats asimètriques. Totes aquestes simetries són globals per al cristall, i el defineixen com a tal, però a més, en una unitat asimètrica, hi poden haver altres simetries locals, que només són vàlides dins d'ella. Aquestes simetries s'anomenen *simetries no cristal·logràfiques* (NCS). Generalment hi són per característiques pròpies de la molècula. És la unitat d'empaquetament que ha format el cristall (fig. 3.1).

3.1.2 Els cristalls de proteïna

Fins l'any 1934 es creia que les proteïnes eren objectes globulars sense estructura interna. Les primeres difraccions, eren de molt baixa resolució i suggerien un ordenament en fibra. Aquest any J.D. Bernal i D.Crowfoot varen poder obtenir imatges de difracció de raigs X de prou resolució com per demostrar que una proteïna tenia una estructura interna [41].

Les proteïnes són compostos formats per una seqüència d'aminoàcids (fig. 3.2 i taula 3.1) [20] que es pleguen formant una estructura secundària, principalment hèlixs α i fulles β . Aquesta estructura secundària es replega en l'anomenada terciària que és el que dóna a les proteïnes la seva forma i funció característica. Moltes vegades una proteïna necessita estar unida amb altres

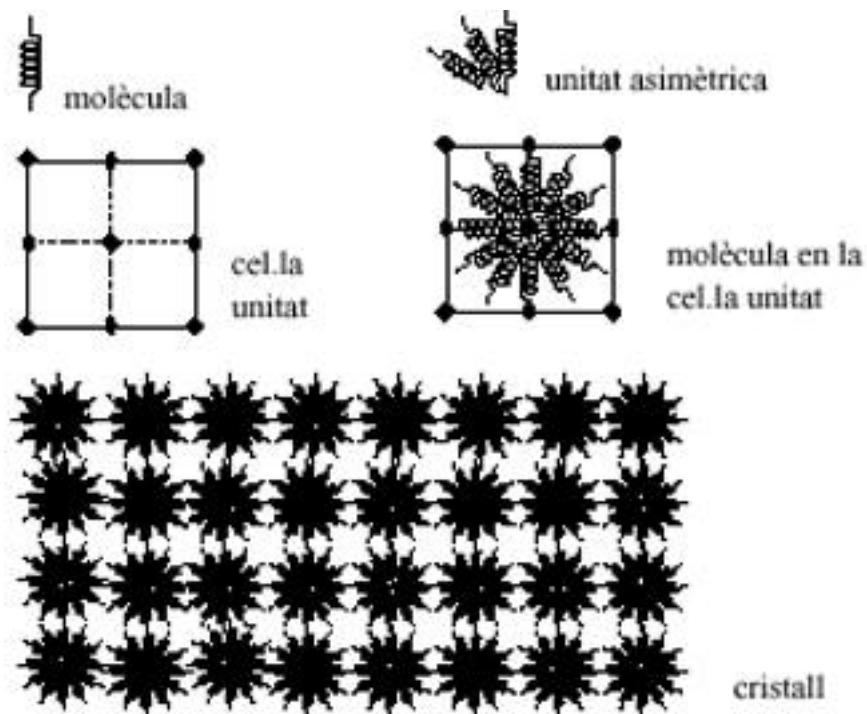
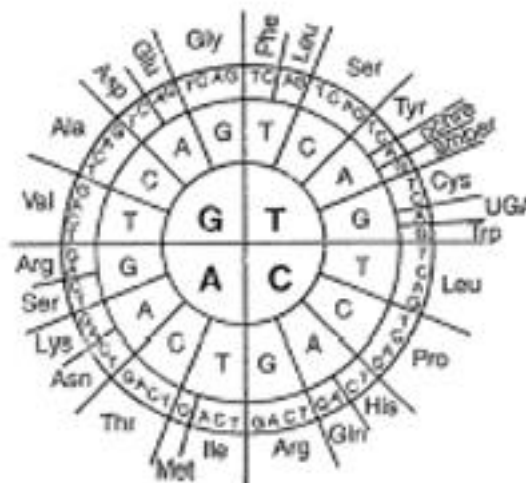


Figura 3.1: Simetries dels cristalls.

per fer la seva funció. Aquesta es la seva estructura quaternària. Veurem que per al cas del connector del $\Phi 29$, són necessàries 12 còpies de la proteïna p10 per formar la unitat funcional.

Un cristall de proteïna està format, a més del material proteic, per una quantitat gran de solvent, generalment de l'ordre del 50%, tot i que és molt variable. Segons les condicions de cristallització i segons el mètode emprat, també s'hi poden trobar agents cristallitzadors (MPD, glicerol, etc. fig. 3.10, plana 55), i àtoms pesants afegits per ajudar a resoldre l'estructura, o bé que formen part del complex actiu. En altres casos, òbviament també hi poden haver presents altres compostos amb els que es vulguin estudiar les interaccions amb la proteïna, com fàrmacs, DNA o RNA. Quan s'introdueixen àtoms pesants per ajudar a resoldre una estructura, els cristalls resultants s'anomenen derivats, respecte del cristall format només per la proteïna pura, que s'anomena natiu. Aquesta gran quantitat de solvent provoca moltes vegades que un cristall de proteïna sigui més aviat esponjós. Es degrada

Figura 3.2: El codi genètic. A la perifèria del cercle hi ha els noms dels aminoàcids, al centre el de les bases (seqüència del DNA). El DNA es transcriu directament a RNA missatger, que està format pel mateix conjunt de bases, però amb U en lloc de T. El Ribosoma és una gran proteïna que té per funció traduir l'RNA missatger a proteïna. El codi genètic es llegeix radialment des del centre del cercle i cada grup 3 bases dóna lloc a un aminoàcid. UGA, ochre i amber són combinacions de bases que no corresponen a cap aminoàcid i marquen el final del codi d'una proteïna.



aminoàcids						bases	
Alanina	Ala	A	Leucina	Leu	L	Adenina	A
Arginina	Arg	R	Lisina	Lys	K	Tirosina	T
Asparagina	Asn	N	Metionina	Met	M	Guanidina	G
Aspàrtic	Asp	D	Fenilalanina	Phe	F	Citosina	C
Cisteïna	Cys	C	Prolina	Pro	P	Uracil	U
Glutamina	Gln	Q	Serina	Ser	S		
Glutàmic	Glu	E	Treonina	Thr	T		
Glicina	Gly	G	Triptòfan	Trp	W		
Histidina	His	H	Tirosina	Tyr	Y		
Isoleucina	Ile	I	Valina	Val	V		

Taula 3.1: Aminoàcids i Bases. Esquerra: Nom dels aminoàcids, junt amb les abreviacions de tres lletres i d'una lletra. Dreta: Nom de les bases del DNA i de l'RNA i el seu codi d'una lletra. El DNA està format per ATGC i l'RNA per AUGC. Les tirosines es transcriuen directament a uracils.

molt ràpidament si no es troba en la solució on ha crescut, o bé congelat.

El nombre de grups espacials permesos en un cristall de proteïna és més reduït que en un cristall convencional, ja que no són permeses operacions d'inversió. Dels 230 grups espacials, només en són possibles 65.

3.2 Obtenció de la proteïna, dels cristalls i dels derivats

3.2.1 Obtenció de la proteïna p10

La proteïna es pot obtenir aïllant-la directament del seu organisme [11] o bé per sobreexpressió en organismes que porten un plasmidi recombinant [30] (fig. 3.3) (DNA circular no genòmic que conté el gen que la codifica). En tots dos casos cal purificar la proteïna. El procés realitzat per la proteïna del connector està explicat a la plana 77 [24].

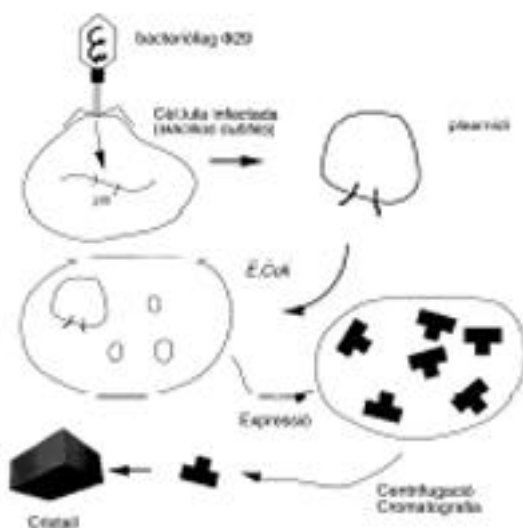


Figura 3.3: Obtenció de la proteïna. S'insereix el gen que codifica la proteïna en un plasmidi, que és DNA circular no genòmic que contenen els bacteris i se l'expressa en una cèl·lula. La proteïna clonada s'aïlla de la cèl·lula per diversos mètodes. Es purifica i se n'obtenen els cristalls posant-la en una solució sobresaturada, junt amb altres agents que ajudaran a la cristal·lització.

3.2.2 Els cristalls

Molt breument, per a obtenir cristalls cal tenir una mostra de proteïna molt pura, i deixar-la en repòs, en una solució sobresaturada, que conté també agents precipitants que ajudaran a la oligomerització de la molècula, i a la posterior formació del cristall. La solució es manté segellada per mantenir un equilibri entre les fases líquida i gasosa. Cal notar que, així com a la preparació de la proteïna es fan servir tècniques de biologia molecular més o menys establertes, amb un coneixement força bo dels mecanismes que hi participen; la cristal·lització és un fenomen encara poc conegut com per poder establir tècniques de cristal·lització amb alta fiabilitat. Per això, per a obtenir

cristalls s'han de fer servir cribatges de conjunts d'agents cristal·litzadors que han funcionat en altres proteïnes de característiques similars. Hi ha cribatges generals, on es varien les condicions al llarg d'un gran rang, i d'altres més precisos que tracten d'afinar condicions prometedores.

El número de cribatges necessaris per obtenir cristalls pot variar molt segons la “dificultat” dels cas. Generalment passen mesos fins a obtenir cristalls analitzables. A més s'ha d'afegir el temps necessari per a tal que es desenvolupin els cristalls. Un cristall pot necessitar fins i tot anys per créixer correctament. El temps d'espera habitual és de l'ordre d'una a poques setmanes. Per suposat també es pot donar el cas que el cristall es degradi si està massa temps en la solució. En el nostre cas, els cristalls es degradaven als dos dies, després de créixer molt poc.

3.2.3 Els derivats

Una estructura pot ser resolta de moltes maneres en cristal·lografia de raigs X. Quan no es coneix cap estructura similar, cal fer servir mètodes basats en la difracció més intensa, o bé anisotròpica, d'àtoms pesants (secció 3.5.1, plana 42).

Com que la major part de les proteïnes no contenen àtoms pesants amb aquestes propietats, cal introduir-los expressament. Una de les tècniques és la de xopar el cristall en una solució que contingui l'ió pesant. D'aquesta manera l'ió pesant es podrà difondre en el cristall i fixar-se en llocs on hi trobi afinitat per fer-ho. L'avantatge d'aquest mètode és que la geometria de la cel·la unitat no sol variar, condició indispensable per a segons quin mètode de resolució es vulgui afrontar. El principal problema és que no sempre l'ió pot penetrar en el cristall i situar-se ordenadament. De vegades queda desordenat en les regions de solvent, o de vegades arriba a llocs d'unió en molt poques cel·les (ocupació baixa).

Una altra tècnica és la *co-cristal·lització*. Aquest cop, proteïna i ió pesant s'intenten cristal·litzar a la vegada. Afavorint així que les unions es formin en la solució, abans de formar-se el cristall. L'avantatge és que augmenta la probabilitat de formar compostos ió pesant-proteïna, gràcies a l'agitació tèrmica. El problema és que segurament caldrà trobar condicions de cristal·lització noves. Un altre problema és que potser cristal·litza amb un empaquetament diferent i els cristalls no es poden fer servir per comparar amb els cristalls nadius.

Una tercera tècnica consisteix a substituir certs aminoàcids per altres que han estat modificats per contenir un àtom pesant. El cas més típic és el de la seleni-metionina (SeMet), on el sofre de les metionines se substitueix per seleni, un àtom que difracta anòmalament a una longitud d'ona molt conve-

nient per a cristalls de proteïna i, a més, típica en sincrotrons. L'avantatge d'aquest mètode està en que la localització dels àtoms pesants és molt precisa, i aquests estaran, en principi, idènticament presents en cadascuna de les proteïnes (ocupació alta). La dificultat consisteix, és clar, a obtenir proteïna així modificada, i que aquesta cristal·litzi.

G.E	Ió pesant	Línia	Data	Referència
<i>C</i> 2		X31	1997	[24], plana 78
<i>P</i> 4 ₂ 2 ₁ 2		X31,X11,id2,5.2R	1998-1999	[24], plana 78
<i>P</i> 4 ₂ 2 ₁ 2	<i>Ta</i> ₆ <i>Br</i> ₁₂ ²⁺	X31,X11,id2	1998-1999	[24], plana 78
<i>I</i> 422		bm30,id2,id14EH2 id14EH3,id14EH4,bm14	1999	3.5, plana 36
<i>I</i> 422	<i>Ag</i> ⁺	id2	1999	3.6, plana 37
<i>I</i> 422	<i>Au</i> ⁺	id14EH2	1999	3.7, plana 38
<i>I</i> 422	<i>Pt</i> ²⁺	id14EH2	1999	3.8, plana 39
<i>I</i> 422	<i>Ir</i> ⁺	id14EH2	1999	
<i>I</i> 422	Tk (<i>Hg</i> ²⁺)	id14EH2	1999	
<i>I</i> 422	<i>Ta</i> ₆ <i>Br</i> ₁₂ ²⁺	bm30, id14EH4, bm14	1999	3.9, plana 40
<i>I</i> 422	<i>Hg</i> ²⁺	id14EH2	1999	[42], plana 102
<i>I</i> 422	SeMet	id14EH4, bm14	2000	

Taula 3.2: Llista dels cristalls analitzats. Tipus de cristalls del connector analitzats per difracció de raigs X, ió pesant i línia de difracció usada. G.E: grup espacial. Les línies de difracció estan llistades a la taula 3.3. Les referències corresponen a les taules on hi ha les estadístiques de cada cristall processat. Les taules dels cristalls fets servir als resultats estan referenciades cap a l'article corresponent. El cristall que es va fer servir a l'afinament és el I422 Hg, tractat com natiu.

3.3 Raigs X: generació, òptica, interacció amb els cristalls i detecció

3.3.1 Generació

Per a l'anàlisi de cristalls de proteïna per difracció de raigs X, aquests es generen de dues maneres: amb generadors d'ànode rotatori i amb la radiació de sincrotró.

Un generador d'ànode rotatori, [20] consisteix en una cambra al buit, on un filament (el càtode) emet electrons que són accelerats per una diferència de

potencial, fins que xoquen amb un ànode fet d'un material adequat per a que emeti raigs X a la longitud d'ona necessària per a l'experiment. Generalment, per a proteïnes, l'ànode és de coure, que té una K_α d'1.5418Å. Aquest procés és molt ineficient, ja que la majoria de l'energia consumida es transforma en calor. Per evitar que l'ànode es fongui, se'l manté rotant per anar canviant el punt d'incidència dels electrons i se l'envolta amb un circuit de refrigeració. L'avantatge és que l'aparell té unes dimensions raonables per instal·lar en un laboratori. És adequat per a cristalls relativament grans, amb cel·les relativament petites. El problema és l'elevada quantitat de temps necessari per realitzar un experiment (d'uns pocs dies a setmanes), a causa de la poca intensitat del feix de raigs X. Un altre problema és que no es pot canviar la longitud d'ona, per la qual cosa no són possibles, en general, experiments de difracció anòmala.

La radiació de sincrotró [58] és la que es desprén en fer canviar de direcció partícules carregades que estan viatjant a velocitats properes a la de la llum. Les partícules carregades, en un sincrotró típicament electrons o positrons, circulen per un tub poligonal que està al buit. En els vèrtexs d'aquests polígons, els *bending magnets* corben la trajectòria de les partícules, generant alhora els raigs X. Un dispositiu que genera raigs més intensos és l'*insertion device magnet* (ondulador), que s'intercala en les seccions rectes. En un ondulador, uns camps magnètics en sèrie fan oscil·lar els electrons generant una radiació a una energia relacionada amb el període de l'oscil·lació. Un *wiggler* és un ondulador on els camps magnètics són d'intensitat molt més alta. El resultat és la generació d'harmònics del període de l'ondulador. Aquests harmònics formen un espectre similar al del *bending magnet*, però d'intensitat proporcional al nombre d'imants.

En aquest treball s'ha fet servir únicament radiació de sincrotró, ja que la mida petita dels cristalls i la gran mosaïcitat, obligava a tenir feixos intensos, i de secció petita.

3.3.2 Òptica

Els raigs X surten tangencialment de l'anell del sincrotró amb un espectre d'energies molt gran. Abans que arribin als aparells per fer l'experiment, s'han de preparar adequadament.

L'òptica consisteix en una sèrie d'aparells situats tangencialment a l'anell. Principalment, els monocromadors filtren les longituds d'ona sobrants, i els col·limadors afinen el raig [20].

El típic monocromador és un cristall de silici. Com que l'angle de difracció varia amb la longitud d'ona, es pot seleccionar la longitud d'ona orientant aquest cristall convenientment. Hi ha diverses configuracions estudiades per

Nom	Tipus	Sincrotró	Detector
BW7A	mad	Doris	MarCCD
BW7B	fixe	Doris	Mar 345 Image Plate
X31	mad	Doris	Mar 345 Image Plate
X11	mad	Doris	MarCCD
id2	fixe	ESRF	Image Plate
id14EH2	fixe	ESRF	Quantum CCD
id13	mad	ESRF	Quantum CCD
id14EH3	fixe	ESRF	MarCCD/Mar 345 Image Plate
id14EH4	mad	ESRF	Quantum CCD
bm14	mad	ESRF	MarCCD/Mar 345 Image Plate
bm30	mad	ESRF	Mar 345
5.2R	fixe	Elettra	Mar 300 Image Plate

Taula 3.3: Línies de difracció. fixe/mad: longitud d'ona fixe/ajustable. Doris: Desy, Hamburg. ESRF: Grenoble. Elettra: Trieste. id: *insertion device*. bm: *bending magnet*. X, BW i R: *wiggler*.

millorar l'enfoc, monocromacitat i intensitat del feix. Típicament el cristall monocromador es corba mecànicament per a comprimir el feix, i se'n munten dos en seqüència per a resoldre millor la longitud d'ona i orientar el feix en direcció a l'estació experimental.

Un col·limador serveix per definir un feix el més paral·lel possible. Els més senzills consisteixen en un tub allargat per on passa el raig. Aquest tub té unes obertures (*slits*) ajustables que aturen els raigs més divergents. En calen almenys tres convenientment separats per aturar també els raigs difractats per les pròpies obertures. Aquestes obertures es poden ajustar per a definir la secció del raig.

La divergència és l'angle d'obertura que té el raig en sortir de les últimes obertures. Aquesta divergència ha de ser la menor possible, ja que provocaria reflexions escampades en el detector, a causa del diferent angle d'incidència dels raigs sobre el cristall.

3.3.3 Interacció amb els cristalls

Els raigs X interaccionen amb les partícules carregades elèctricament. A les energies típiques usades (9-12 KeV), corresponents a una longitud d'ona al voltant d'1Å, l'índex de refracció està molt a prop de la unitat [20]. És per això que els raigs X no es poden enfocar, i per tant no és possible la formació

directa de la imatge. El tipus d'interacció que promou la difracció, a les energies típiques, és l'*scattering Thomson*. És tracta de la dispersió elàstica dels fotons pels electrons dels àtoms.

A l'*scattering Thomson*, si la radiació incident no està polaritzada, té una intensitat I_i , i interacciona amb una partícula lliure de càrrega e i massa m , la distribució angular de la intensitat de la radiació emesa és:

$$I_{eTh} = I_i \frac{e^4}{m^2 r^2 c^4} \frac{1 + \cos^2(2\theta)}{2},$$

essent c la velocitat de la llum, i r i 2θ les coordenades polars del punt d'observació, respecte de la direcció incident del raig.

La relació és inversament proporcional al quadrat de la massa de la partícula. Els electrons són 2000 cops més lleugers que els protons, i el valor absolut de la seva càrrega és el mateix. Per tant, la contribució dels electrons és 4 milions de vegades major que la dels protons. La difracció estarà només relacionada amb la distribució electrònica dins del cristall.

En un cristall els electrons no estan aïllats, pertanyen a àtoms, que alhora pertanyen a molècules i a cristalls. Cada raig difractat interfereix amb els difractats pels altres electrons, d'una manera característica donada per la geometria de tot el cristall. En particular, l'ordre cristal·lí afavorirà unes poques direccions en les que es produiran una interferència constructiva dels raigs (*les reflexions*) i l'estructura de la molècula modularà les intensitats de cadascuna d'aquestes reflexions.

Si f_j , $\mathbf{r}_j$ i β_j són, respectivament, el factor de difusió atòmic, les coordenades i el tensor dels desplaçaments tèrmics de cadascun dels $j = 1..N$ àtoms de la cel·la unitat, i $\mathbf{h} = (\mathbf{k}_D - \mathbf{k}_I)/2\pi$ és l'anomenat vector recíproc que representa la diferència entre els vectors de propagació del raig difractat ($\mathbf{k}_D$) i l'incident ($\mathbf{k}_I$), la intensitat difractada és el mòdul al quadrat de [43]:

$$F_{\mathbf{h}} = \sum_{j=1}^N f_j e^{2\pi i(\mathbf{h}^T \mathbf{r}_j - \mathbf{h}^T \beta_j \mathbf{h})}$$

$$I_{\mathbf{h}} = |F_{\mathbf{h}}|^2$$

$F_{\mathbf{h}}$ s'anomena factor d'estructura o amplitud.

Si considerem un cristall com una distribució periòdica de densitat electrònica, també es pot posar $F_{\mathbf{h}}$ com:

$$F_{\mathbf{h}} = \int_V \rho(\mathbf{r}) e^{2\pi i \mathbf{h} \cdot \mathbf{r}} d\mathbf{r}$$

on la integració s'estén només al volum de la cel·la unitat.

El factor d'estructura és, doncs, la transformada de Fourier de la distribució electrònica en la cel·la unitat. Fent la transformada inversa, obtindrem el *mapa* de la densitat electrònica en la cel·la:

$$\rho(\mathbf{r}) = \frac{1}{V} \sum_{\mathbf{h}} |F_{\mathbf{h}}| e^{2\pi i \Phi_{\mathbf{h}}} e^{-2\pi i \mathbf{h} \cdot \mathbf{r}}$$

on s'ha posat $F_{\mathbf{h}} = |F_{\mathbf{h}}| e^{2\pi i \Phi_{\mathbf{h}}}$ per resaltar que en fer el mòdul al quadrat la part de l'exponencial desapareix i les intensitats observades no contenen la informació de la fase ($\Phi_{\mathbf{h}}$), necessària per construir els mapes. Aquest és l'anomenat *problema de les fases* (plana 42).

3.3.4 Detectors

Per detectar els raigs difractats, en el nostre camp s'han imposat els detectors del tipus *image plate* i del tipus *CCD*. Altres mètodes, fotogràfics, estan en desús o es fan servir en casos molt específics.

Ambdós detectors tenen un aspecte exterior similar, on una pantalla d'entre 15 i 30 cm de radi se situa en línia amb la sortida dels raigs X i la mostra. La pantalla està muntada en un suport que permet variar la distància, i de vegades també l'angle del detector respecte el feix incident. Entre la mostra i el detector hi ha un *beam stop*, una peça molt petita de plom que serveix per no saturar el detector amb els raigs que no han difractat.

Els detectors del tipus *image plate* consisteixen en un disc de fòsfor d'uns 30 a 35 cm de diàmetre que va recollint la direcció i l'energia cedida pels raigs difractats durant el temps de recollida d'una imatge. Abans de recollir el següent diagrama, un làser escaneja la superfície del detector i excita un altre cop les partícules de fòsfor. La llum emesa es recull amb un fotodetector i es transforma en una imatge bidimensional [20].

Els del tipus *CCD* també tenen per pantalla un disc de fòsfor de mida similar, però en aquest cas, el fòsfor no amagatzema l'energia dels raigs difractats, si no que la converteix en llum visible. Aquesta llum es condueix per fibres òptiques a un chip *CCD* (*charge coupled device*) que l'amagatzema en forma de càrrega elèctrica fins que es digitalitza al final de la recollida de cada imatge [13].

En l'estat actual dels detectors, els *image plates* proporcionen un àrea de recollida més gran, i això es converteix en més facilitat per discernir les reflexions (*overlapping*), ja que, a igual resolució, es pot muntar més lluny. Altrament, el rang dinàmic de la recollida (diferència entre el valor més baix i el més alt) és molt més alt que en un *CCD*, que a més, quan se satura pertorba els *píxels* del voltant. També el soroll de fons de l'*image plate* és

menor que en un *CCD*. Per altra banda, el temps de *readout* és molt menor en un *CCD* (de l'ordre de 2 segons en front de més d'1 o 2 minuts). En el cas de recollides en un sincrotró, on els temps d'exposició per imatge són de l'ordre de pocs segons, l'ús de *CCDs* es fa imprescindible per accelerar el procés de recollida. En un laboratori, a causa de la menor intensitat dels raigs X, les recollides són d'hores per imatge, per la qual cosa l'increment de temps que representa un *image plate*, un parell de minuts per imatge, no és significatiu.

Ambdues tecnologies són força noves i encara poden evolucionar molt. Els *CCDs* tendeixen a augmentar la seva mida en aparells múltiples, i els *image plates* tendeixen a ser més ràpids.

3.4 L'experiment

3.4.1 Congelació i muntatge dels cristalls

Per difractar un cristall, se l'ha de *muntar* en un dispositiu adequat durant el temps que estarà exposat als raigs X. Un cristall es pot col·locar en un capil·lar de vidre, junt amb part de la gota on ha crescut. Aquesta tècnica es fa servir quan els cristalls poden aguantar la radiació sense degradar-se. També s'ha de fer servir quan els cristalls no aguanten la congelació. Llavors cal tenir una gran quantitat de cristalls per poder fer l'experiment.

Generalment els cristalls no aguanten la radiació, i per evitar *cremar-los* s'han de muntar congelats. La tècnica consisteix en mantenir-los en un flux de nitrogen a uns 100K durant tot l'experiment. Per al transport dels cristalls també és convenient la seva congelació. En aquest cas, s'enllacen amb un *pin* que consisteix en una barra metàl·lica acabada en un petit llaç de niló de les dimensions adequades [19]. La barra té unes dimensions estàndard de 18mm i el llaç té un diàmetre d'entre 0.2 a 0.05 mm segons la mida del cristall. El cristall es *pesca* amb el llaç de la gota on ha crescut, i segons les condicions en què hagi crescut se'l banya en una solució crioprotectora. Llavors se'l congela per immersió ràpida (*flash freezing*) en nitrogen o en propà líquids. Els cristalls es transporten en un *dewar* ple de nitrogen líquid al lloc de l'experiment.

Un cristall congelat aguanta generalment de l'ordre de 100 imatges en un sincrotró. En canvi un cristall no congelat només aguanta de l'ordre d'un parell d'imatges (fig. 3.4).

A l'hora de l'experiment, el cristall es col·loca davant d'on sortirà el feix, junt amb el *pin* o el capil·lar, muntat en un goniòmetre per poder ajustar l'alineament del cristall davant del feix. És important que a mesura que el

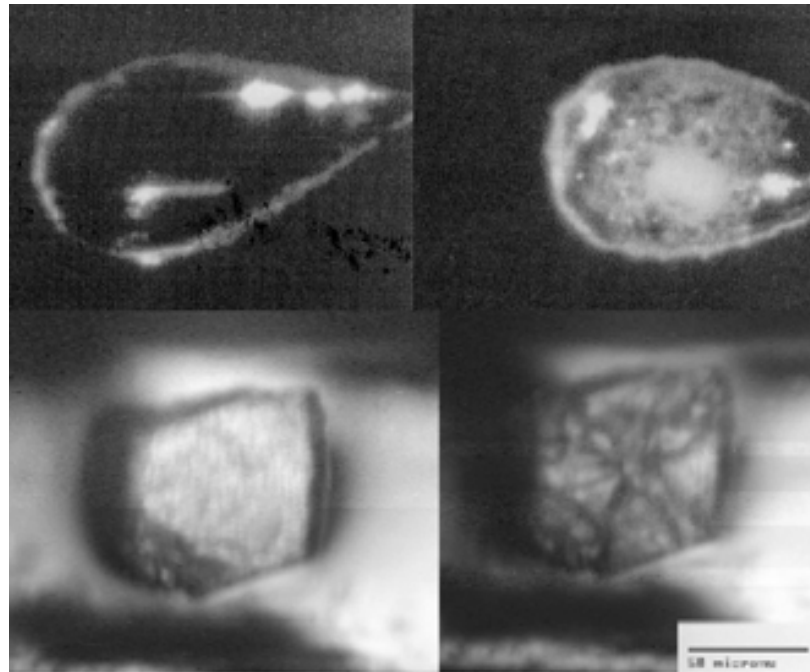


Figura 3.4: Efecte dels raigs X sobre els cristalls. Les dues imatges superiors corresponen a un dels cristalls que s'han fet servir per resoldre l'estructura. Està muntat en un loop, banyat en el criotampó que el protegeix i exposat a un flux de nitrogen a 100K. A l'esquerra, abans de començar l'experiment. El cristall és pràcticament transparent. A la dreta, després de recollir dos conjunts de dades, unes 200 imatges. El cristall està pràcticament cremat. El quadre blanc central és el punt d'incidència del feix. Les dues imatges inferiors corresponen a un cristall del virus HRV2 [56]. També abans i després de l'experiment, però el cristall no s'ha congelat, i només s'ha pogut recollir una imatge.

cristall rotí durant la difracció no es traslladi. D'aquesta manera les dades representaran una rotació completa. En el cas que hi hagués hagut translacions, s'haurien de processar les dades per separat, i probablement se'n perdrien moltes a causa dels ajustatges que caldria fer.

Quan es treu el cristall del seu contenidor de nitrogen líquid i es posa en el goniòmetre, moltes vegades hi ha un sobtat canvi de temperatura que pot formar cristalls d'aigua o de la solució que també difractaran, generalment en forma d'anells, i que segurament emascararan la difracció de la proteïna. Si passa això, de vegades es pot intentar solucionar tallant per un segon el

flux de nitrogen (*fast annealing*). Generalment això és suficient per a què la solució es torni a recongelar o vitrificar correctament. De vegades, a més, el canvi de temperatura ajusta l'empaquetament de les proteïnes en el cristall i s'obté una difracció de millor qualitat.

Hi ha moltes variants en aquesta tècnica. La més senzilla consisteix a interrompre el feix amb una targeta de plàstic. Últimament s'estan afegint dispositius mecànics als goniòmetres que permeten graduar el temps que s'interromp el flux de nitrogen amb un pedal, i a la vegada deixar les mans lliures per treballar.

Una altra tècnica, adequada si aquesta falla és la de reescalfar el cristall a temperatura ambient. Un cop congelat el cristall se'l torna banyar en la solució crioprotectora. L'escalfament sol trencar el cristall pels punts més febles, generalment els plans de macla [46].

No és freqüent tenir gaires cristalls per fer un experiment, per la qual cosa totes aquestes tècniques tenen el risc de malmetre del tot el cristall. Si un cristall difracta mitjanament bé, és millor assegurar abans un conjunt de dades abans d'intentar aquestes millores. La poca secció transversal d'un feix de raigs X de sincrotró pot permetre tornar a difractar un cristall per una altra zona.

3.4.2 Recollida i processament de les dades

La llei de la difracció prediu raigs difractats en moltes direccions, però un detector només recull les que es troben en una direcció determinada ("endavant"). En aquestes direccions els raigs són més intensos segons la llei de Thomsom, i per altres causes com l'absorció i la correcció de Lorentz [20]. Els raigs es projecten sobre el detector donant lloc a un patró característic que permet identificar el grup espacial. (A la plana 80 es pot veure un diagrama de difracció corresponent a un dels cristalls analitzats.)

A una donada orientació del cristall respecte el raig incident, i depenent de característiques del cristall (la mida de la cel·la unitat, la mida del propi cristall, la qualitat del cristall, etc.) i de l'experiment (mida del detector, distància, longitud d'ona, etc.), el nombre de reflexions mesurades sol ser baix i les mesures de cada reflexió no són completes: En un cristall de mida finita, les reflexions no són puntuals, s'extenen sobre un angle sòlid. A més a més, els petits desordres en l'orientació de les cel·les augmenten aquest angle (*mosaïcitat*). A causa d'això les reflexions no interferiran sempre constructivament i serà necessari rotar una mica el cristall per a recollir una reflexió sencera.

Per recollir un conjunt de dades complet d'un cristall (totes les reflexions necessàries per poder resoldre l'estructura) es recullen, doncs, dades en di-

I422							
Dmin	Nmes	Nref	I/sigma	Comp(%)	C.A.(%)	Mult	Rsym
6.64	12989	1955	10.3	99.3	99.3	6.6	0.058
4.70	24301	3405	10.1	100.0	99.7	7.1	0.060
3.83	31349	4353	11.0	100.0	99.9	7.2	0.055
3.32	36993	5090	6.3	100.0	99.9	7.3	0.076
2.97	42056	5724	6.6	100.0	99.9	7.3	0.112
2.71	46583	6336	3.9	99.9	99.9	7.4	0.187
2.51	50329	6840	2.5	99.9	99.9	7.4	0.289
2.35	38109	7315	1.8	99.7	99.9	5.2	0.378
2.21	26351	7744	0.5	99.2	99.8	3.4	0.521
2.10	18880	7712	1.4	94.1	98.9	2.4	0.505
Global	327940	56474	5.4	98.9	98.9	5.8	0.100

C2							
12.65	11178	3089	4.8	98.9	98.9	3.6	0.097
8.94	20880	5602	5.2	100.0	99.6	3.7	0.111
7.30	26903	7164	3.9	100.0	99.8	3.8	0.177
6.32	31833	8463	2.4	100.0	99.9	3.8	0.313
5.66	34956	9581	1.6	99.9	99.9	3.6	0.466
5.16	36951	10536	1.8	99.5	99.8	3.5	0.430
4.78	38830	11312	1.9	98.6	99.5	3.4	0.396
4.47	41067	12061	1.8	97.4	99.2	3.4	0.418
4.22	42244	12600	1.3	96.4	98.7	3.4	0.575
4.00	43043	13154	1.0	94.9	98.2	3.3	0.730
Global	327885	93562	2.3	98.2	98.2	3.5	0.316

Taula 3.4: Capes de reflexions. Estadístiques per capes del processament de les reflexions del cristall del tercer article (I422, plana 91), i d'un dels cristalls amb simetria C2. Dmin: resolució màxima de la capa. Nmes: Nombre de reflexions mesurades. Nref: Nombre de reflexions independents. I/sigma: $\langle I \rangle / \langle \sigma(I) \rangle$. Comp: Completesa. C.A.: Completesa acumulada. Mult: Multiplacitat. Rsym: Similitud entre reflexions (Veure text a plana 41)

ferents orientacions del cristall. El mètode més utilitzat és de la rotació, en què es recullen les dades mentre el cristall va girant al voltant d'un eix perpendicular al raig incident. Aquesta rotació és petita, de l'ordre d'un grau, i permet recollir tota la intensitat que ha *escampat* la mosaicitat. A continuació s'avança una quantitat similar a la de l'angle de rotació i es torna a recollir un altre cop oscil·lant al voltant d'aquesta orientació nova. Depenent

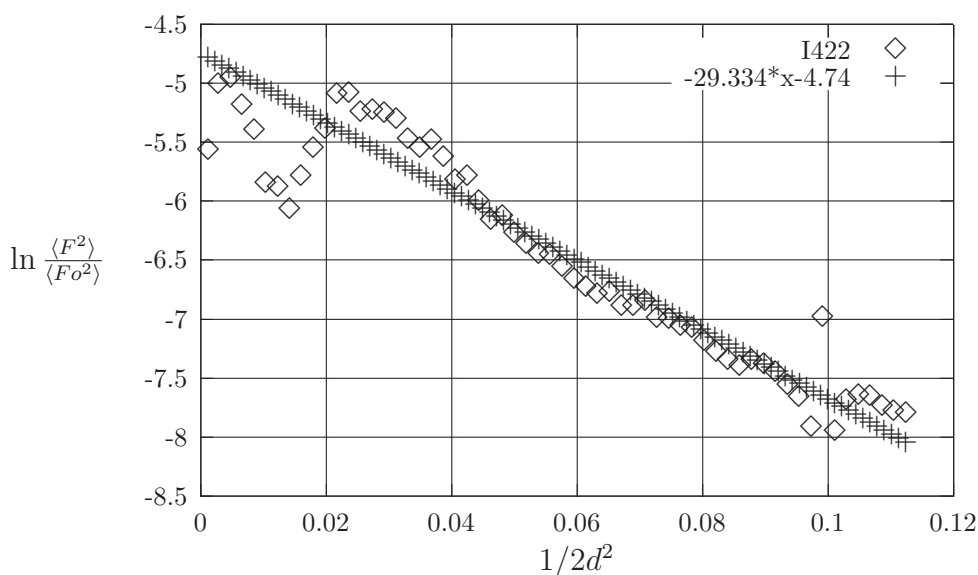


Figura 3.5: Wilson Plot. Ajustant per mínims quadrats una línia recta a la gràfica de $\ln(\frac{\langle F^2 \rangle}{\langle F_o^2 \rangle})$ enfront de $1/2d^2$ es pot calcular un B-factor mig pel cristall. Aquesta gràfica pertany al cristall del tercer article (I422, plana 91). En aquest cas $B = 29.33$.

del grup espacial i del tipus d'experiment, caldrà recollir entre 30 i 180 graus per a completar les dades.

L'experiment comença amb la recollida d'una o unes poques orientacions, les necessàries per determinar el paràmetres per la recollida completa: l'orientació del cristall, la cel·la unitat, la mosaïcitat, el rang d'angles que s'ha de cobrir i amb quina rotació. També s'ajusta la distància del detector al cristall, depenent de la resolució observada i de la qualitat de les reflexions de més alta resolució. Si són molt febles, és millor retirar el detector i recollir només les reflexions de baixa resolució amb bona qualitat i separació. També és important que les reflexions no es sobreposin, per la qual cosa s'ha de jugar amb la distància i la rotació. Aquestes reflexions no es podrien mesurar acuradament i s'haurien de descartar. Cal ajustar també el temps d'exposició de cada imatge, per tal de *resoldre* bé les reflexions i alhora no *cremar* ni el cristall ni saturar el detector.

Un cop preparades les característiques de l'experiment, es procedeix a la recollida de tot el conjunt de dades. Generalment es recullen els angles seqüencialment, per facilitar l'escalat de les dades, però es poden fer servir altres estratègies com recollir primer les zones que difracten millor, o bé, en

I422

Estadístiques de la recollida			
Tipus	Natiu		
Línia de difracció	id14EH3		
Data	1999		
Longitud d'ona (Å)	0.93103		
Distància (mm)	600, 399.48, 179.48 (*)		
Temperatura (K)	100		
Detector	MAR CCD		
N. d'imatges recollides	180, 74, 180		
N. graus de rotació en cada imatge	1.0		
Mosaïcitat (°)	0.3		
Grup Espacial	I422		
Cel·la unitat (Å,°)	$a = 155.751$	$b = 155.751$	$c = 160.845$
	$\alpha = 90.0$	$\beta = 90.0$	$\gamma = 90.0$
Resolució (Å)	111.9 - 2.78 (2.93 - 2.78)		
Completeness (%)	99.9 (99.9)		
$\langle I \rangle / \langle \sigma(I) \rangle$	2.9 (5.3)		
Multiplicitat	15.9 (14.6)		
N. de reflexions observades	399674 (52778)		
N. de reflexions úniques	25179 (3607)		
Rmerge (%)	7.4 (27.7)		
B-factor (Wilson plot, Å ²)	52.21		

Taula 3.5: Estadístiques de la recollida i del processament dels cristalls I422. Els valor entre parèntesis es refereixen a la capa de més resolució. (*) tres conjunts de dades

el cas de MAD (secció 3.5.1, plana 42), recollir reflexions oposades (parells de Bijvoet, plana 43) per tal d’aconseguir dades que representin el cristall en un mateix estat.

Això també s’aplica en els canvis de longitud d’ona. Si s’ha recollit una mateixa zona (més l’oposada) en cada longitud d’ona, les dades recollides seran de major precisió. El problema és que amb aquest mètode, en fer moviments més grans del goniòmetre, s’augmenta l’error en l’angle de recollida (deriva) i que el canvi de longitud d’ona és un procés delicat fins i tot en les línies més avançades. Pot escalfar-se el monocromador, o bé desalinear-se algun altre aparell de l’òptica, pel que cal evitar canvis innecessaris en la longitud d’ona. L’estratègia a seguir vindrà donada, en la pràctica, en les

I422Ag

Estadístiques de la recollida	
Tipus	Derivat Ag
Línia de difracció	EH2
Data	1999
Longitud d'ona (Å)	0.932
Distància (mm)	215.2
Temperatura (K)	100
Detector	MAR CCD
N. d'imatges recollides	108
N. graus de rotació en cada imatge	1.0
Mosaïcitat (°)	0.65
Grup Espacial	I4
Cel·la unitat (Å, °)	$a = 155.5$ $b = 155.5$ $c = 160.2$ $\alpha = 90.0$ $\beta = 90.0$ $\gamma = 90.0$
Resolució (Å)	64.7 - 2.75(*) (2.92 - 2.75)
Completeness (%)	94.6 (66.8)
$\langle I \rangle / \langle \sigma(I) \rangle$	5.7 (1.8)
Multiplicitat	4.4 (3.8)
N. de reflexions observades	206461 (20411)
N. de reflexions <i>úniques</i>	46788 (5324)
Rmerge (%)	7.5 (38.4)
B-factor (Wilson plot, Å ²)	72.2

Taula 3.6: Estadístiques de la recollida i del processament dels cristalls I422Ag. Els valor entre parèntesis es refereixen a la capa de més resolució. (*) Es van observar reflexions fins a 2.6 Å

possibilitats que ofereixi la línia de difracció.

Per obtenir un bona estimació de la intensitat d'una reflexió, és necessari recollir-la diverses vegades. Generalment això s'aconsegueix augmentant el rang de rotació de cristall. En la mesura experimental, se sol considerar bona una multiplicitat mitjana de l'ordre de 4. No sempre és possible obtenir-la, sobretot en l'alta resolució, però cal tenir en compte que l'error en la mesura d'una reflexió ($\sigma(I_h)$) influenciarà totes les fases properes. En les reflexions de baixa resolució se solen obtenir valors molt alts d' $\langle I \rangle / \langle \sigma(I) \rangle$, però no és així en les que representen l'alta resolució. Generalment es fa servir aquest criteri per determinar a quina resolució màxima s'ha de recollir, que sol ser d' $\langle I \rangle / \langle \sigma(I) \rangle = 1$ o 2.

I422Au

Estadístiques de la recollida	
Tipus	Derivat Au
Línia de difracció	EH2
Data	1999
Longitud d'ona (Å)	0.932
Distància (mm)	309,5
Temperatura (K)	100
Detector	MAR CCD
N. d'imatges recollides	90
N. graus de rotació en cada imatge	1
Mosaïcitat (°)	1.1
Grup Espacial	I4
Cel·la unitat (Å,°)	$a = 155.6 \quad b = 155.6 \quad c = 159.2$ $\alpha = 90.0 \quad \beta = 90.0 \quad \gamma = 90.0$
Resolució (Å)	79.6-4.0 (4.22-4.0)
Completeness (%)	99.9 (99.7)
$\langle I \rangle / \langle \sigma(I) \rangle$	5.8 (5.1)
Multiplicitat	3.6 (3.6)
N. de reflexions observades	57464 (8402)
N. de reflexions úniques	16052 (2326)
Rmerge (%)	9.0 (13.3)
B-factor (Wilson plot, Å ²)	21.35

Taula 3.7: Estadístiques de la recollida i del processament dels cristalls I422Au. Els valor entre parèntesis es refereixen a la capa de més resolució.

El procés de la indexació ja s’ha fet en recollir les primeres imatges del cristall. Molt breument, consisteix a determinar el grup espacial, l’orientació del cristall i la cel·la a partir de la distribució de reflexions d’un grup petit d’imatges. Donades les característiques físiques de l’experiment (distància del detector, longitud d’ona, polarització del feix, etc) i de les de la imatge (centre, resolució, separació entre reflexions, etc.) es pot predir la mida i orientació de la cel·la unitat:

La projecció de les reflexions en el detector forma unes *llunes* que permeten ordenar les reflexions i calcular la orientació del cristall. A una longitud d’ona fixa, la distancia entre les reflexions és inversament proporcional a la mida de la cel·la unitat i proporcional a la distància entre el detector i el cristall. Un cop predita una cel·la i una orientació, cal comprovar si és correcta,

I422Pt

Estadístiques de la recollida	
Tipus	Derivat Pt
Línia de difracció	EH2
Data	1999
Longitud d'ona (Å)	0.932
Distància (mm)	309.7
Temperatura (K)	100
Detector	MAR CCD
N. d'imatges recollides	100
N. graus de rotació en cada imatge	1
Mosaïcitat (°)	1.1
Grup Espacial	I4
Cel·la unitat (Å, °)	$a = 155.4$ $b = 155.4$ $c = 158.7$ $\alpha = 90.0$ $\beta = 90.0$ $\gamma = 90.0$
Resolució (Å)	79.4-4.0 (4.22-4.0)
Completeness (%)	96.7 (98.2)
$\langle I \rangle / \langle \sigma(I) \rangle$	3.6 (4.4)
Multiplicitat	3.0 (3.0)
N. de reflexions observades	46075 (6729)
N. de reflexions <i>úniques</i>	15436 (2279)
Rmerge (%)	8.3 (13.4)
B-factor (Wilson plot, Å ²)	73.6

Taula 3.8: Estadístiques de la recollida i del processament dels cristalls I422Pt. Els valor entre parèntesis es refereixen a la capa de més resolució.

tractant de predir el patró en altres imatges no fetes servir en la indexació. Moltes mesures de la fiabilitat dels resultats, tant en la indexació com en càlculs posteriors es basen en la predicció de reflexions no fetes servir en els càlculs.

Un cop acabada la recollida de totes les dades, a partir de la indexació, es podrà predir la situació de cada reflexió en cada una de les imatges recollides. La integració consisteix a escanejar la quantitat d'energia dipositada en cadascuna de les reflexions. Per això es limita cada reflexió dins d'una capsa i s'ajusta el perfil de la reflexió a una gaussiana. Aquesta capsa està formada per dues zones, la que pertany pròpiament a la reflexió, i la que forma el soroll de fons de cada reflexió, que es farà servir per modular l'error de la mesura. No es pot fer servir un soroll de fons general per a tota la

I422TaBr

Estadístiques de la recollida			
Tipus	MAD Ta_6Br_{12}		
Línia de difracció	bm30		
Data	1999		
Longitud d'ona (Å)	1.2651	1.2565	1.2559 1.2435
Distància (mm)	350.0		
Temperatura (K)	100		
Detector	MAR 345		
N. d'imatges recollides	30+60(*)		
N. graus de rotació en cada imatge	1.0		
Mosaïcitat (°)	0.4		
Grup Espacial	I422		
Cel·la unitat (Å,°)	$a = 156.085$ $\alpha = 90.0$	$b = 156.085$ $\beta = 90.0$	$c = 161.080$ $\gamma = 90.0$
Resolució (Å)	55.9 - 3.0		(3.16 - 3.0)
Completeness (%)	99.2		(99.8)
$\langle I \rangle / \langle \sigma(I) \rangle$	4.9		(2.8)
Multiplicitat	5.9		(5.9)
N. de reflexions observades	117274		(16973)
N. de reflexions úniques	19990		(2890)
Rmerge (%)	9.2		(26.4)
B-factor (Wilson plot, Å ²)	56.25		

Taula 3.9: Estadístiques de la recollida i del processament dels cristalls I422TaBr. Els valor entre parèntesis es refereixen a la capa de més resolució. (*) Es va recollir en dues parts a causa de que estava maclat

imatge per que aquest és molt variable, principalment amb la resolució. El soroll de fons en la zona de baixa resolució sol ser més alt que la intensitat de les reflexions d'alta resolució. Per això cal que el detector tingui un rang dinàmic molt alt. La presència de solvent desordenat en els cristalls provoca anells difusos que augmenten el soroll de fons.

El resultat de la integració és un arxiu per a cada imatge recollida on s'especifiquen les coordenades de la reflexió, la intensitat integrada, el seu σ , i altres característiques que seran necessàries per escalar-les posteriorment: angle de rotació, quina proporció de la reflexió està tota continguda en la imatge, les coordenades en el difractòmetre, etc.

L'escalat és la etapa del processament de les dades en la qual es tracta

d'ajuntar totes les reflexions repetides que s'han recollit i de reconstruir les que estan disperses en varies imatges, a causa de la mosaïcitat. Es fan servir les reflexions que estan senceres en una imatge i les equivalents en altres imatges per poder calcular un factor d'escala entre elles. Si el cristall té molta mosaïcitat, es pot intentar escalar també a partir de les reflexions parcials. Petites diferències en el temps de recollida, en la intensitat del feix, o per degradació de la mostra fan que calgui aplicar una escala entre elles, i obtenir un perfil de l'evolució de la difracció. Es per tant important definir un model d'escalat que tingui en compte si s'ha de fer fent servir el soroll de fons per determinar com evoluciona la difracció, o bé si cal un escalat independent de cada reflexió i de les seves simètriques. També cal fer separar els parells de Bijvoet si la difracció és anòmala i es volen fer servir les diferències entre aquests per resoldre l'estructura.

La qualitat de l'escalat es jutja a partir d'uns residuals que avaluen la desviació de les reflexions, ordenades per capes de resolució, i també per intensitats. Valors baixos indicaran que les reflexions relacionades per simetria, o per què s'han recollit diverses vegades, són similars i per tant es tracta de reflexions ben processades o recollides.

La similitud entre reflexions recollides diverses vegades es mesura amb un residual anomenat *Rmerge*:

$$R_{merge} = \frac{\sum_{i,h} I_{i,h} - \langle I_h \rangle}{\sum_h \langle I_h \rangle}$$

on $I_{i,h}$ és la mesura i de la reflexió h , i $\langle I_h \rangle = \frac{\sum_i I_{i,h}}{\sum_i 1}$.

Anàlogament, *Rsym* és un indicador de la similitud entre les reflexions que estan relacionades pel grup espacial. L'expressió és idèntica, només que l'índex i representa les reflexions simètriques.

Ja s'ha mencionat que una altra mesura de la qualitat és $\langle I \rangle / \langle \sigma(I) \rangle$, també tabulada en capes i per intensitats, que ajuda a discernir fins a quina resolució la difracció és prou intensa, o si hi ha àrees cremades o no aprofitables per culpa de la presència d'un anell de solvent. La última quantitat important és la completa, distribuïda per capes de resolució, que mostra quin tant per cent del total de reflexions s'han recollit en cadascuna d'elles (taula 3.4).

Una reflexió vindrà caracteritzada per 3 coordenades ($\mathbf{h} = (h, k, l)$), una intensitat $I_{\mathbf{h}}$ i una mesura de la seva qualitat $\sigma(I_{\mathbf{h}})$. En el cas d'un experiment de MAD, també s'hauran d'incloure per separat els valors dels parells de Bijvoet i les seves *sigmes*.

Generalment a l'hora d'afinar una estructura es treballa amb les amplituds, que són proporcionals a l'arrel quadrada de les intensitats. En un últim

pas del processament de les reflexions cal aplicar aquesta proporció i alhora s'aprofita per calcular el *Wilson Plot*, que és una estadística de la intensitat mitjana de les reflexions en cada capa de resolució. El pendent mitjà a la zona central indicarà el valor mitjà de la temperatura (*B-factor*) o desordre del cristall. Aquest valor mig s'ha de conservar en l'estructura afinada final. El *Wilson Plot* permet també aplicar una escala absoluta a les reflexions.

Si els àtoms estan distribuïts aleatòriament en la unitat asimètrica, llavors

$$\langle F^2 \rangle = k \langle Fo^2 \rangle e^{-2B \frac{\sin^2(\theta)}{\lambda^2}}$$

on θ és l'angle de difracció segons la llei de Bragg

$$2d \sin(\theta) = n\lambda$$

amb d la resolució i 2θ l'angle entre el raig incident i el difractat (fig. 3.5) [1].

Els àtoms no estan distribuïts aleatòriament en estructures reals. Aquest efecte s'aprecia principalment amb les reflexions de baixa resolució. Normalment es fan servir només les reflexions d'alta resolució per a calcular el *Wilson Plot* (fig. 3.5).

De cara a treballar amb mètodes de màxima versemblança, és important tenir un conjunt de dades complet, encara que no s'hagi mesurat tot. Les reflexions que falten s'afegeixen, junt amb un senyal que indiqui que no s'han observat (*missing flag*) i se n'assigna a tota la part observada un altre que indicarà si la reflexió s'haurà de separar o considerar a l'hora de afinar, per poder calcular l'*Rfree* ("reflexions *Rfree*", plana 58).

3.5 Solució de l'estructura tridimensional

3.5.1 MIR i MAD

Quan no hi ha cap model de partida (una estructura similar), el problema de les fases (plana 30) obliga a reduir l'estructura a un nombre petit d'àtoms i tractar d'afrontar el problema amb mètodes directes o de Patterson. Els àtoms pesants tenen una contribució més gran en la difracció, i presenten anisotropicitat quan la longitud d'ona dels raigs X correspon a un dels seus *edges* d'absorció.

La tècnica del MIR (*Multiple Isomorphous Replacement*) és de les més antigues que es coneix, i consisteix en recollir per separat dades de difracció de diferents cristalls isomorfs (mateixa cel·la i mateix empaquetament) que contenen àtoms pesants afegits. En principi calen un mínim de tres

cristalls (hi ha variants, plana 46). Un pot ser el natiu i els altres dos han de ser derivats diferents. Si els cristalls són prou isomorfs i es poden posar les intensitats difractades de cadascun en una mateixa escala, una síntesi de fourier de les diferències entre les intensitats del natiu i cadascun dels derivats donarà lloc a un mapa (*Patterson* plana 47) amb densitats altes en els punts corresponents a les distàncies entre els àtoms pesants. Com que és una diferència entre dos cristalls que només difereixen pels àtoms pesants, les distàncies entre els àtoms de la proteïna no apareixeran. Per això és important l'isomorfisme. Per consideracions amb el grup espacial, les diferències apareixeran formant uns motius (plans i línies de Harker) que permetran el càlcul de la posició dels àtoms pesants. La relació amb les fases de la proteïna és directa, però cal un segon cristall derivat perquè hi ha una incertesa en la fase, ja que d'aquesta relació només pot trobar el cosinus de la fase:

$$\phi_P = \phi_{H1} + \cos^{-1} \left[\frac{F_{PH1}^2 - F_P^2 - F_{H1}^2}{2F_P F_{H1}} \right]$$

$$\phi_P = \phi_{H2} + \cos^{-1} \left[\frac{F_{PH2}^2 - F_P^2 - F_{H2}^2}{2F_P F_{H2}} \right]$$

on F són els factors d'estructura i ϕ són les fases. El subíndex P representa el cristall de proteïna sol, $PH1$ i $PH2$, els cristalls de proteïna i l'ió pesant, i $H1$ i $H2$ les amplituds i fases calculades pels àtoms pesants, a partir de la seva posició en el cristall.

El principal problema del MIR és la necessitat d'obtenir dos derivats d'una proteïna nativa, i que a més hi siguin isomorfs. Si es pot aconseguir una radiació a una longitud d'ona corresponent a l'*edge* d'absorció d'un ió pesant, llavors el problema es redueix a obtenir un sol cristall derivat amb l'àtom pesant, i optativament, un natiu si cal verificar que no hi hagi hagut modificacions importants en l'estructura. Això podria passar, per exemple, si l'àtom pesant interacciona amb la proteïna en aminoàcids actius. A aquestes longituds d'ona els electrons absorbeixen l'energia dels fotons per canviar a una òrbita superior i, per després emetre fluorescència en tornar al nivell energètic original. El resultat és que la difracció es torna anisòtropa (anòmala), i les reflexions oposades (parells de Friedel) ja no són d'igual intensitat. En aquest cas els parells de reflexions, i els seus simètrics, s'anomenen parells de Bijvoet.

Un experiment de MAD (*Multiwavelength Anomalous Dispersion*) consisteix a recollir dades de difracció del cristall més d'un cop (típicament tres), cada vegada a una longitud d'ona diferent. Les longituds d'ona es trien de manera que una d'elles estigui en el màxim d'absorció, l'altra en el punt

d'inflexió i l'altra després d'aquest màxim (figs. 3.6, plana 44 i 3.7, plana 45).

Per comprovar la presència de l'àtom pesant dins del cristall, i calcular quina és l'energia d'absorció, se sotmet el cristall a un rang d'energies variable i es mesura la fluorescència produïda (XAFS, X ray Absorption Fine Structure). Aquesta fluorescència és màxima en l'*edge* d'absorció. La fluorescència és una radiació isotròpica que no depèn de l'estat cristal·lí, per la qual cosa es pot produir tant per la presència d'àtoms pesants en la gota que protegeix el cristall com per atòms pesants desordenats dins del cristall (fig. 3.6).

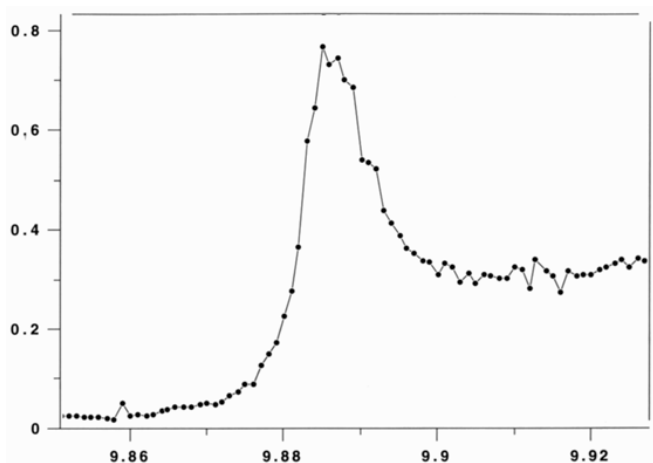


Figura 3.6: Fluorescència TaBr. Màxim d'absorció per un cristall del connector derivat amb TaBr ($1422 Ta_6Br_{12}$). A l'eix horitzontal es representa el rang d'energies en KeV i al vertical la fluorescència en unitats relatives.

A partir de les dades recollides, es poden tractar els resultats com en un experiment de MIR, substituint les diferències entre cristalls, per diferències entre longituds d'ona i entre parells de Bijvoet. Els factors de difusió atòmics s'han d'ajustar a les propietats de l'àtom pesant en cada longitud d'ona.

Un factor d'estructura en una dispersió anòmala es pot separar en dues parts,

$$F(\mathbf{h}) = F'(\mathbf{h}) + iF''(\mathbf{h})$$

(fig. 3.7) on F' representa la contribució no anòmala (proteïna i àtoms pesants) i F'' l'anòmala, a causa només dels àtoms pesants (o dispersor anòmal). A causa de l'anisotropicitat hi haurà un factor d'estructura diferent per a cada parell de reflexions oposades. Aquesta contribució es pot trobar en taules, o millor per fluorescència, ja que l'entorn molecular la modifica una mica. Si F^+ i F^- són les amplituds mesurades per a cada parell de Bijvoet, es pot calcular el sinus de la fase i trencar l'ambigüitat aprofitant una de les

expressions d'abans pel MIR:

$$\phi_{PH} = \phi_H \pm \sin^{-1} \left[\frac{F_{PH}^+ - F_{PH}^-}{2F_H''} \right]$$

$$\phi_P = \phi_H + \cos^{-1} \left[\frac{F_{PH}^2 - F_P^2 - F_H^2}{2F_P F_H} \right]$$

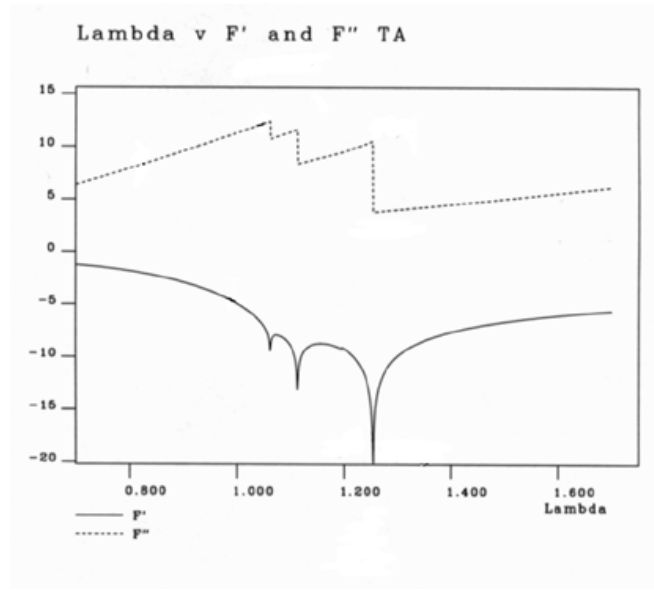


Figura 3.7: Factors de difusió atòmics. A partir de la corba 3.6 es poden calcular els factors de difusió atòmics dels àtoms pesants. Un factor de difusió al voltant d'un *edge* d'absorció té la forma $f = f_o + \Delta f' + i f''$ on f_o és la contribució *normal* i $\Delta f'$ i f'' són, respectivament les parts real i imaginària degudes al moment dipolar dels electrons que oscil·len entre dos nivells energètics. Els factors d'estructura F' i F'' s'obtenen agrupant les parts real i imaginària en fer la suma sobre tots els àtoms. A la figura F' i F'' representen $\Delta f'$ i f'' . Els coeficients d'absorció $\mu(E)$ de la corba 3.6, estan relacionats amb $f''(E)$ segons $\mu(E) = (4\pi N \hbar e^2 / m E) f''(E)$ (N és el nombre atòmic, m i e són la massa i la càrrega de l'electró). A partir d'una de les relacions de Kramers-Krönig s'obté l'altre factor de difusió: $f'(E) = \frac{2}{\pi} P \int_0^\infty \frac{E' f''(E')}{E^2 - E'^2} dE'$ amb P , el valor principal de la integral a la singularitat, i es suposa que no hi ha dependència entre f' i f'' respecte la direcció en l'espai recíproc [40].

Com que es tenen diferències entre longituds d'ona i entre parells de reflexions, per calcular el Patterson es poden fer més combinacions que en

el cas del MIR. Per això, de vegades un experiment de MAD es pot resoldre amb menys longituds d'ona, si les dades són prou bones. Un requisit per fer un experiment de MAD és disposar d'una línia de sincrotró que permeti ajustar la longitud d'ona. En un aparell d'ànode rotatori no es poden fer aquest tipus d'experiments ja que la longitud d'ona és fixa.

Per la similitud entre els dos mètodes, existeixen altres combinacions com el SIRAS (*Single Isomorphous Replacement and Anomalous Scattering*), on es fa servir un natiu i un derivat que tingui un àtom pesant amb contribució anòmala significativa. A la pràctica, però, el que s'intenta és preparar alhora cristalls natis, derivats normals i derivats per MAD i combinar els que s'hagin pogut obtenir i fer un experiment de MIRAS (*Multiple Isomorphous Replacement and Anomalous Scattering*). La qualitat (o l'existència) de la difracció determinarà *a posteriori* el mètode a utilitzar per resoldre l'estructura.

3.5.2 Reemplaçament Molecular

El reemplaçament molecular és la tècnica que més es fa servir actualment per a resoldre el problema de les fases. A mesura que el nombre d'estructures conegudes augmenta, el nombre de plegaments o estructures terciàries noves augmenta a un ritme cada cop més lent. A l'any 1999 només un 9% de les molècules resoltes contienien un *folding* nou [34]. El reemplaçament molecular és capaç de determinar l'orientació i posició d'una proteïna en el cristall a partir de l'estructura tridimensional d'una altra proteïna o fragment similar, una solució de RMN (ressonància magnètica nuclear), o d'un model de baixa resolució provinent de microscòpia electrònica, que és el que s'ha intentat en aquest treball.

Consisteix en la cerca sistemàtica per correlació entre les intensitats difractades i les que produiria el model, tractat com un cos rígid, en successius empaquetaments en la cel·la cristal·logràfica. Es tracta d'un problema de 6 dimensions per molècula en la unitat asimètrica (tres rotacions i tres translacions), per la qual cosa cal dividir-lo en cerques més senzilles, o seria impracticable el temps de CPU necessari, ja que per a cada empaquetament cal calcular els factors d'estructura, per FFT, i després calcular-ne la correlació amb les intensitats observades.

El problema es divideix en dos. La primera part consisteix en una cerca de l'orientació de cada molècula model posada a l'origen de la cel·la. Molt bàsicament, es calculen les intensitats per cadascuna de les rotacions i es correlacionen amb les observades. Els valors de millor correlació correspondran a l'orientació de les molècules en el cristall. Es treballa amb intensitats per que la transformada de Fourier d'aquestes dona lloc un mapa que és l'auto-

convolució de la cel·la unitat. En aquests mapes es mantenen les orientacions de la molècula, però totes les translacions es desplacen a l'origen (mapes de Patterson):

$$P(\mathbf{u}) = \int_V \rho(\mathbf{r})\rho(\mathbf{r} + \mathbf{u})d\mathbf{r} = \frac{1}{V} \sum_{\mathbf{h}} |F_{\mathbf{h}}|^2 \cos(2\pi\mathbf{h} \cdot \mathbf{u})$$

$\mathbf{u}$ (vector de Patterson) representa totes les distàncies interatòmiques, traslladades a l'origen. Es calcula una funció de rotació

$$\mathcal{R}(R) = \frac{1}{V} \int_V P(\mathbf{u})P_m(\mathbf{R}^{-1}\mathbf{u})d^3\mathbf{u}$$

entre el Patterson natiu de les reflexions (P) i el del model situat a l'origen de coordenades i en diverses orientacions (P_m).

De tots els valors de correlació, corresponents a orientacions de la molècula s'extreuen tantes orientacions com molècules en la unitat asimètrica. Generalment s'observen grups de solucions similars que corresponen a petites diferències en l'orientació, a causa del soroll. El nombre de molècules per unitat asimètrica es pot estimar senzillament a partir del volum de Matthews (V_m):

$$1.66 \leq V_m = \frac{V(\text{\AA}^3)}{M(Da)N_{asu}N_{mol}} \leq 4$$

on V és el volum de la cel·la en $\text{\AA}^3$, M és la massa de la proteïna en Dalton, N_{asu} és el nombre d'unitats asimètriques i N_{mol} és el nombre de molècules per unitat asimètrica. Es tracta d'una llei empírica corresponent a cristalls amb un rang del 30% al 75% de solvent, pel que l'última paraula la té l'empaquetament trobat.

Un cop resolt el problema de les orientacions, s'aplica un procediment similar per les translacions. Aquest és més ràpid, per que una translació només implica un canvi en la fase de les amplituds calculades a partir de les rotacions. Es trasllada, tridimensionalment, primer una de les molècules fins arribar a la millor correlació, i després es fixa en aquesta posició per tenir-la en compte en el càlcul de les amplituds de les següents molècules. Els factors d'estructura de cada molècula i orientació s'han calculat prèviament en la rotació. Si apliquem una translació T a una molècula orientada segons R , els nous factors d'estructura seran:

$$F_{\mathbf{h}}^c(\mathbf{T}) = \sum_s^{N_{asu}} g_s(\mathbf{h})e^{2\pi i\mathbf{h}R\mathbf{T}}$$

on $g_s(\mathbf{h})$ és un factor d'estructura global calculat prèviament per a cadascuna de les N_{asu} molècules de la unitat asimètrica.