

Part I: la proteïna p4

4.1. Clonació, expressió i purificació de la proteïna p4

Aquesta part del treball es va dur a terme al laboratori de la Prof. Margarita Salas, i la proteïna ens era enviada precipitada en sulfat amònic.

4.1.1. Solubilització de la proteïna

La solubilització es va realitzar mitjançant una resuspensió del precipitat de sulfat amònic de la proteïna, amb el tampó (NaCl 200 mM, Tris 20 mM pH=7,5, EDTA 1 mM i DTT 1 mM) que vam utilitzar al llarg de tots els experiments de formació de complexos i les purificacions cromatogràfiques dels complexos.

El següent pas va ser eliminar el sulfat amònic de la mostra, ja que podria interferir en els experiments de cristal·lització, degut a la naturalesa caotrópica d'aquest agent.

Es va optar per fer un canvi de tampó mitjançant centrifugació en tub centricon. Aquest mètode és eficient ja que en el mateix procés es fa un canvi de tampó i es concentra la mostra a la concentració apropiada per fer els experiments de cristal·lització, que en el nostre cas va ser de 13 mg/ml.

4.1.2. Cristal·lització

Les primeres proves de cristal·lització de la proteïna p4 es van dur a terme a una concentració de 13 mg/ml. S'obtingueren els primers cristalls a partir de la condició número 1 del CRYSTAL SCREEN II (Hampton) a 4°C. Aquesta condició conté NaCl 2M i PEG 6K al 10%(v/v) i els cristalls apareixien al cap de dues setmanes.

Aquests cristalls, de dimensions aproximades 0,4 x 0,2 x 0,1 mm, ja no van ser millorats i són els cristalls que es van utilitzar per recollir dades de difracció (Figura 4.1).

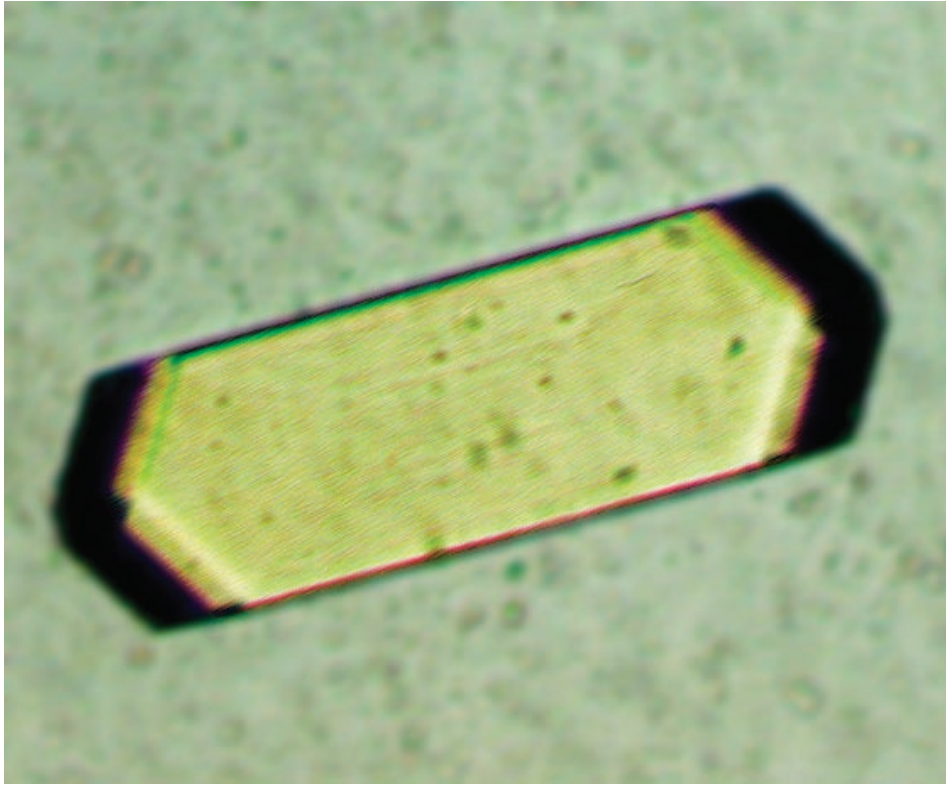


Figura 4.1: Cristall de proteïna p4

4.1.3. Crioprotecció i congelació dels cristalls

La congelació dels cristalls de proteïna té com a objectiu reduir al màxim tant el dany per radiació com la dessecació del cristall durant la recollida de les dades de difracció. El dany per radiació dels biocristalls sembla estar relacionat amb la formació de radicals lliures. Encara que els processos fotoquímics que produeixen aquests radicals lliures (dany per radiació primari) són fenòmens localitzats, es poden induir subseqüents reaccions químiques a llocs relativament remots degut a la propagació via difusió dels radicals lliures en les regions de solvent del cristall de proteïna (dany per radiació secundari). D'aquesta manera, el dany per radiació s'estén i provoca que la intensitat de la difracció decaigui, procés que sol anar

acompanyat per canvis en els patrons de difracció i en les dimensions de la cel·la unitat. La reducció del dany per radiació dels cristalls que es recullen a temperatures criogèniques és deguda a que, en disminuir la temperatura, el processos de difusió i, en conseqüència, la propagació dels radicals lliures altament reactius dins del cristall, s'alenteix.

Quan un cristall d'una macromolècula biològica es refreda fins a temperatures criogèniques, el problema principal és evitar la cristallització de qualsevol molècula d'aigua present en el sistema, tant sigui externa com interna al cristall. Per això s'ha d'escollir un procés de refredament que porti a la vitrificació del solvent, és a dir, a que aquest assoleixi una fase amorfa. Per aconseguir-ho es modifiquen les propietats fisicoquímiques de la solució en que ha crescut el cristall afegint-hi crioprotectors, de manera que la vitrificació es pot assolir a taxes de refredament moderades. S'han de superar dos problemes: el crioprotector s'ha d'introduir sense que això produeixi un dany significatiu al cristall, i el dany durant la congelació s'ha de minimitzar.

Tal com s'ha explicat a material i mètodes, es va provar un sol tipus de crioprotector, el glicerol, un dels més utilitzats en criocristal·lografia.

4.1.3.1. Soakings

En aquest procediment forcem que un compost que conté l'àtom pesant que ens interessa, reaccioni amb la proteïna, incorporant l'àtom pesant de manera permanent i en una regió concreta de la proteïna. S'ha d'arribar a un compromís a l'hora de triar el compost entre la capacitat d'integració estable que mostra i l'agressivitat amb que ho fa, ja que ens interessa integrar un àtom pesant sense que el cristall es malmeti ni canviï en excés els seus paràmetres de cel·la ni de simetria, és a dir que el cristall amb la proteïna derivatitzada sigui isomorf amb el cristall natiu.

La intenció de derivatitzar la proteïna era poder realitzar un experiment de tipus MAD, pel que s'havia d'utilitzar un àtom pesant que, a més, fos un bon dispersor

anòmal. Així, vam triar 3 elements : platí, mercuri i or. De cadascun d'ells en vam utilitzar 3 compostos:

Pel mercuri: clorur de mercuri, clorur d'etilmercuri i thimerosal.

Pel platí: tetraclorplatínat potàssic, tetracianoplatínat potàssic i tetranitroplatínat potàssic.

Per l'or: dicianoaurat potàssic, clorur d'or i tetracloraurat sòdic.

Es van fer proves de tots tres a 10mM de concentració i a tres temps diferents: 6h, 24h i 48h, i es va procedir al congelat dels cristalls. Alguns dels compostos, sobretot els de mercuri, van mostrar la seva capacitat de reaccionar amb la proteïna, desintegrant alguns dels cristalls.

4.2. Recollida de les dades de difracció del cristall derivatitzat

De tots els cristalls provats, l'únic cristall amb que es va detectar el cantó d'absorció corresponent al dispersor anòmal introduït, mesurant la fluorescència i un bon patró de difracció va ser el cristall que estava derivatitzat amb tetraclorplatínat potàssic, a una concentració final de 10mM durant 24h.

Amb aquest cristall va realitzar un experiment de tipus MAD, recollint-se tres jocs complets de dades, amb una resolució màxima de 2.9 Å. Els factors de dispersió anòmala f' i f'' derivats de l'espectre de fluorescència es mostren en la Figura 4.2. La primera longitud d'ona es va recollir on f'' era màxima (pic), la segona on f' era mínima (punt d'inflexió) i la tercera a una longitud d'ona allunyada de les dues (remota). Aquestes longituds d'ona tenien un valor de 1,0717 Å pel pic; 1,0721 Å per la inflexió i 0,9762 Å per la remota.

4.2.1. Processament de les dades de difracció C222₁

En realitzar-se l'indexat de les dades de difracció de raigs X del cristall de proteïna p4 derivatitzat amb hexaclorplatínat, es va veure que pertanyia a un

grup espacial ròmbic centrat. Un cop examinades les extincions sistemàtiques per determinar amb més exactitud el grup espacial, vam determinar que es tractava del grup espacial $C222_1$.

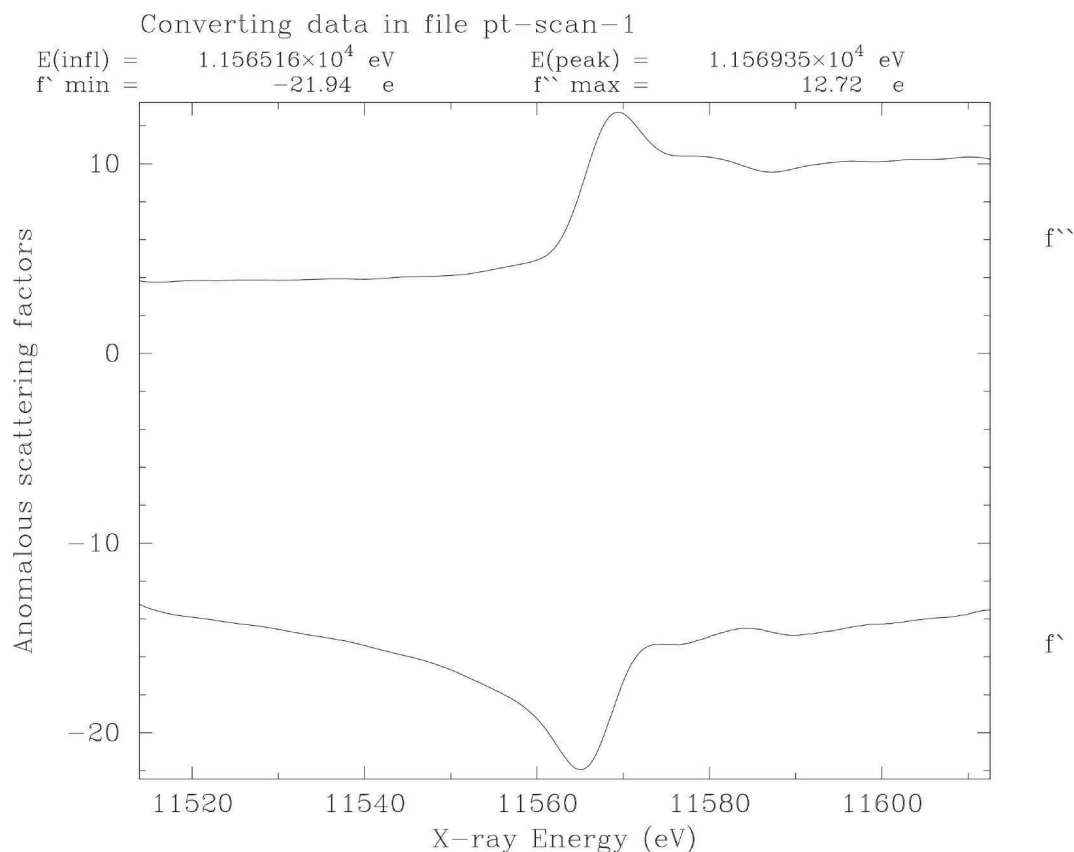


Figura 4.2: Factors de dispersió f' i f'' determinats a partir d'un espectre de fluorescència d'un cristall derivatitzat amb platí

Amb el processament de les dades fins a 2.9 \AA i el seu posterior escalat es va obtenir una cel·la unitat afinada, els paràmetres de la qual són: $a=135,57$; $b=140,51$; $c=66,92 \text{ \AA}$ i $\alpha=\beta=\gamma=90^\circ$.

El càlcul del paràmetre de Matthews (Matthews, 1968) dóna uns valors de $V_m=2.6 \text{ \AA}^3 \text{ Da}^{-1}$ si considerem 4 molècules per unitat asimètrica i $3.5 \text{ \AA}^3 \text{ Da}^{-1}$, si considerem 3 molècules per unitat asimètrica. Això suggereix la presència de 4 ó 3 molècules a la unitat asimètrica, amb un percentatge de contingut de solvent de

52% i 64,5%, respectivament.

Els conjunts de dades corresponents a les 3 longituds d'ona es van processar per separat, obtenint-se unes estadístiques detallades en la Taula 4.1.

Conjunt de dades	Derivat de platí		
	Pic	Inflexió	Remota
Grup espacial	C222 ₁	C222 ₁	C222 ₁
Cel·la unitat	a=135.57, b=140.51, c=66.92		
Longitud d'ona (Å)	1.07170	1.07210	0.97620
Reflexions mesurades	117195	87546	100928
Reflexions úniques	14526	14411	14928
Rang de resolució (Å)	55.2 - 2.9	55.2 - 2.9	55.2 - 2.9
Completeness(%)	97.4 (97.4)	98.9 (98.9)	100 (100)
Multiplicitat	8.0 (5.7)	5.9 (4.0)	6.7 (6.3)
Rmerge (%)	9.1 (44.3)	9.2 (47.3)	10.4 (48.6)
$\langle I \rangle / \langle \sigma(I) \rangle$	7.9 (1.7)	7.7 (1.6)	6.9 (1.5)
Ranomal (%)	5.0 (22.3)	4.9 (27.2)	5.3 (22.9)

Taula 4.1: Processat de les dades del cristall C222₁.

$R_{merge} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_{ii} |I_i(hkl)|$, per a tot el rang. On:

hkl : índexs de les reflexions úniques

$I_i(hkl)$: és la i mesura d'intensitat de la reflexió hkl

$\langle I(hkl) \rangle$: és la mitja de les mesures de la reflexió hkl

Els valors entre parèntesi són els de la última capa

El procés d'escalat de les dades es va fer conjuntament pels tres conjunts de dades. Les estadístiques obtingudes es detallen en la Taula 4.2 per les diferents capes de resolució.

Resolució	Reflexions Mesurades	Reflexions Úniques	Completeness %	Multiplicitat	R_{merge}	R_{sym}
9.17	9340	524	99.7	17.8	6.1	5.9
6.48	18142	876	99.9	20.7	7.2	7.0
5.29	23958	1125	99.9	21.3	9.3	9.1
4.59	27797	1293	100.0	21.5	8.1	7.9
4.10	31758	1467	100.0	21.6	8.6	8.4
3.74	34757	1609	100.0	21.6	11.9	11.6
3.47	37672	1737	100.0	21.7	14.8	14.4
3.24	41257	1883	100.0	21.9	21.4	20.9
3.06	43114	1963	100.0	22.0	37.4	36.5
2.90	36671	2094	100.0	17.5	53.8	52.3
Totals	304466	14571	100.0	20.9	12.6	12.3

Taula 4.2: Escalat de les dades del cristall C222₁.

4.2.2. Resolució del problema de la fase pel mètode MAD

El mètode utilitzat per la resolució del problema de la fase va ser el de la dispersió anòmala a múltiples longituds d'ona o MAD.

Per fer la cerca de les posicions dels àtoms pesants vam utilitzar el programa SOLVE (Terwilliger i Berendzen, 1999). Hi ha un paràmetre que s'ha de tenir en compte a l'hora de fer la cerca, i és la resolució. Normalment s'utilitzarà la resolució màxima, però cal mirar la figura de mèrit corresponent a cada resolució, per saber quines regions de resolució són significatives. Per tal d'obtenir un càlcul de la figura de mèrit en funció de la resolució, vam fer una primera cerca utilitzant el màxim de resolució (Taula 4.3).

FIGURE OF MERIT WITH RESOLUTION									
DMIN:	TOTAL	10.47	6.60	5.16	4.37	3.86	3.49	3.21	2.99
N:	14560	735	1211	1539	1811	2033	2227	2424	2580
MEAN FIG MERIT:	0.33	0.50	0.58	0.57	0.49	0.35	0.26	0.17	0.10

Taula 4.3: Figura de mèrit en funció de la resolució

La figura de mèrit és un valor indicatiu de la fiabilitat de la fase, on la diferència fins a 1 ens mostra la component aleatòria. Així ens interessarà utilitzar una figura de mèrit com més gran millor, però s'ha d'arribar a un compromís amb la resolució, ja que a baixa resolució es poc probable trobar posicions que ens permetin fer una bona extensió de fases. A la resolució màxima la figura de mèrit era de 0.1, així que vam realitzar la cerca a una resolució de 3.5 Å, amb una figura de mèrit de 0.26, obtenint 3 posicions d'àtoms pesants (Taula 4.4):

SITE	ATOM	OCCUP	X	Y	Z	B
1	Pt	0.1963	0.6407	0.3144	0.2269	60.0000
2	Pt	0.2241	0.2671	0.4964	0.0408	60.0000
3	Pt	0.1318	0.6328	0.3215	0.1459	50.3074

Taula 4.4: Posicions trobades pels àtoms de platí

Aquestes posicions van ser utilitzades per calcular les fases de cada reflexió fins a 3.5 Å amb el programa SOLVE, i calculant-se a partir d'elles un mapa de densitat electrònica.

4.2.3. Inspecció i validació de les posicions

El procediment d'inspecció de les posicions d'àtoms pesants es basa en l'observació del mapa de densitat electrònica obtingut a partir del càlcul de les fases. L'observació d'un mapa de densitat electrònica amb fases primerenques és sovint una tasca complicada, ja que normalment aquestes fases no són molt precises i el mapa que s'obté es força confós.

De totes maneres, a l'hora d'inspeccionar el mapa de densitat electrònica observarem primer les posicions trobades dels àtoms pesants. Aquestes posicions han d'estar fortament envoltades de densitat electrònica (ρ) elevada (per exemple, en el nostre cas, amb $\rho > 5\sigma(\rho)$).

Si les posicions estan ben trobades llavors inspeccionem la resta del mapa, mirant que la densitat electrònica no estigui distribuïda de manera atzarosa, el que ens indicaria que les fases tenen una gran component aleatòria, buscant les zones de proteïna i les zones de solvent (Figura 4.4), i si el mapa és molt bo, zones d'estructura secundària.

En el nostre cas, les posicions van ser ben trobades, tal i com es mostra en la Figura 4.3, podent-se intuir on es troba la cadena principal d'aminoàcids.

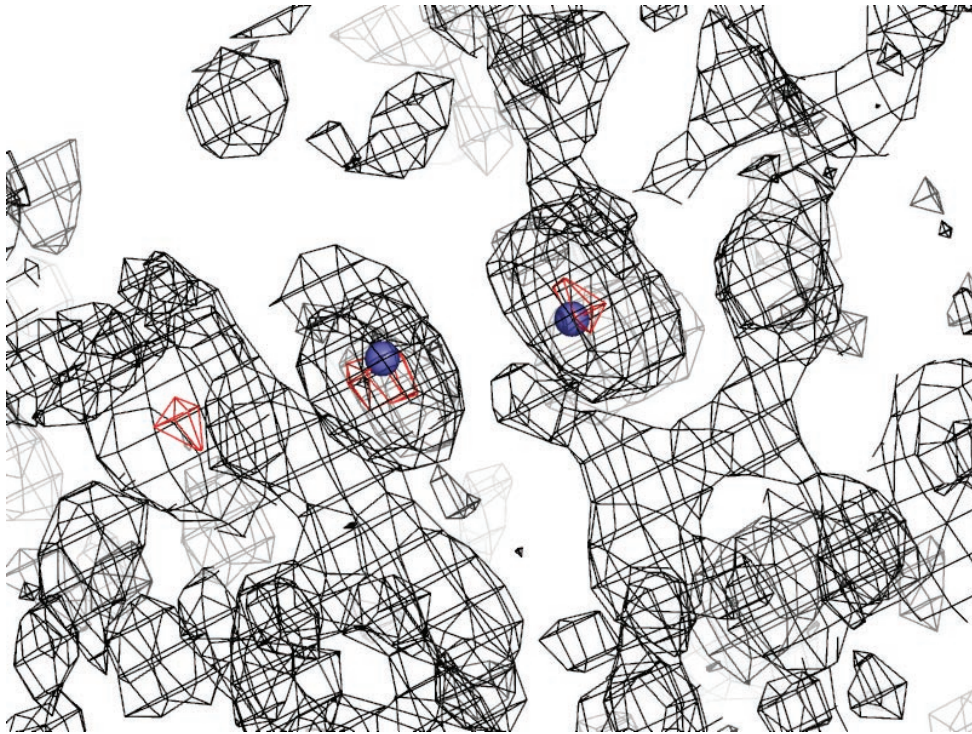


Figura 4.3: Posicions dels àtoms de platí obtingudes amb el programa SOLVE. En vermell es mostra la densitat electrònica contornejada a un valor de 5 sigma.

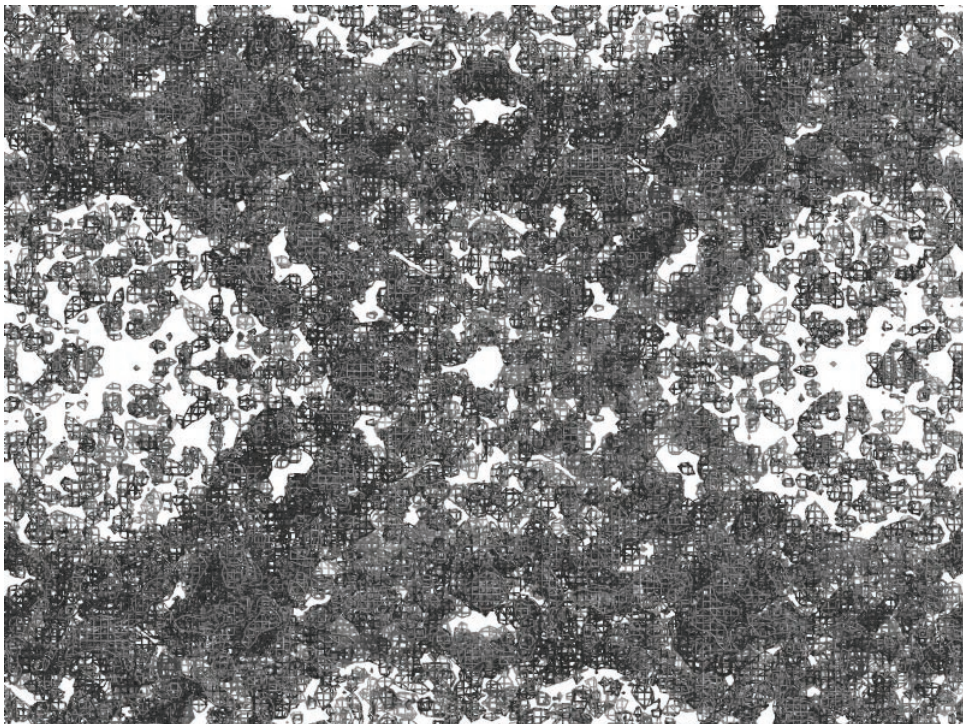


Figura 4.4: Mapa experimental MAD a 3.5 Å de resolució. En una visió global podem veure on es troben les zones de solvent i les de proteïna.

4.2.4. Extensió de les fases

Un cop validades les posicions dels àtoms pesants, i així considerant les fases obtingudes com a correctes, el següent pas va consistir en estendre les fases fins al màxim de resolució del que disposàvem.

El fet d'estendre fases ens era necessari, ja que la resolució a la qual treballàvem (3.5 Å) era molt baixa i no ens permetia utilitzar cap programa de traçat automàtic, ni tampoc poder traçar manualment el model de manera més o menys àgil.

L'extensió de fases, tal i com s'explica en l'apartat de materials i mètodes, és un procés iteratiu en el qual es defineix una màscara de la densitat corresponent a la proteïna i aplicant mètodes de modificació de densitat es calculen les fases de les reflexions de major resolució que les determinades inicialment.

El procés d'extensió de les fases es va realitzar amb el programa RESOLVE (Terwilliger, 2000). Per fer l'extensió de fases es van fer servir les tècniques de modificació de densitat de *solvent flattening* o aplanament de solvent, el *histogram matching* i promitjat. Les fases es van estendre de 3.5 Å a 3 Å.

L'ús combinat de les tècniques per millorar les fases, junt amb l'extensió de les mateixes, ens va donar com a resultat un mapa de molt millor aspecte, on les regions de proteïna i les de solvent estaven clarament diferenciades (Figura 4.5). En la regió de proteïna, es podia distingir clarament la cadena principal, per bé que només algunes cadenes laterals eren identificables, majoritàriament les dels aminoàcids aromàtics.

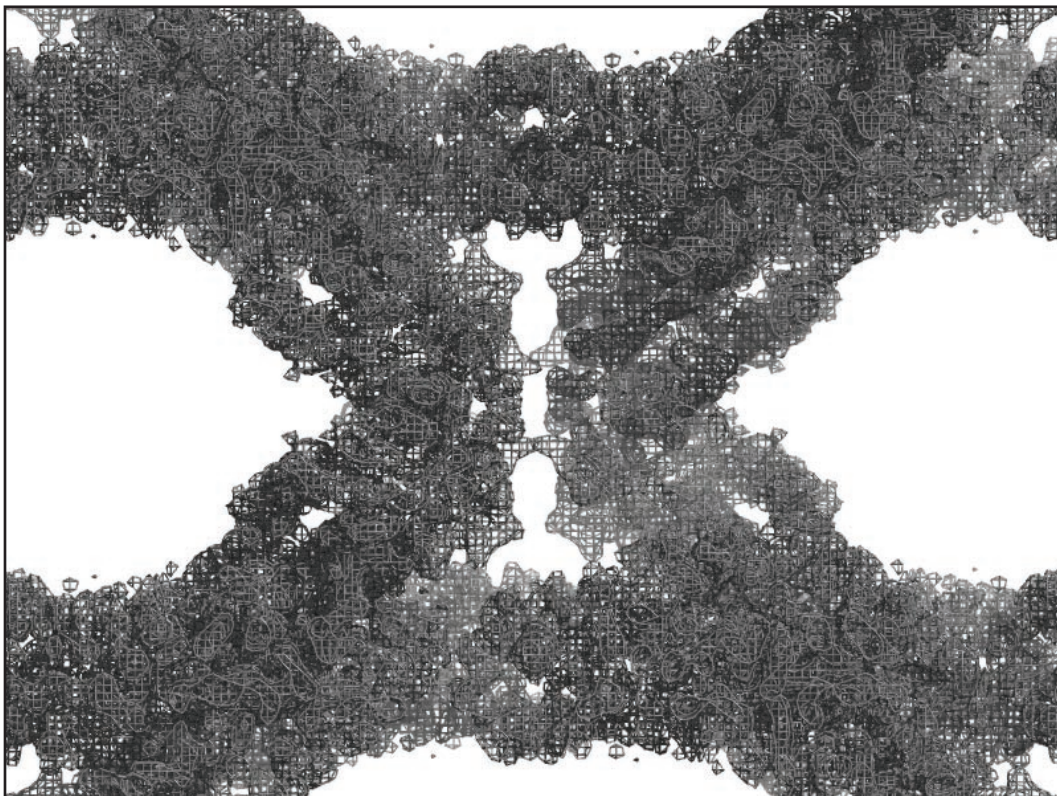


Figura 4.5: Mapa corresponent a la mateixa regió que en la Figura 4.4 després de la modificació de densitat i extensió de fases fins a 3 Å amb RESOLVE. La zona de solvent s'ha netejat, definint-se amb més precisió els marges amb la regió de proteïna.

4.2.5. Construcció del model a 3 Å

La construcció del model es va fer en 2 passos ben diferenciats. El primer pas va ser automàtic i va ser fet pel programa RESOLVE (Terwilliger, 2002). El resultat va ser una construcció del 59,4% (223 de 375) del model de proteïna (Figura 4.6). D'aquests 223 residus només 9 tenien cadena lateral, i el model parcial tenia moltes mancances, sobretot en les zones d'inici i final de regions d'estructura secundària.

El model inicial estava compost, essencialment, pels elements d'estructura secundària. El segon pas, totalment manual, va consistir en corregir el model construït, finalitzar la cadena principal i assignar a cada residu la seva cadena lateral i número dins la seqüència de la proteïna. Ens vam centrar en completar un sol dels

monòmers, combinant la utilització de parts ja traçades en els altres dos i construcció manual. L'assignació la cadena lateral va començar en els residus de densitat més clara, és a dir, els residus amb cadena lateral aromàtica. A partir d'aquests residus, bastant abundants en la proteïna, es va poder definir la seqüència correcta de la proteïna. En el moment en que la seqüència va ser assignada i van ser col·locades algunes cadenes laterals va començar el procés de refinat.

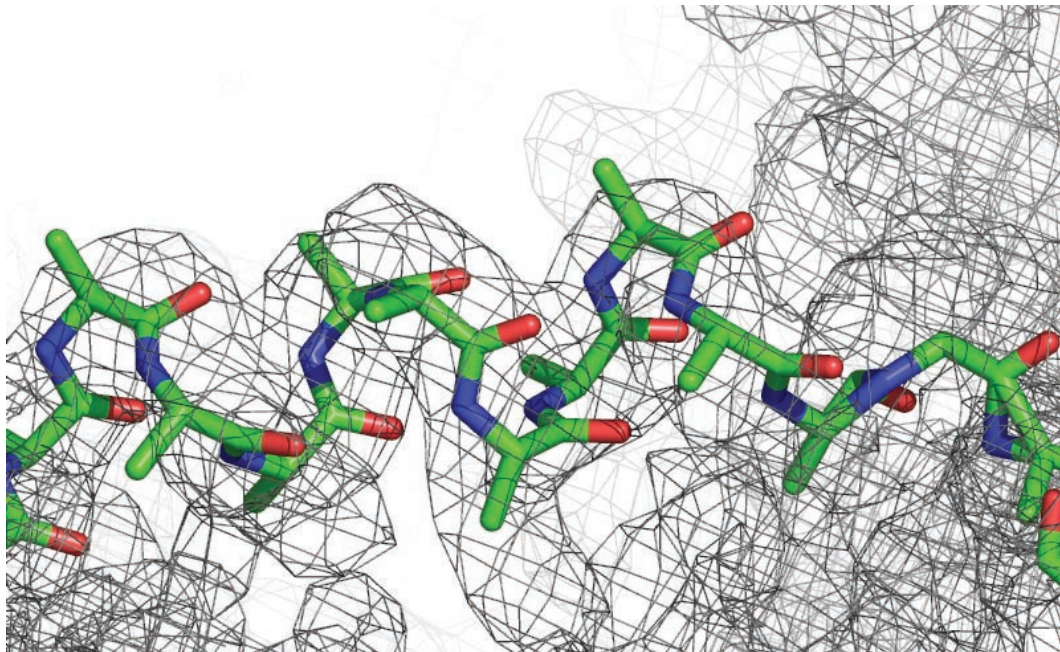


Figura 4.6: Detall d'una hèlix α del model traçat automàticament pel programa RESOLVE

L'únic monòmer del que disposàvem va ser utilitzat com a model de cerca amb el programa de reemplaçament molecular MOLREP (Vagin i Teplyakov, 1997). Vam utilitzar l'opció de cercar un model en un mapa de Fourier, calculat amb les fases ja disponibles, realitzant-se una correlació amb la densitat existent. Així vam poder situar correctament els 3 monòmers dins de la unitat asimètrica.

4.2.6. Afinament de l'estructura de la proteïna derivatitzada

El primer pas que vam realitzar en l'afinament va ser del tipus cos rígid, en el que cada protòmer es mou de manera rígida per fer-lo coincidir amb la densitat electrònica. En aquest tipus d'afinament no hi intervé cap tipus de minimització

energètica ni estereoquímica. Aquest tipus d'afinament també es realitza en l'últim pas del reemplaçament molecular. El vam fer utilitzant el programa CNS (Brüger et al, 1998).

La inspecció de la funció d'autorrotació (Figura 4.7) ens va indicar la presència de dos eixos de simetria no cristal·logràfica que ens relacionaven les tres molècules de la unitat asimètrica. Amb el programa CNS es van introduir les matrius que relacionaven cada protòmer amb els altres dos per tal de poder fer servir la simetria no cristal·logràfica des dels primers passos d'afinament.

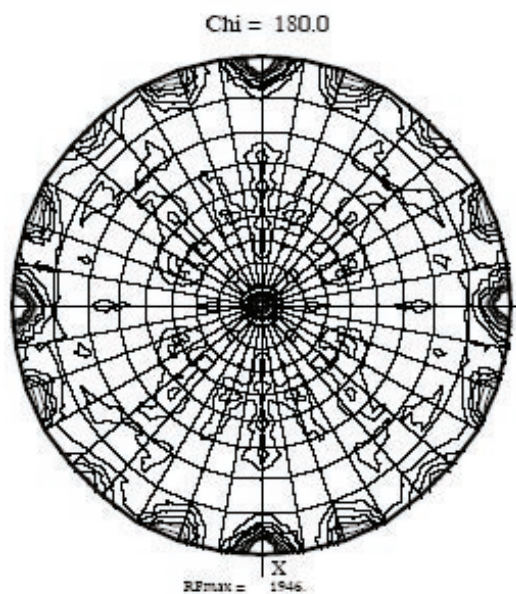


Figura 4.7: Funció d'autorrotació del cristall C222₁. Es poden observar els tres eixos cristal·logràfics binaris i, a més, 2 eixos, a $\chi=180^\circ$, $\psi=90^\circ$ i $\phi=22,5^\circ$ i $\phi=45^\circ$ corresponents a la simetria no cristal·logràfica.

El procediment habitual és introduir les restriccions de simetria no cristal·logràfica amb un cert pes associat, el qual anirem relaxant a mida que avança l'afinament.

En el procés d'afinament es van realitzar passos successius de minimització combinats amb la inspecció del model i dels mapes de densitat electrònica 2Fo-

Fc i Fo-Fc. Durant l'observació dels mapes i del model es van anar col·locant les cadenes laterals que corresponien a cada aminoàcid segons es veia una demanda de densitat en els mapes Fo-Fc i 2Fo-Fc.

Quan el model va estar complet, vam introduir l'afinament dels factors de temperatura, assignant a cada àtom un valor que indica la seva mobilitat tèrmica.

<i>Processat</i>	Derivat hexaclorplatínat		
	Pic	Inflexió	Remota
Grup espacial	C222 ₁	C222 ₁	C222 ₁
Cel·la	a=135.57, b=140.51, c=66.92 Å		
Longitud d'ona (Å)	1.07170	1.07210	0.97620
Reflexions mesurades	117195	87546	100928
Reflexions úniques	14526	14411	14928
Rang de resolució (Å)	55.184 - 2.9	55.184 - 2.9	55.184 - 2.9
Completeness (%)	97.4 (97.4)	98.9 (98.9)	100 (100)
Multiplicitat	8.0 (5.7)	5.9 (4.0)	6.7 (6.3)
Rmerge (%)	9.1 (44.3)	9.2 (47.3)	10.4 (48.6)
$\langle I \rangle / \langle \sigma(I) \rangle$	7.9 (1.7)	7.7 (1.6)	6.9 (1.5)
Ranòmal (%)	5.0 (22.3)	4.9 (27.2)	5.3 (22.9)
<i>Fasejat</i>			
Figura de mèrit	0.45		
Figura de mèrit (després de l'aplanament de solvent)	0.68		
<i>Afinament</i>			
Rang de resolució	55.0 – 3.0		
Reflexions	13192		
R _{treball}	23.94		
R _{free}	26.90		
Àtoms totals	2895		
Àtoms de proteïna	2892		
Àtoms de solvent	0		
Àtoms de Pt	3		
<i>R.m.s.d. dels valor objectiu</i>			
Enllaços (Å)	0.010		
Angles (°)	1.42		

Taula 4.5: Estadística de processat, fasejat i afinament del model de la proteïna p4 en la seva forma derivatitzada amb platí

El model es va completar i afinar en els tres monòmers presents en la unitat

asimètrica, i aquest contenia els residus Pro2 fins Pro115. Quan els indicadors de qualitat de l'afinament, els factors R, van mantenir-se estables vam concloure el procés d'afinament. El model de proteïna presentava unes estadístiques que es detallen en la Taula 4.5.

4.3. Processament de les dades de difracció $P2_12_12_1$

L'indexat de les dades de difracció de raigs X del cristall de proteïna p4 va mostrar que el cristall pertanyia a un grup espacial ròmbic primitiu i segons les extincions sistemàtiques es tractava del grup espacial $P2_12_12_1$. Les dades es van processar fins a 2 \AA i la cel·la unitat té uns paràmetres de $a=66.35$, $b=136.23$, $c=140.4 \text{ \AA}$ i $\alpha=\beta=\gamma=90^\circ$.

El càlcul del coeficient de Matthews (Matthews, 1968), ens va suggerir un valor de 6 molècules per unitat asimètrica i un contingut de solvent del 64,5 %. Veiem que la derivatització de la proteïna amb l'àtom pesant induïx un canvi en el grup espacial, tot i que la cel·la unitat pràcticament no canvia. Les estadístiques del processat, escalat s'inclouen en la Taula 4.6.

	Natiu-apo
Processat	
Grup espacial	$P2_12_12_1$
Cel·la unitat	$a=66.35$, $b=136.23$, $c=140.40$
Longitud d'ona (Å)	0.934
Reflexions mesurades	515736
Reflexions úniques	79464
Rang de resolució (Å)	54.9 – 2.0
Completeness(%)	94.2 (94.2)
Multiplicitat	6.3 (5.1)
Rmerge (%)	7.8 (30.9)
$\langle I \rangle / \langle \sigma(I) \rangle$	6.3 (2.2)

Taula 4.6.: Processat de les dades del cristall $P2_12_12_1$.
 $R_{merge} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_{ii} |I_i(hkl)|}$, per a tot el rang. On:
 hkl : índexs de les reflexions úniques
 $I_i(hkl)$: és la i mesura d'intensitat de la reflexió hkl
 $\langle I(hkl) \rangle$: és la mitja de les mesures de la reflexió hkl
Els valors entre parèntesi són els de la última capa

4.3.1. Resolució del problema de la fase pel mètode del reemplaçament molecular

Degut a la diferència de grups espacials entre el cristall de la proteïna p4 i el cristall de la proteïna p4 derivatitzat, no es podia fer un ús directe de les fases determinades per aquest últim cristall. Per això vam haver d'utilitzar el model de la proteïna derivatitzada, afinament fins a 3 Å, per resoldre el problema de la fase, per reemplaçament molecular, en el conjunt de dades de difracció del cristall de proteïna sense derivatitzar.

La cerca es va realitzar amb el programa AMoRe (Navaza, 1994) i com a model de cerca vàrem utilitzar un dímer de proteïna p4 en el rang de resolució 55-4 Å (Taula 4.7). Recordar que en el grup espacial $P2_12_12_1$ la unitat asimètrica és el doble de gran que en el grup espacial $C222_1$ i, per tant, tindrem el doble de molècules per unitat asimètrica, així que vàrem fer la cerca per 6 molècules, o 3 dímers.

En una primera cerca per un sol dímer els resultats obtinguts van ser:

	α	β	γ	tx	ty	tz	CCF	RF	CCI	Dist
PRO	173.8	48.5	183.1	0.0105	0.0062	0.0082	23.8	54.9	16.9	66.3
PRO	173.1	48.5	183.4	0.0048	0.1446	0.0079	18.0	56.6	12.3	43.9
PRO	175.0	48.4	181.9	0.0112	0.2372	0.0075	19.3	55.8	12.0	33.4
PRO	174.1	48.6	182.6	0.0037	0.3065	0.0080	18.0	56.2	11.8	36.5
PRO	173.7	48.7	182.9	0.0108	0.0375	0.0078	18.2	56.5	12.6	66.3
PRO	173.1	48.5	183.4	0.0113	0.4410	0.0080	17.5	56.5	12.1	61.5

Taula 4.7: Resultat de la cerca de posicions de reemplaçament molecular.

(α, β, γ): angles d'Euler; (tx, ty, tz): translacions

CCF: coeficient de correlació entre les F_o i les F_c del model

RF: R factor entre ells

CCI: coeficient de correlació amb les intensitats

Dist: distància entre els centres de masses entre molècules simètriques

PRO: solucions per la proteïna

Utilitzant la primera solució com a correcta, vam fixar-la i vam fer la cerca de nou (Taula 4.8). Cal destacar que utilitzant una solució correcta, les correlacions es veuen reforçades.

	α	β	γ	tx	ty	tz	CCF	RF	CCI	Dist
FPRO	176.1	48.5	182.2	0.0060	0.0060	0.0074	27.5	59.7	20.7	66.3
PRO	151.1	65.0	231.6	0.8236	0.8837	0.6929	34.7	50.3	32.1	36.9
FPRO	175.7	48.5	182.5	0.0044	0.0061	0.0069	27.5	59.7	20.7	66.3
PRO	151.0	64.5	231.9	0.8217	0.3818	0.1920	32.8	50.7	32.0	37.4

Taula 4.8: Resultats de la cerca de reemplaçament molecular amb una solució fixa.

(α, β, γ): angles d'Euler ; (tx,ty,tz): translacions

CCF: coeficient de correlació entre les Fo i les Fc del model

RF: R factor entre ells

CCI: coeficient de correlació amb les intensitats

Dist: distància entre els centres de masses entre molècules simètriques

FPRO: solucions fixes per la proteïna; PRO: solucions per la proteïna

Les dues solucions primeres són molt semblants, podria ser que les dues fossin solucions correctes. I ja, per últim, fixant les dues solucions i fent una última cerca (Taula 4.9):

	α	β	γ	tx	ty	tz	CCF	RF	CCI	Dist
FPRO	176.1	48.5	182.3	0.0029	0.0061	0.0084	27.5	59.7	20.7	66.3
FPRO	151.6	65.9	230.8	0.8249	0.8872	0.6941	31.3	51.9	29.8	36.9
PRO	151.0	63.8	232.2	0.8266	0.3736	0.1898	56.9	41.8	50.3	37.3

Taula 4.9: Resultats de la cerca de reemplaçament molecular amb dues solucions fixes.

(α, β, γ): angles d'Euler; (tx,ty,tz): translacions

CCF: coeficient de correlació entre les Fo i les Fc del model

RF: R factor entre ells

CCI: coeficient de correlació amb les intensitats

Dist: distància entre els centres de masses entre molècules simètriques

FPRO: solucions fixes per la proteïna; PRO: solucions per la proteïna

Efectivament, la solució de la tercera molècula correspon a la segona solució de la cerca per 2 dímers.

Un altre fet que sorprèn és que els dos últims dímers trobats pertanyen pràcticament a la mateixa rotació, però amb diferent translació separades aproximadament mitja cel·la en y i z. El fet d'introduir l'àtom pesant ha fet que les rotacions coincidissin exactament i les diferències en les coordenades y i z sigui exactament 0.5, provocant el canvi de grup espacial de $P2_12_12_1$ a $C222_1$.

4.3.2. Construcció i afinament del model a 2 Å

La inspecció dels mapes de densitat electrònica 2Fo-Fc i Fo-Fc amb el model obtingut per reemplaçament molecular, sense afinar, ens va mostrar que la solució era correcta.

Com ja havíem fet en el model a 3 Å, vam realitzar un afinament de tipus cos rígid, utilitzant el programa CNS, per tal d'acabar d'ajustar el model a la densitat i aconseguir un bon punt de partida per iniciar l'afinament més fi. En aquest pas, vam passar d'un factor R del 44,8 % a un del 43,1%, i un factor *R_{free}* del 44,3 % a un del 42,9 %, mostrant-nos que és un pas útil, encara que els programes de reemplaçament molecular també ho realitzin en el seu procés d'afinament de la solució.

Posteriorment, i després d'inspeccionar la funció d'autorotació (Figura 4.8), vam calcular les matrius que relacionen cada monòmer amb tots els demés i així poder introduir les restriccions de simetria no cristal·logràfica en l'afinament de l'estructura.

El següent pas va consistir en fer un afinament de tipus *simulated annealing* o rescalfament simulat. Tal i com s'explica en l'apartat de materials i mètodes, aquesta tècnica consisteix en una simulació d'un escalfament de la molècula fins a temperatures de uns milers de graus Kelvin (°K) per tal d'optimitzar geomètricament el model, superant barreres energètiques que impedeixen el canvi d'una conformació a una altra.

El model, després del *simulated annealing*, va presentar una gran millora en els indicadors de qualitat, els factors R, ja que van passar del 42,9 % al 35,4 %, i el factor *R_{free}* del 42,8 % al 36,9 %.

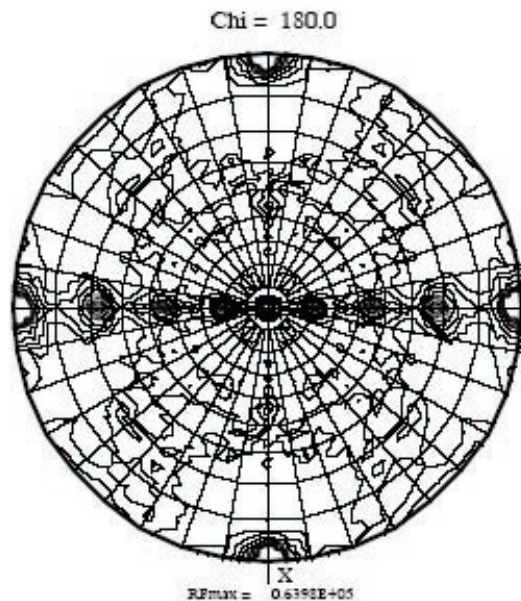


Figura 4.8: Funció d'autorrotació del cristall de p4 corresponent al grup espacial $p2_12_12_1$. Es mostren 3 eixos binaris locals a chi 180° i $\phi=90^\circ$ i $\psi=21'6''$; $\phi=90^\circ$ i $\psi=45^\circ$, i $\phi=90^\circ$ i $\psi=68,4^\circ$.

A partir d'aquest punt l'afinament va consistir en fer cicles successius de minimització alternats amb inspeccions manuals del model i de la seva concordança amb els mapes de densitat. Es va afegir un residu d'arginina en l'extrem carboxi-terminal que no s'observava en el model afinament fins a 3 Å.

Després de que s'estabilitzessin els factors R propers al 34 % i acabéssim de construir el model, vàrem fer una cerca de molècules d'aigua. En un primer pas es van trobar al 254 molècules d'aigua, observant-se una lleugera millora dels factors R.

Un cop traçat tot el model i iniciat el procés de cerca de molècules d'aigua, vàrem començar a afinar també els factors de temperatura, observant una millora substancial en els indicadors deixant-los, en un primer pas, al voltant del 28 %.

Es van cercar més molècules d'aigua varis cops més, comprovant una per una en els mapes de densitat electrònica que cada molècula interaccionés, mitjançant

una interacció de tipus pont d'hidrogen, amb algun àtom de la proteïna. Vàrem comprovar que el tipus d'àtom fos l'adequat per aquest tipus d'interacció (O o N) i que la distància fos la correcta (fins a 4 Å des de l'àtom d'oxigen de l'aigua fins a l'altre àtom implicat en l'interacció). En l'últim pas d'afinament teníem un total de 561 molècules d'aigua

L'afinament va prosseguir amb més cicles de minimització i afinament dels factors de temperatura fins a uns estadístics finals que es detallen en la Taula 4.10.

	Natiu
<i>Processat</i>	
Grup espacial	P2 ₁ 2 ₁ 2 ₁
Cel·la	a=66.35, b=136.23, c=140.40
Longitud d'ona (nm)	0.934
Reflexions mesurades	515736
Reflexions úniques	79464
Rang de resolució (Å)	54.902 – 2.0
Completeness (%)	94.2 (94.2)
Multiplicitat	6.3 (5.1)
Rmerge (%)	7.8 (30.9)
<I>/<σ(I)>	6.3 (2.2)
<i>Afinament</i>	
Rang de resolució (Å)	55.0 – 2.0
Reflexions	86693
R _{treball}	22.67
R _{free}	25.57
Àtoms totals	6401
Àtoms de proteïna	5839
Àtoms de solvent	562
<i>R.m.s.d. dels valors objectiu</i>	
Enllaços (Å)	0.007
Angles (°)	1.24

Taula 4.10: Estadística del processat i afinament de l'estructura de la proteïna p4 en la seva forma lliure

4.3.3. Descripció de l'estructura

La molècula de p4 té unes mesures aproximades de 45 x 20 x 20 Å³. L'estructura del monòmer, de 125 residus, està formada per una fulla β central de 5 cadenes antiparal·leles, que parteix l'estructura en dues parts. A la banda interna de la cadena β tenim 3 hèlixs α (α1-α3). L'hèlix α1 és la més llarga de totes i està involucrada

en la dimerització i en la unió al DNA, mentre que les altres dues estan parcialment implicades en la dimerització ($\alpha 3$), però el seu paper principal és el de formar el nucli hidrofòbic al voltant del qual s'acoblen els demés elements d'estructura secundària (Figura 4.9).

A la part externa de la fulla β , trobem els 2 elements més importants de l'estructura. Per una banda tenim un *loop* situat a l'extrem N-terminal de la molècula i que compren els residus 3-9, formant una estructura molt rígida estabilitzada per varis ponts d'hidrogen, molt important per realitzar el reconeixement específic de la seqüència de DNA (Rojo i Salas, 1990). Molt proper al *loop* trobem també una hèlix 3_{10} d'un sol gir (residus Leu12 a Glu14)

Per altra banda hi ha els residus 117-125, que els trobem desordenats en l'estructura i que tenen la important funció d'interaccionar amb la RNA polimerasa (Rojo et al. 1990).

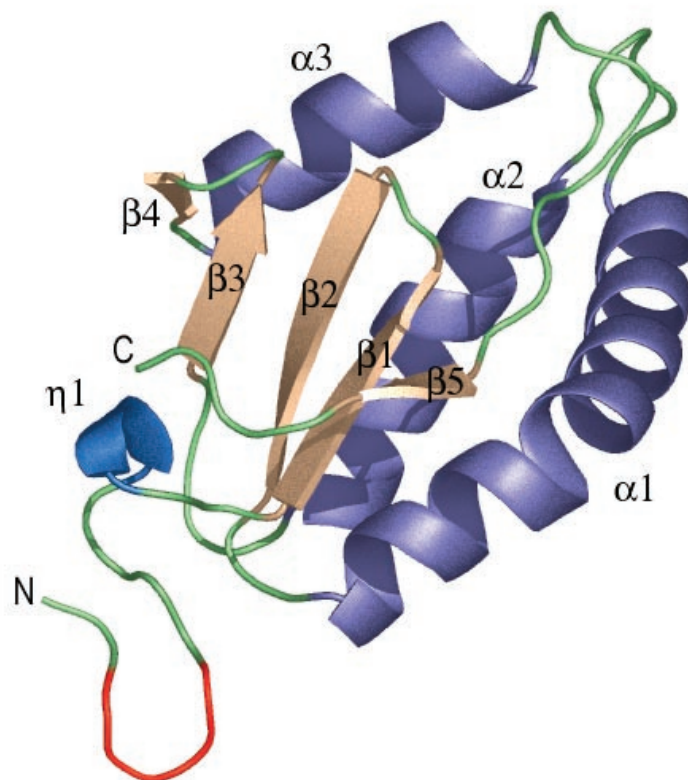


Figura 4.9: Estructura d'un monòmer de p4.

4.3.4. Dímer

En la dimerització hi intervenen una gran quantitat de contactes proteïna-proteïna amb una superfície d'interacció de 808 Å². És una superfície totalment hidrofòbica, evitant així la interacció amb el solvent. També mostra una gran complementarietat (Figura 4.10). Malgrat això, s'observa una cavitat interna a la zona de dimerització.

Els contactes entre monòmers consisteixen principalment en interaccions de tipus de van der Waals entre l'hèlix $\alpha 1$ d'un monòmer i l'hèlix $\alpha 3$ de l'altre monòmer.

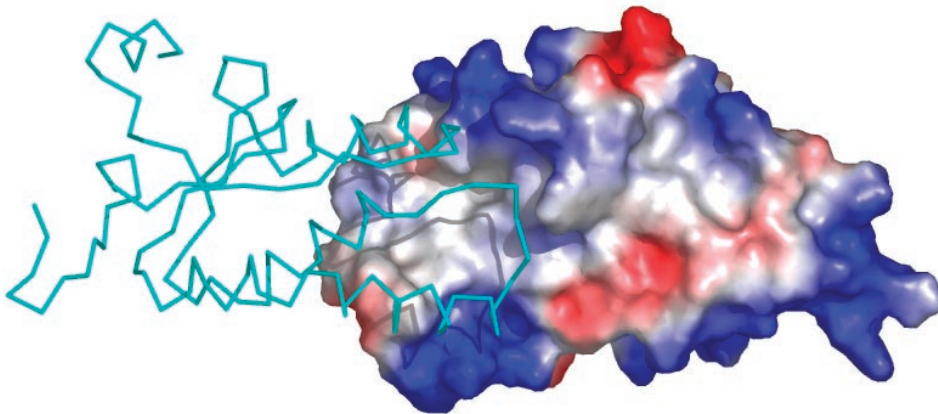


Figura 4.10: Representació en superfície de Connolly d'un dels monòmers i en varilles de l'altre, per mostrar la superfície de contacte entre monòmers, totalment hidrofòbica.

Troblem fins a un total de 17 residus implicats en la dimerització, obtenint-se 89 contactes de van der Waals entre els monòmers. Els residus Pro59 i Trp60 en el

loop que connecta les hèlix $\alpha 1$ i $\alpha 2$, i els residus Trp90 i His94 en l'hèlix $\alpha 3$, són els principals responsables de formar la regió hidrofòbica d'interacció amb l'altre monòmer. Diversos contactes, com Leu98-Leu98, Thr58-Leu65 i Ala56-Thr69, ajuden a tancar la interfase (Figura 4.11).

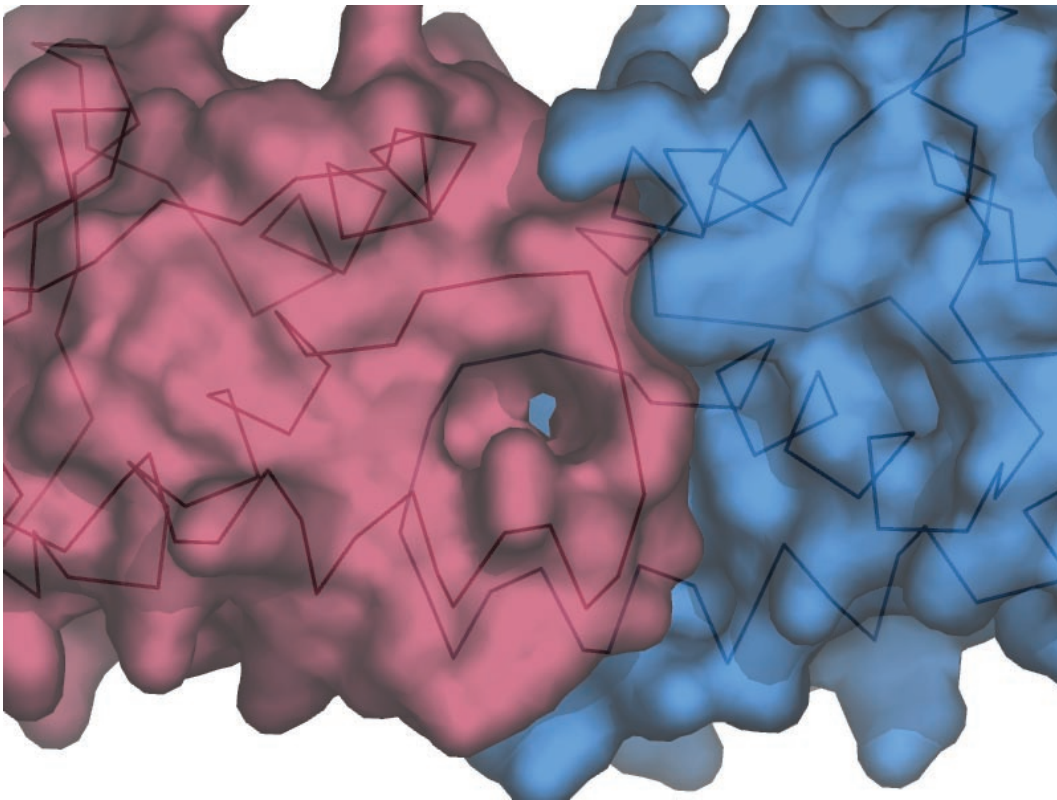


Figura 4.11: Detall de les superfícies de contacte entre monòmers, observant-se una gran complementarietat en les superfícies.

4.3.5. Cavitat interior

En la zona de dimerització de p4 es forma una cavitat interna sense solvent envoltada de residus hidrofòbics (Figura 4.12). En aquesta cavitat és on vam trobar el compost de platí, situat entre dos residus de metionina. Aquesta cavitat podria conferir a l'estructura una certa esponjositat, útil en el moment d'unir-se al DNA, ja que permetria un petit moviment entre monòmers.

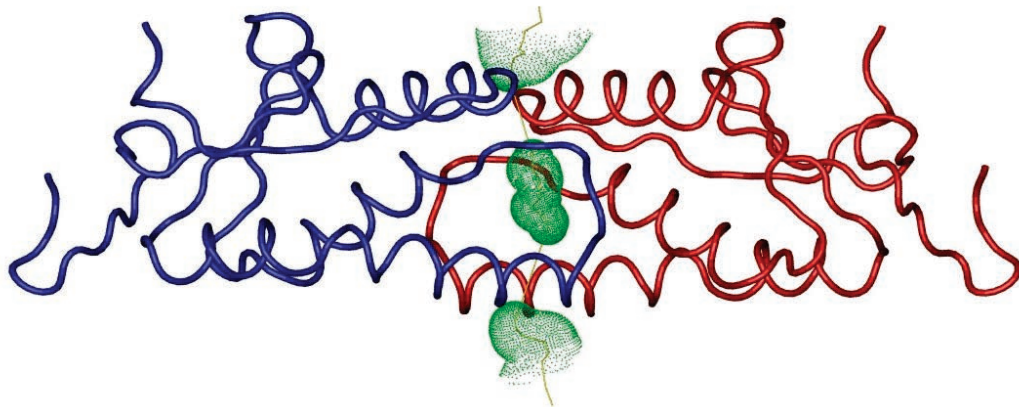


Figura 4.12: Representació de la cavitat interna present entre els dos monòmers. Calculat amb el programa HOLE (Smart *et al.*, 1993) i representat amb VMD (Humphrey *et al.*, 1996).

4.3.6. L'extrem N-terminal

La regió N-terminal, important en el reconeixement del DNA, forma una estructura molt compacta. Consisteix en un gir β de tipus I on participen moltes interaccions accessòries, conferint una gran rigidesa a aquesta zona (Figura 4.13).

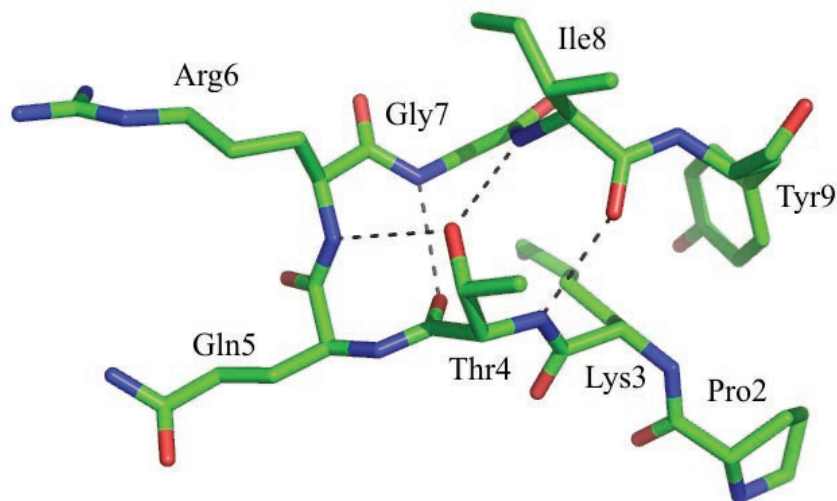


Figura 4.13: Esquema del *loop* N-terminal, on vàries interaccions li aporten una gran rigidesa.

El residu clau és el Thr4, que amb la seva cadena lateral estableix la conformació de la cadena principal mitjançant dues interaccions de pont d'hidrogen. La primera interacció és amb el grup NH que forma l'enllaç peptídic del residu Arg6, i la segona és amb el NH del residu Ile8.

Hi ha una tercera interacció realitzada pel residu Thr4, pròpia del gir β , on el grup CO interacciona amb el grup NH del residu Gly7.

4.3.7. Comparació entre monòmers

Si superposem els 6 monòmers presents a la unitat asimètrica (Figura 4.14) veiem que gairebé no hi ha diferències entre ells, obtenint-se un r.m.s.d. de 0.17 Å (el menor) i 0.38 Å (el major).

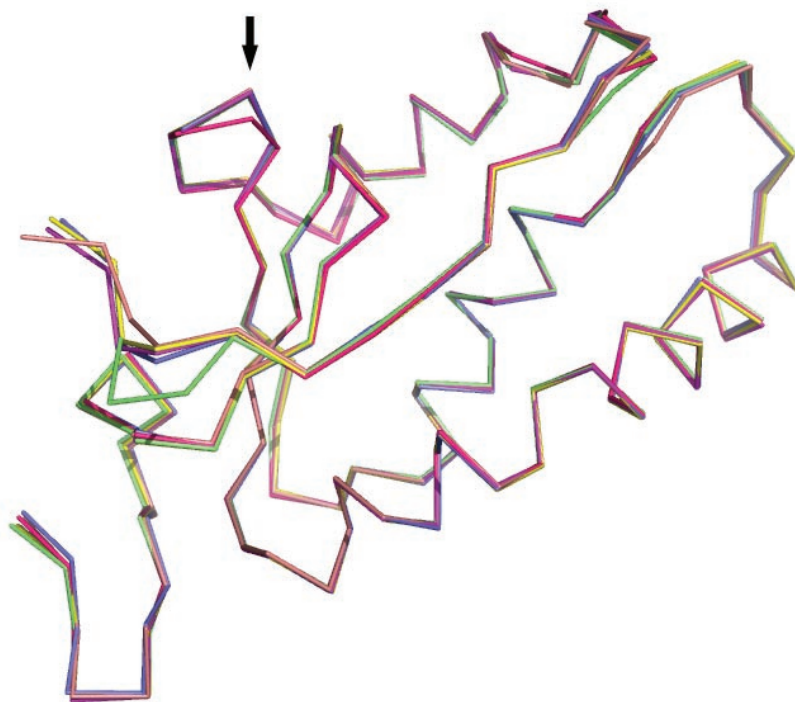


Figura 4.14: Superposició dels 6 monòmers presents en la unitat asimètrica. No s'observa cap diferència significativa entre ells, a excepció de la regió C-terminal i de la zona marcada amb una fletxa.

Observant amb més deteniment, veiem que les diferències més grans es troben a l'extrem C-terminal de la molècula. Com ja s'ha comentat anteriorment en el punt 4.3.3., la zona que compren els residus 118-125 no és visible en els mapes de densitat electrònica, degut probablement a que presenta una gran flexibilitat i els últims residus que veiem, Arg116 i Leu117, funcionarien com a frontissa.

A part, el residu Lys84 de la molècula F també mostra una petita variació, però aquest canvi és degut a un efecte de l'empaquetament del cristall, com es mostra en la Figura 4.15. El residu Lys84 contacta de manera diferent en cada monòmer, explicant així les diferències observades en la superposició de monòmers.

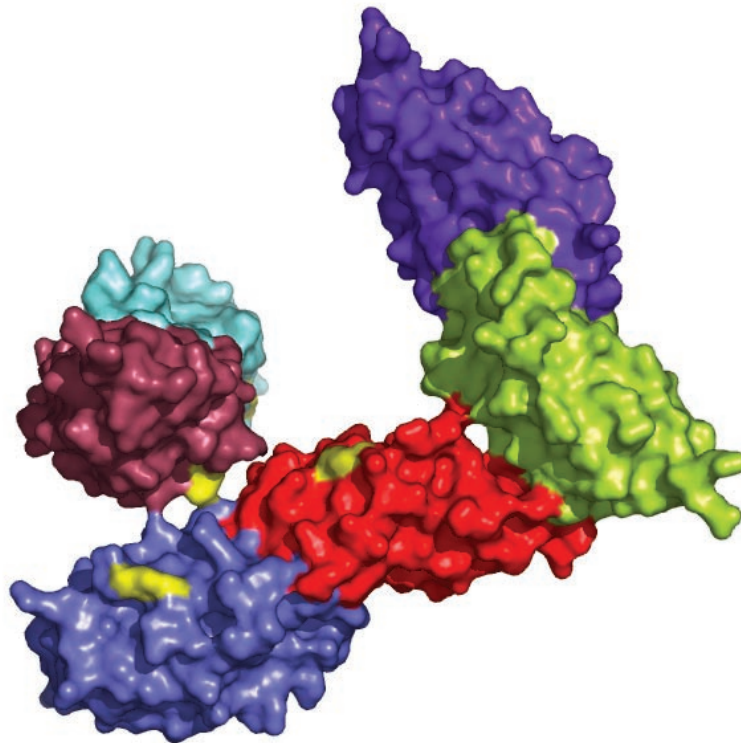


Figura 4.15: Esquema de la unitat asimètrica en el cristall $P2_12_12_1$, on es representa amb groc el residu Lys84. Es pot veure com en els diferents monòmers té contactes cristal·logràfics diferents. En morat es mostra la molècula F, on el residu Lys84 contacta amb una molècula de proteïna veïna.

4.3.8. Comparació amb altres estructures

Vam fer una cerca per intentar de trobar homòlegs estructurals de la proteïna p4. La cerca es va realitzar al servidor DALI (Holm i Sander, 1993; <http://www.ebi.ac.uk/dali/>). El resultat obtingut va ser que no hi havia cap tipus d'homologia estructural amb cap estructura coneguda.

4.3.9. Identificació de la regió d'unió a DNA

Degut a l'absència d'un motiu conegut d'unió a DNA i a cap homologia estructural amb cap proteïna resolta per cristal·lografia de raigs X, vam realitzar un càlcul de mapes de superfícies electrostàtiques amb el programa GRASP (Nicholls, 1991) per poder identificar zones amb densitat de càrrega (Figura 4.16). Observant les zones carregades positivament i de potencial interacció amb els grups fosfat (carregats negativament i constituents de l'esquelet del DNA), vam trobar una regió potencial d'unió al DNA.

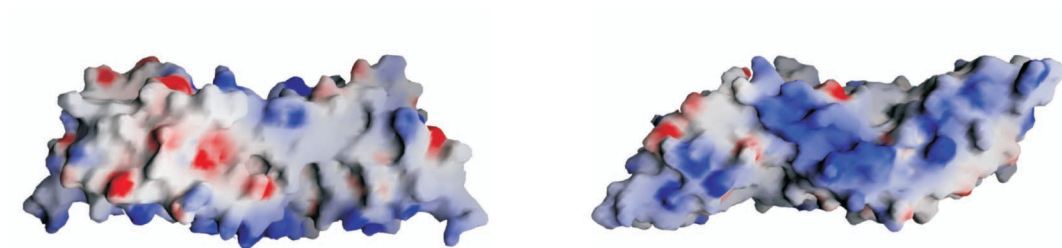


Figura 4.16: Visió lateral i inferior del dímer en una representació de superfície electrostàtica, elaborada amb GRASP (Nicholls, 1991). Es pot veure les 3 regions carregades positivament (color blau) en la visió inferior, que podrien interaccionar amb el DNA

Les zones identificades com a zones de potencial interacció corresponen a la zona del *loop* N-terminal i a l'hèlix $\alpha 1$. Aquesta potencial regió està d'acord amb les evidències que tenim de l'estructura, ja que s'ajustaria a un reconeixement d'una gran regió de DNA i presentaria l'extrem C-terminal lliure per interaccionar amb la subunitat α de la RNA polimerasa.

4.3.10. El diagrama de Ramachandran

El diagrama de Ramachandran (Ramachandran, 1968) mostra tots els residus en les regions permeses, a excepció del residu Thr22 (Figura 4.17).

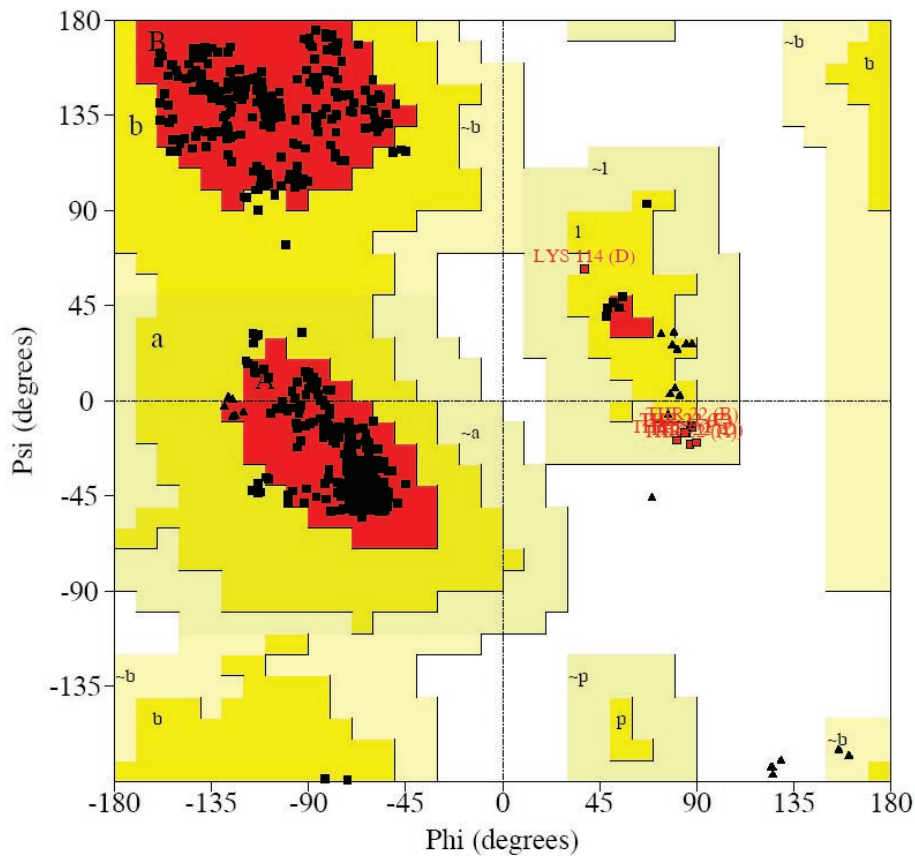


Figura 4.17: Diagrama de Ramachandran de la proteïna p4 en el seu estat no unit a DNA. La majoria de residus estan situats en les regions permeses, a excepció del residu Thr22

En estudiar detingudament el residu Thr22, ens adonem que està situat formant part del *loop* que connecta la cadena β_1 i la β_2 de la fulla β central. Entre les dues cadenes s'estableix un gir de tipus β , estabilitzat per la interacció de tipus pont d'hidrogen de la cadena lateral del residu Asn21 amb el nitrogen amida del residu Val24. També ens trobem que el residu Glu23 estabilitza la seva conformació per un pont d'hidrogen amb el residu Tyr96.

El residu Thr22 es troba al mig de dos residus fortament fixats (Figura 4.18), el que probablement compensa la seva conformació poc favorable energèticament.

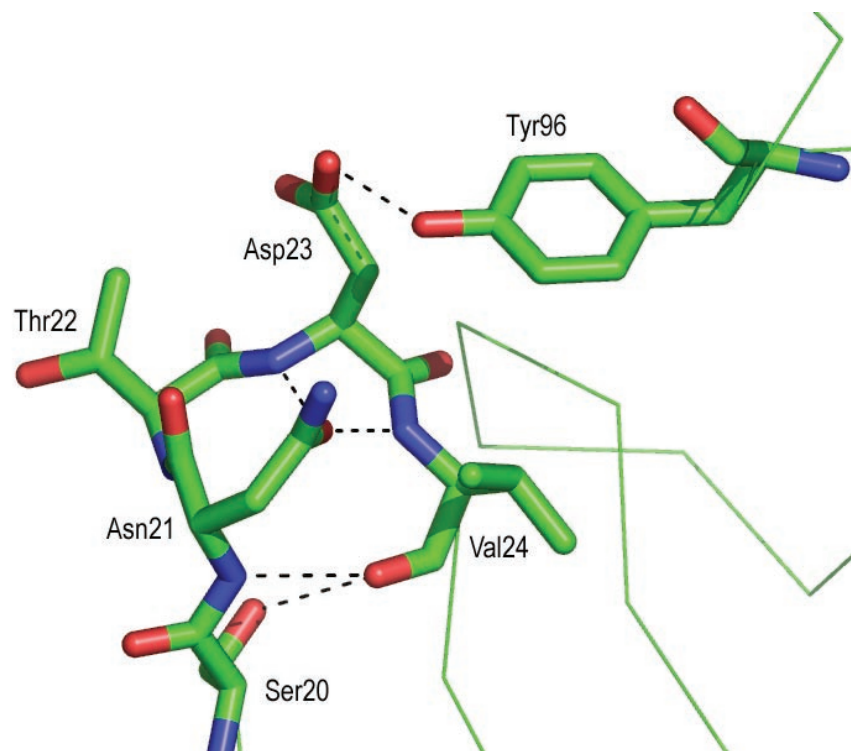


Figura 4.18: Representació en *sticks* de la regió on es troba el residu Thr22, envoltat de nombroses interaccions que estableixen la seva conformació en un estat d'alta energia.

Part II: El complex p4 - DNA

4.4.Preparació dels complexos amb DNA

Els complexos es van preparar tal i com es detalla a l'apartat de materials i mètodes. Vam utilitzar 4 jocs diferents d'oligonucleòtids, 2 d'ells amb extrems roms i 2 d'ells amb extrems cohesius. Utilitzar extrems roms o cohesius és un aspecte que pot influir en la cristal·lització, donant cristalls de qualitat diferent.

Els oligonucleòtids provats van ser:

p4_41roma

5'-TAA CTT TTT GCA AGA CTT TTT TAT AAA ATG TTG ACG TTT TT-3'

p4_41romb

5'-AAA AAC GTC AAC ATT TTA TAA AAA AGT CTT GCA AAA AGT TA-3'

p4_40coa

5'-AAC TTT TTG CAA GAC TTT TTT ATA AAA TGT TGA CGT TTT T-3'

p4_40cob

5'-AAA ACG TCA ACA TTT TAT AAA AAA GTC TTG CAA AAA GTT A-3'

p4_31coa

5'-TAA CTT TTT GCA AGA CTT TTT TAT AAA ATG T -3'

p4_31cob

5'-CAT TTT ATA AAA AAG TCT TGC AAA AAG TTA A-3'

p4_31roma

5'-TAA CTT TTT GCA AGA CTT TTT TAT AAA ATG T -3'

p4_31romb

5'-ACA TTT TAT AAA AAA GTC TTG CAA AAA GTT A-3'

En aparellar-los, obteníem els següents dúplex:

p4_41rom

```
TAA CTT TTT GCA AGA CTT TTT TAT AAA ATG TTG ACG TTT TT
AAT TAA AAA CGT TCT GAA AAA ATA TTT TAC AAC TGC AAA AA
```

p4_40co

```
AA CTT TTT GCA AGA CTT TTT TAT AAA ATG TTG ACG TTT TT
ATT GAA AAA CGT TCT GAA AAA ATA TTT TAC AAC TGC AAA A
```

p4_31co

```
TA ACT TTT TGC AAG ACT TTT TTA TAA AAT GT
AAT TGA AAA ACG TTC TGA AAA AAT ATT TTA C
```

p4_31rom

```
TAA CTT TTT GCA AGA CTT TTT TAT AAA ATG T
ATT GAA AAA CGT TCT GAA AAA ATA TTT TAC A
```

4.4.1. Purificació i cristal·lització dels complexos amb DNA

La cromatografia d'exclusió molecular (o filtració en gel) permet separar les proteïnes segons els seu pes molecular. No es fa cap canvi de tampó al llarg de la mateixa. Es va utilitzar per purificar els complexos de proteïna amb DNA de la proteïna i/o el DNA no units, i amb un tampó de salinitat moderada-baixa que ens va permetre fer els experiments de cristal·lització sense tornar a canviar el mateix. El cromatograma mostrava un pic molt clar corresponent al complex de p4-DNA (Figura 4.19).

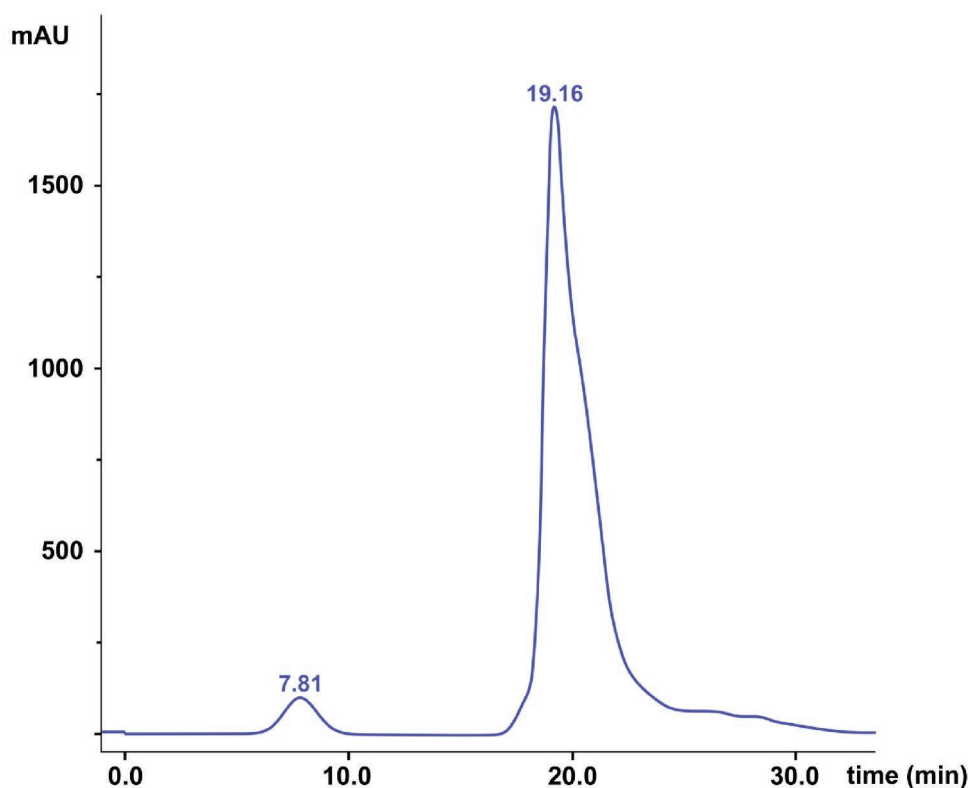


Figura 4.19: cromatograma de la columna de gel filtració Superdex 75 HR 10/30

Vàrem analitzar les fraccions eluïdes per electroforesi en gel de poliacrilamida (SDS-PAGE) i tinció amb plata, comprovant que les fraccions corresponents al pic contenen proteïna i DNA (Figura 4.20). Es van ajuntar les fraccions de la 14 a la 17, i es van concentrar mitjançant centrífons per fer els assajos de cristal·lització.

Les primeres proves de cristal·lització del complex format per p4 i per l'oligonucleòtid 41rom es van dur a terme a una concentració de 6 mg/ml. Els primers cristalls es van obtenir amb la condició número 35 del *screening* PROTEIN DNA, que conté cacodilat pH=7 100 mM; MPD al 30% (v/v) i MgCl₂ 100mM. A la Figura 4.21 es mostren aquests cristalls que presentaven una de les dimensions molt prima.

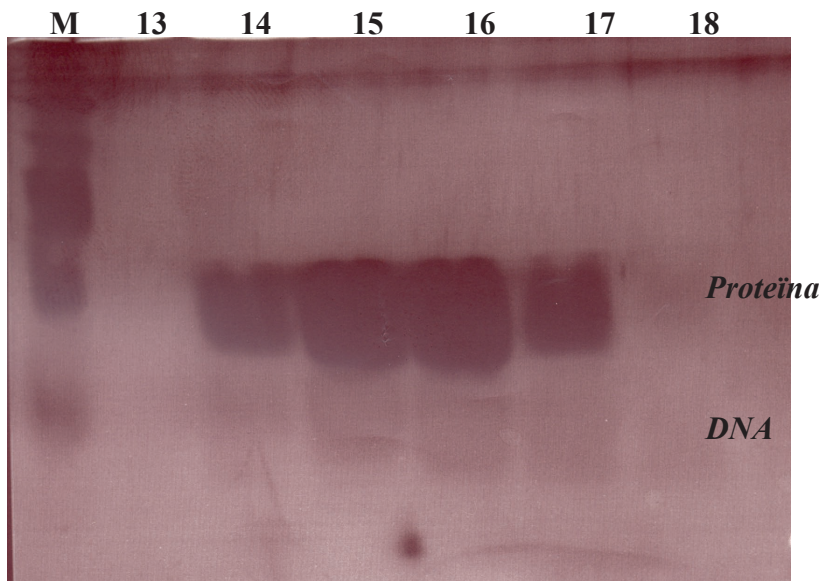


Figura 4.20: Gel tenyit amb plata. Es pot veure la proteïna tenyida de color fosc a la part superior i el DNA tenyit més tènueament, a la part inferior.

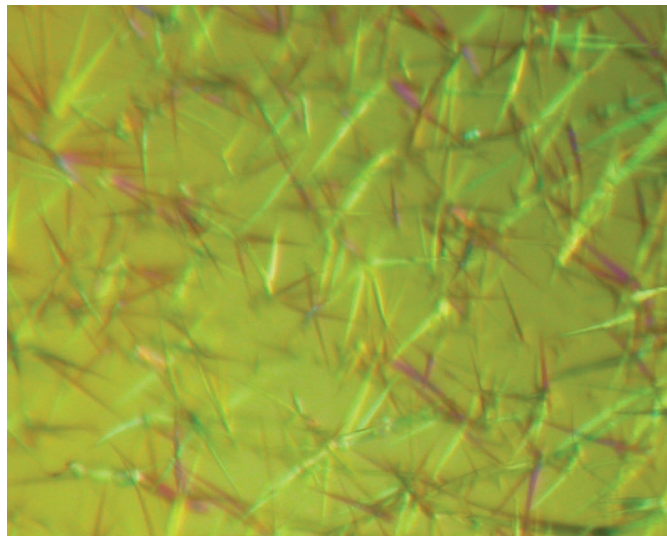


Figura 4.21: Cristalls inicials del complex p4-41rom

Aquests cristalls van ser provats a la línia ID13 del sincrotró ESRF (Grenoble), la difracció observada va ser extremadament anisotròpica, pel qual vàrem haver de millorar aquests cristalls.

Es varià la concentració de precipitant, el pH de la solució i el volum de la gota; però no s'obtingueren resultats positius. Una altra condició del *screening* PROTEIN DNA va donar-nos cristalls amb la mateixa morfologia. Va ser la condició número 75, que conté Tris pH=8,5 100 mM; PEG 2K MME al 20% (v/v) i MgCl₂ 100 mM. En aquest cas, l'exploració de les condicions al voltant d'aquesta ens va permetre d'optimitzar els cristalls. La condició final va ser Tris pH=8,5 50 mM; PEG 2K MME al 20% (v/v) i MgCl₂ 80 mM (Figura 4.22).

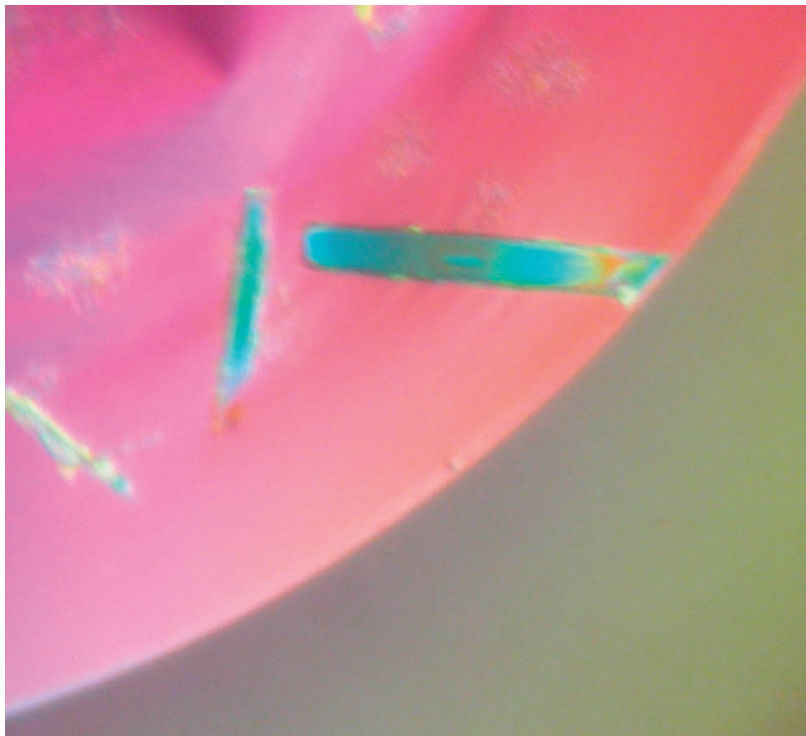


Figura 4.22: Cristalls millorats del complex p4-41rom

Amb l'oligonucleòtid 40co no es va aconseguir cap cristall, després de realitzar els 2 *screenings* específics de complexos de proteïna amb DNA, com són el Natrix (Hampton) i el PROTEIN-DNA.

Amb l'oligonucleòtid 31co es van fer proves de cristallització a una concentració de 7 mg/ml ; obtenint-se cristalls en les condicions del PROTEIN DNA que contenen Tris pH=8 50 mM; PEG 3K al 12,5% (v/v) i MgCl₂ 100 mM, i en una altra que contenia Tris pH=8,5 50 mM; PEG 4K al 20% (v/v) i KCl 200 mM. Aquests cristalls tenien una morfologia de prisma hexagonal (Figura 4.23 i 4.24).

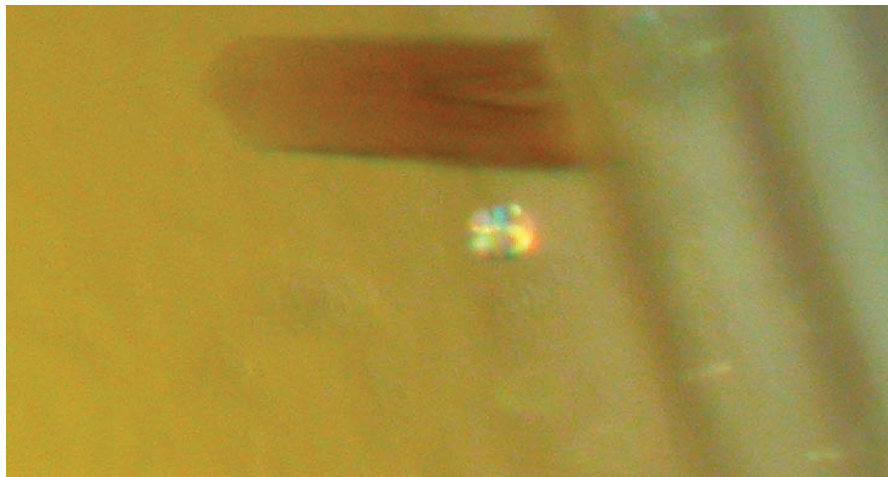


Figura 4.23: Cristalls hexagonals del complex p4-31co, en créixer presentaven un buit a l'interior.

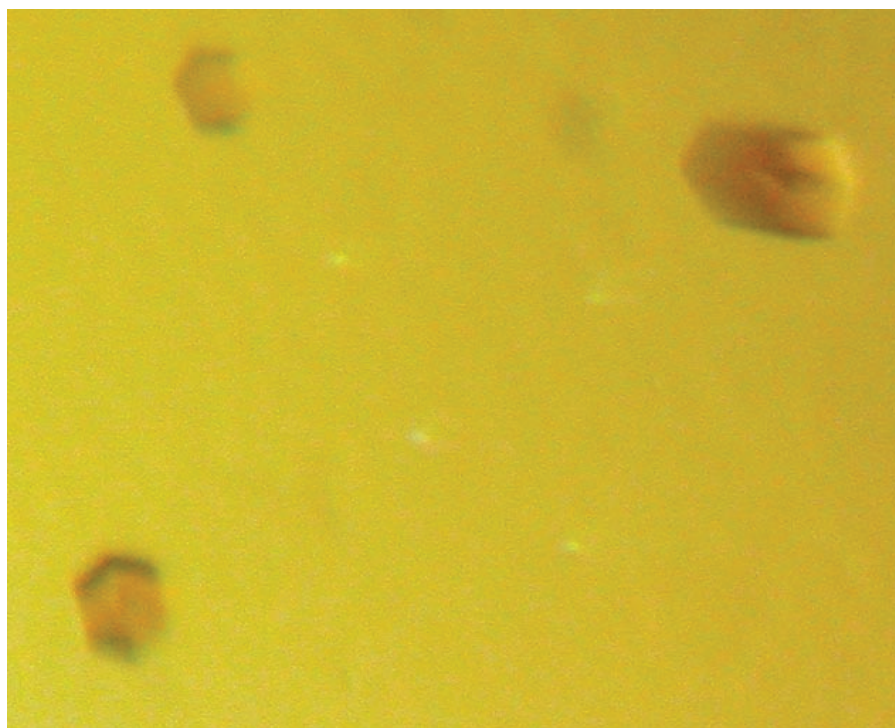


Figura 4.24: Cristalls del complex p4-31co. Un cop havien crescut exteriorment, llavors creixien interiorment, obtenint-se un cristall compacte.

4.5. Processament de les dades de difracció P2

Dels dos complexos provats, únicament els cristalls del que contenia l'oligonucleòtid 41rom van difractar a alta resolució. Aquest oligonucleòtid conté la seqüència corresponent al lloc 3 d'unió de p4.

Es van indexar les dades de difracció de raigs X d'aquest cristall arribant a la conclusió que pertanyia al grup espacial monoclínic primitiu P2. Els cristalls van difractar els raigs X fins a 2.7 Å i tenien una cel·la unitat de paràmetres $a=52.4$; $b=27.8$, $c=192.8$ Å i $\beta=97.4^\circ$. El càlcul del paràmetre de Matthews (Matthews 1968) ens donà uns valors de $V_m=2.8$ Å³ Da⁻¹, considerant 1 complex per unitat asimètrica i un contingut en solvent del 56.1 %.

En el processament de les dades del complex no es va seguir el mateix protocol que en els anteriors cristalls degut al procediment que s'havia seguit en la recollida de les dades. Aquest cristall va ser recollit a la línia ID-13, la qual té una òptica microfocalitzadora que permet recollir dades de cristalls molt petits o, com és el cas, molt estrets. La contrapartida és que el cristall pateix extraordinàriament el dany per radiació, observant-se una ràpida decaiguda en la difracció dels raigs X al llarg de la recollida de dades. Per superar aquest problema, durant el procés de recollida es va recentrant el cristall i irradiant en diferents punts en quant s'observa una pèrdua evident en la difracció. Per assegurar-nos de no tenir problemes en el processament i l'escalat de les dades, cada vegada que fem un recentrat del cristall, la recollida de dades es va reemprendre 4 ó 5 graus abans de l'aturada, obtenint-se un solapament entre els diferents paquets de dades.

El processament es va fer amb el programa XDS (Kabsch, 1993), utilitzant els 9 segments per separat, els quals no es van ajuntar fins a l'escalat de les dades. Les dades finals es detallen en la Taula 4.11.

	Complex amb DNA
Processat	
Grup espacial	P2
Cel·la unitat	a=52.4, b=27.8, c=192.8, β = 97.4
Longitud d'ona (Å)	0.978
Reflexions mesurades	63103
Reflexions úniques	15759
Rang de resolució (Å)	30.0 – 2.7
Completeness(%)	99.0 (96.8)
Multiplicitat	4 (2.97)
Rmerge (%)	12.5 (79.8)
$\langle I \rangle / \langle \sigma(I) \rangle$	8.96 (2.52)

Taula 4.11.: Processat de les dades del cristall P2.

$R_{merge} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_{ii} |I_i(hkl)|}$, per a tot el rang. On:

hkl : índexs de les reflexions úniques

$I_i(hkl)$: és la i mesura d'intensitat de la reflexió hkl

$\langle I(hkl) \rangle$: és la mitja de les mesures de la reflexió hkl

4.5.1. Resolució del problema de la fase pel mètode del reemplaçament molecular

A l'hora d'afrontar la resolució del problema de la fase es va optar per utilitzar el mètode del reemplaçament molecular, donat de que disposàvem del model de la proteïna lliure, resolta prèviament i afinada fins a 2 Å.

4.5.1.1. Tria dels models de cerca

A la cel·la unitat, presumiblement, tenim un complex de p4 amb el seu promotor natural, compostat d'un dímer de p4 i una molècula de DNA de 41 parells de bases. Així doncs vam utilitzar dos models de cerca. Per una banda un dímer de proteïna p4. Per altra banda havíem de trobar una solució pel fragment de DNA. Per aquesta part del complex no es disposava d'un model tan clar com pel cas de la proteïna. Sabíem que el fragment de DNA es trobava corbat (Rojo, 1990). Aquest fet ens va fer desistir d'utilitzar un model per tot el fragment de DNA, ja que la forma de la molècula diferiria molt de la d'un DNA ideal de 41 parells de bases.

Ens vam decantar per utilitzar un fragment de DNA (Figura 4.25) de 10 parells de bases en una conformació ideal de B-DNA generat amb el programa TURBO-FRODO (Roussel i Cambilleau, 1989)

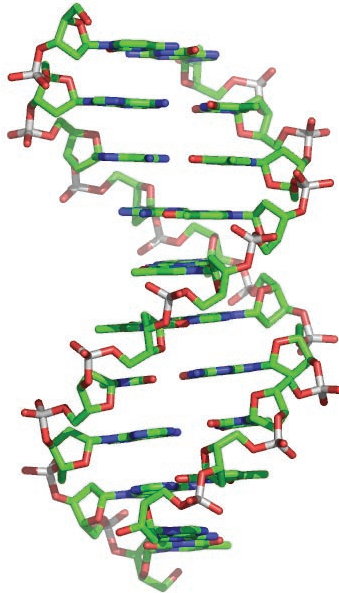


Figura 4.25: Model de DNA ideal de 10 parells de bases en conformació B utilitzat com a model de cerca en el reemplaçament molecular.

Això ens permetria trobar solucions parcials per la molècula de DNA, podent-se buscar més d'una molècula.

En una primera cerca (Taula 4.12), vam utilitzar només el model del dímer de proteïna. Vam utilitzar el model obtingut en la solució del reemplaçament molecular per calcular les fases i vam inspeccionar els mapes de densitat electrònica juntament amb el model de proteïna per comprovar que es tractava de la solució correcta.

A més de la densitat corresponent a la proteïna també hi havia certa densitat extra, en una disposició que suggeria una “escala de cargol”, típica de la densitat observada en el DNA, tot i que no era un bon mapa per col·locar manualment els diferents nucleòtids de la molècula de DNA.

α	β	γ	tx	ty	tz	CCF	RF	CCI	Dist
141.1	164.7	64.8	0.4260	0.000	0.2499	41.3	45.1	40.9	27.8
138.0	13.9	118.7	0.0730	0.000	0.2518	40.9	45.3	40.8	27.8
141.5	165.3	64.7	0.0744	0.000	0.2711	31.0	49.6	34.7	27.8
140.3	14.0	116.0	0.3967	0.000	0.2847	30.1	49.5	32.7	27.8
142.3	164.5	65.8	0.0306	0.000	0.4331	30.8	49.6	34.5	26.9
138.4	13.6	118.1	0.4549	0.000	0.0696	30.2	50.2	33.9	27.8
139.3	13.8	116.9	0.1617	0.000	0.4374	29.9	49.6	31.7	27.8

Taula 4.12.: Solucions de reemplaçament molecular per la proteïna.

(α, β, γ): angles d'Euler; (tx, ty, tz): translacions

CCF: coeficient de correlació entre les Fo i les Fc del model

RF: R factor entre ells

CCI: coeficient de correlació amb les intensitats

Dist: distància entre els centres de masses de les molècules simètriques més properes

L'estratègia a seguir va ser fer cerques individuals de manera iterativa i fixar les solucions correctes. Així doncs, un cop fixada la solució pel dímer de p4 vam prosseguir amb un fragment de 10 parells de bases de DNA. Els resultats de la cerca amb AMoRe es mostren a la Taula 4.13.

	α	β	γ	tx	ty	tz	CCF	RF	CCI	Dist
FPRO	140.9	165.1	64.6	0.4276	0.0005	0.2499	40.5	45.5	40.5	27.8
DNA	63.3	23.4	205.7	0.1262	0.2358	0.9146	49.1	43.4	50.4	27.8
FPRO	140.4	165.2	64.1	0.4270	-0.0001	0.2498	40.5	45.5	40.5	27.8
DNA	22.0	51.9	238.5	0.1079	0.1870	0.8990	48.7	43.3	49.0	27.8
FPRO	140.5	164.9	64.2	0.4268	0.0021	0.2497	40.5	45.5	40.5	27.8
DNA	11.3	110.3	59.4	0.2663	0.5520	0.0207	47.9	44.2	49.5	26.0
FPRO	141.0	165.0	64.7	0.4280	0.0012	0.2498	40.5	45.5	40.5	27.8
DNA	149.8	27.8	127.0	0.1396	0.2839	0.9285	47.4	44.1	47.6	27.8
FPRO	141.7	165.0	65.3	0.4268	0.0003	0.2497	40.5	45.5	40.5	27.8
DNA	127.0	24.4	328.0	0.1006	0.2460	0.9111	45.7	44.7	48.6	27.8

FPRO	140.2	165.1	63.9	0.4277	0.0006	0.2499	40.5	45.5	40.5	27.8
DNA	168.0	142.2	46.8	0.7597	0.5100	0.9966	47.0	44.2	49.1	25.9
FPRO	140.6	164.8	64.3	0.4267	0.0044	0.2498	40.5	45.5	40.5	27.8
DNA	146.1	37.5	111.2	0.0687	0.8757	0.7709	46.7	44.3	48.3	25.4
FPRO	140.7	164.6	64.6	0.4270	0.0036	0.2502	40.5	45.5	40.5	27.8
DNA	71.3	14.8	181.1	0.0764	0.8354	0.7568	48.0	43.9	49.1	25.9

Taula 4.13.: Solucions de reemplaçament molecular pel DNA, fixant la solució per la proteïna.
 (α, β, γ) : angles d'Euler; (tx, ty, tz) : translacions
 CCF: coeficient de correlació entre les F_o i les F_c del model
 RF: R factor entre ells
 CCI: coeficient de correlació amb les intensitats
 Dist: distància entre els centres de masses de les molècules simètriques més properes
 FPRO: solució fixada per la proteïna; DNA: solució pel DNA

La comprovació de les solucions dels fragments de DNA les vam fer de diferent manera que la del dímer de proteïna. Les solucions obtingudes no les vam utilitzar per calcular les fases, sinó que vam utilitzar el mateix mapa que havíem calculat a partir del model del dímer de proteïna.

Com que la proteïna era el component conegut del complex, vam preferir que les fases amb les que calculàvem els mapes fossin exclusivament provinents de la proteïna. Així podíem superposar la densitat extra que veiem en el mapa de densitat calculat pel dímer de proteïna amb les diferents solucions obtingudes per reemplaçament molecular pel model de DNA.

Amb el primer fragment es va veure un bon ajust a la densitat electrònica, sobretot en la zona dels grups fosfat (Figura 4.26).

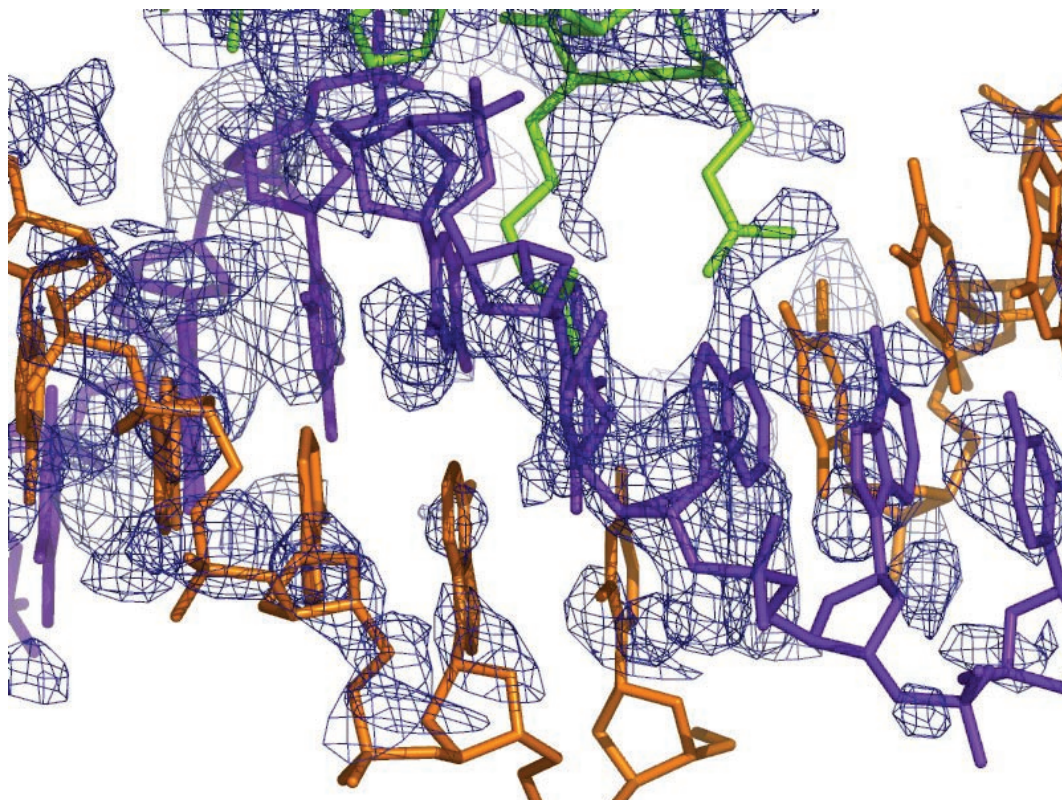


Figura 4.26: Solució obtinguda mitjançant reemplaçament molecular de la molècula de DNA. Es pot observar correspondència amb la densitat, sobretot en les bases i els fosfats.

Vam continuar amb el procés iteratiu de cerca de solucions, fixant ara, tant el dímer de proteïna p4 com el primer fragment de DNA, obtenint el resultat mostrat a la Taula 4.14.

	α	β	γ	tx	ty	tz	CCF	RF	CCI	Dist
FPRO	140.8	164.6	64.5	0.4278	0.0008	0.2500	40.5	45.5	40.5	27.8
FDNA	59.3	22.9	208.9	0.1272	0.2414	0.9144	47.3	44.2	49.6	27.8
DNA	62.3	14.4	188.3	0.0724	0.8381	0.7563	54.5	42.1	55.4	26.4
FPRO	140.1	165.2	63.8	0.4277	0.0034	0.2498	40.5	45.5	40.5	27.8
FDNA	59.4	24.6	209.8	0.1297	0.2392	0.9157	47.3	44.2	49.6	27.8
DNA	10.4	111.1	58.9	0.2638	0.5518	0.0191	52.0	43.8	54.4	22.7

FPRO	140.5	165.0	64.3	0.4263	0.0043	0.2497	40.5	45.5	40.5	27.8
FDNA	61.5	22.6	207.6	0.1305	0.2347	0.9138	47.3	44.2	49.6	27.8
DNA	146.1	36.2	110.2	0.0690	0.8763	0.7727	53.3	43.1	54.6	26.6
FPRO	139.9	165.3	63.6	0.4283	0.0002	0.2499	40.5	45.5	40.5	27.8
FDNA	62.7	23.9	206.8	0.1320	0.2342	0.9158	47.3	44.2	49.6	27.8
DNA	168.9	145.8	49.2	0.7620	0.5183	-0.0008	51.7	43.8	53.8	17.9
FPRO	140.6	165.0	64.4	0.4266	0.0013	0.2498	40.5	45.5	40.5	27.8
FDNA	61.2	24.2	208.8	0.1236	0.2345	0.9157	47.3	44.2	49.6	27.8
DNA	75.7	160.7	327.9	0.8904	0.9793	0.1672	49.3	44.1	50.6	17.2
FPRO	141.0	165.3	64.8	0.4277	0.0031	0.2498	40.5	45.5	40.5	27.8
FDNA	61.9	24.1	207.5	0.1252	0.2387	0.9153	47.3	44.2	49.6	27.8
DNA	11.1	108.4	60.2	0.4451	0.4763	0.0640	48.4	45.0	52.2	23.6

Taula 4.14. Solucions de reemplaçament molecular pel DNA, fixant la proteïna i la primera solució pel DNA

(α, β, γ): angles d'Euler; (tx, ty, tz): translacions

CCF: coeficient de correlació entre les Fo i les Fc del model

RF: R factor entre ells

CCI: coeficient de correlació amb les intensitats

Dist: distància entre els centres de masses de les molècules simètriques més properes

FPRO: solució fixada per la proteïna; DNA: solució pel DNA; FDNA: solució fixada pel DNA.

Altra vegada vam observar la nova solució en el mapa calculat per la proteïna, ajustant-se correctament amb la densitat electrònica, excepte algunes bases mal col·locades. Ja es començà a veure que el DNA en conformació ideal de B-DNA no ens seria útil per trobar totes les solucions ja que s'intuïa una conformació corbada, tal i com s'esperava.

No vàrem trobar cap més solució vàlida pels fragments de DNA, tot i no tenir el model global completat (Figura 4.27). La decisió a seguir va ser començar l'afinament amb el model incomplet però fiable, i la resta de model ja es faria manualment conforme anés apareixent densitat positiva en les zones on el model era incomplet.

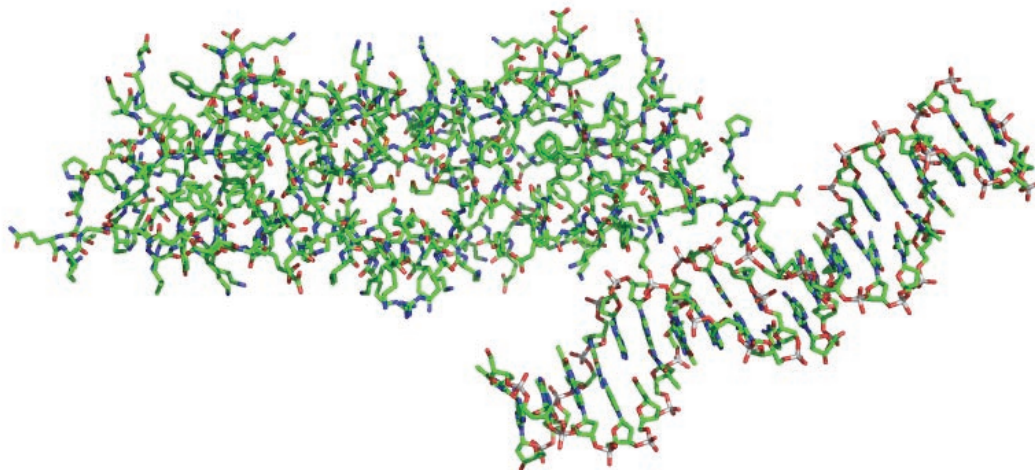


Figura 4.27: Model obtingut a partir de les tres solucions vàlides de reemplaçament molecular.

4.5.2. Afinament i finalització del model de l'estructura

Vàrem començar l'afinament per tal de poder optimitzar les fases disponibles i millorar els mapes, i que ens permetés poder finalitzar la construcció del model més clarament. L'afinament va començar amb un model de proteïna igual que l'afinament a 2 Å i 2 segments de DNA, de 10 parells de bases cada un. Així doncs quedarien 21 parells de bases per construir.

Es van calcular les matrius de simetria no cristal·logràfica que relacionaven els dos monòmers de proteïna (Figura 4.28) per aplicar restriccions durant l'afinament. Juntament amb aquestes restriccions en vàrem afegir una d'addicional entre les bases complementàries del DNA per tal de no perdre la interacció.

En un primer pas d'afinament, com ja havíem fet anteriorment, es va procedir a un afinament de tipus cos rígid. Monitoritzant els factors R, observàrem una lleugera millora, passant del 51,2 % al 48,3% en el factor R i del 50,3 % al 48% en el factor Rfree. Un segon pas de *simulated annealing* i de minimització va fer un nou ajust del model en la densitat, obtenint-se un R factor final de 41,1 % i un Rfree de 46,5 %.

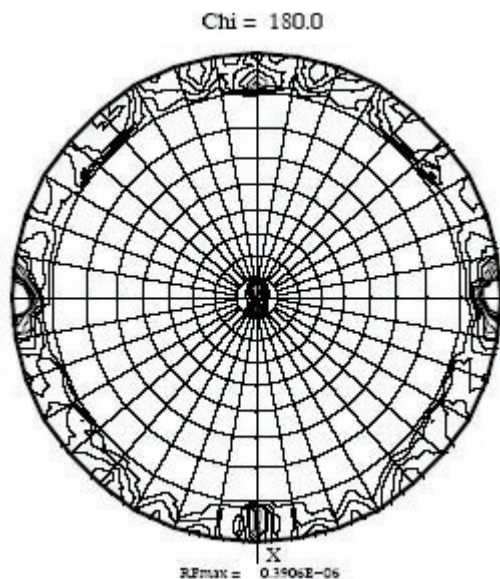


Figura 4.28: Funció d'autorrotació del cristall de p4 en complex amb una molècula de DNA. Es mostra un eix binari local a $\chi=180^\circ$, $\psi=86,2^\circ$ i $\phi=0^\circ$

Partint del model en aquest punt, es van calcular els mapes de densitat electrònica que ja van permetre finalitzar la construcció del model del DNA. Aquesta construcció es va fer utilitzant fragments de 3, 2 i 1 nucleòtid i es col·locant-los manualment encaixant-los en la densitat del mapa.

En aquest moment la seqüència del fragment de DNA no corresponia amb la del cristall, sinó que era aleatòria. A la resolució a la que treballàvem (2.7 Å) era difícil saber la seqüència per observació directa de la densitat electrònica, així que vàrem observar l'empaquetament del cristall per determinar on començava i on acabava el fragment de DNA, combinant-ho amb les dades bioquímiques que teníem sobre els llocs d'unió de p4.

En diversos complexos s'ha observat la formació d'hèlices contínues de DNA que dirigeixen l'empaquetament del cristall fent molt complicat l'assignació de la seqüència. En el nostre cas no era així, sinó que el fragment de DNA contactava, per un extrem amb un altre fragment de DNA (Figura 4.28), just on hi ha l'eix binari de simetria i per tant les direccions de les cadenes són oposades. Per l'altre extrem contactava amb la proteïna.

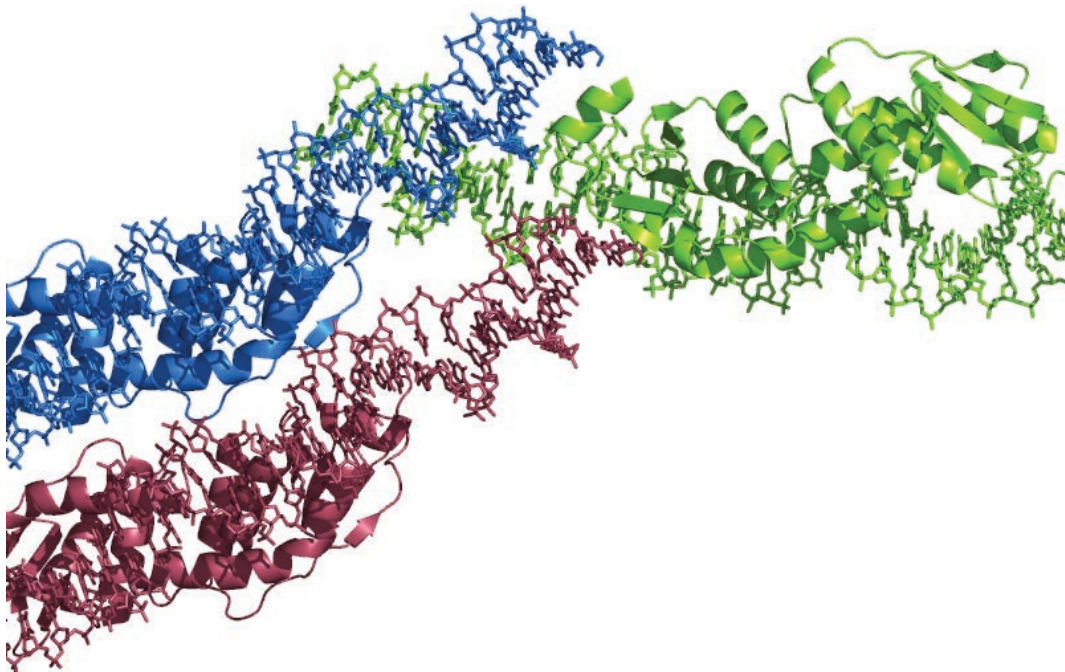


Figura 4.29: Representació de l'empaquetament en el cristall del complex p4-DNA. Un dels extrems de la molècula de DNA contacta amb una molècula de proteïna veïna.

Justament el contacte entre el DNA i la proteïna en l'empaquetament del cristall es realitzava per una zona inicialment no traçada de la proteïna en la seva regió carboxi-terminal. El traçat d'aquesta regió va completar el model de proteïna que contenia des del residu Pro2 fins al residu Met124 dels 125 que formen la proteïna p4. Aquest traçat carboxi-terminal del model de proteïna només es va donar en un dels dos monòmers de la unitat asimètrica.

Després de successius passos iteratius de minimització, inspecció gràfica del model i mapes, i retoc manual de l'estructura, es van establir els factors R a 27,2 % i el Rfree a 33,2 %. Llavors vàrem començar l'afinament dels factors de temperatura. Aquest afinament va suposar una gran devallada en els factors R fins a 24,7 % i 30,9 %. Es va fer una cerca de molècules d'aigua i es van afegir un total de 86 molècules que complien el criteri que havíem establert per validar-les, que consistia en que fes una interacció de tipus pont d'hidrogen. El procés d'afinament va concloure amb unes estadístiques finals detallades en la Taula 4.15.

Complex p4+DNA	
<i>Processat</i>	
Grup espacial	p2
Cel·la	a=52.4 , b=27.8, c=192.8, β= 97.4
Longitud d'ona (nm)	0.978
Reflexions mesurades	63103
Reflexions úniques	15759
Rang de resolució (Å)	30.0 – 2.7
Completeness (%)	99.0 (96.8)
Multiplicitat	4 (2.97)
Rmerge (%)	11.2 (75.4)
<I>/<σ(I)>	8.96 (2.52)
<i>Afinament</i>	
Rang de resolució	30.0 – 2.7
Reflexions	15900
R _{treball}	22.89
R _{free}	27.91
Àtoms totals	3766
Àtoms de proteïna	2005
Àtoms de DNA	1675
Àtoms de solvent	86
<i>R.m.s.d. dels valors objectiu</i>	
Enllaços (Å)	0.007
Angles (°)	1.26

Taula 4.15: estadístics de processat i afinament del model de la proteïna p4 en la seva forma unida a DNA

4.5.3. Descripció de l'estructura

La forma complexada de p4 amb DNA de 41 parells de bases, presenta una estructura secundària que és gairebé igual que la forma lliure. Per tant trobarem una fulla β central de 5 cadenes antiparal·leles que deixa, a una banda 3 llargues hèlix α (α1-α3) i l'altra banda un *loop* situat a l'extrem N-terminal, un gir d'hèlix 3-10 i

una petita hèlix α (α_4) formada pels residus 117-124, que no es veia en l'estructura sense unir (Figura 4.30).

En la proteïna unida a DNA, en la qual ha estat possible traçar tots els residus a excepció del primer residu C- i N-terminal, podem definir aquest plegament com un plegament de tipus α/β *antiparal·lel*, en el que una fulla β antiparal·lela central es troba amb les dues cares enfrontades a hèlixes α .

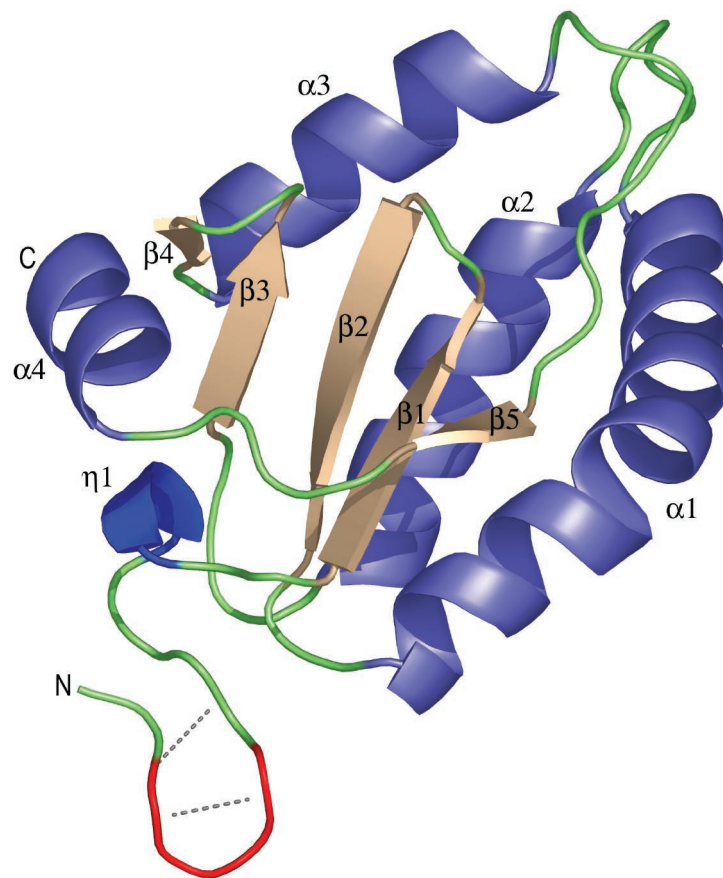


Figura 4.30: Representació de l'estructura secundària corresponent a un monòmer de p4 tal i com la trobem en el complex amb DNA

El fragment de DNA es troba en una conformació de tipus B-DNA i presenta una curvatura de 70° en el seu eix longitudinal, creiem que provocat per la unió de la proteïna p4.

4.5.4. El dímer

El grup espacial en el que va cristal·litzar el complex de p4 unida a una molècula de DNA de 41 parells de bases és el grup P2. En la unitat asimètrica trobem només un complex, format per un dímer de p4 i dues cadenes de DNA de 41 bases cadascuna.

Si fem una superposició d'un monòmer de p4 de la forma unida amb un monòmer qualsevol de la forma no unida, el valor de r.m.s.d. que obtenim és pròxim a 0.4 Å. Aquesta diferència és equiparable a l'obtinguda en la comparació entre els monòmers de p4 de la forma lliure, el que ens porta a afirmar que no hi ha diferències substancials entre les estructures.

Tal i com vam observar en els diferents monòmers de proteïna no unida, les regions que difereixen més les trobem en l'extrem C-terminal del monòmer (Figura 4.31), on a més, en la forma unida a DNA veiem els últims residus de la molècula, que formen una hèlix α ($\alpha 4$).

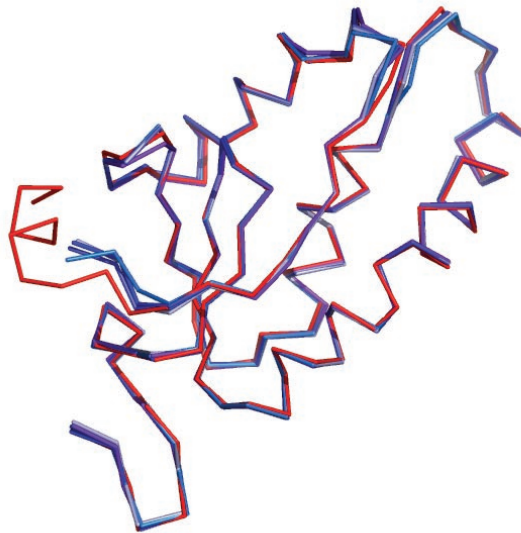


Figura 4.31: Superposició dels monòmers lliures (diferents tons de blaus) i d'un monòmer del complex amb DNA (en vermell). Observem que l'única diferència la trobem en l'extrem C-terminal on, en la proteïna del complex s'ha pogut traçar 8 residus més.

4.5.5. L'hèlix $\alpha 4$

El monòmer A de p4 del cristall del complex mostra densitat electrònica en l'extrem C-terminal, el que ens va permetre de traçar el model fins al penúltim residu, el Met124. Aquesta regió que comprèn els residus Leu117-Met124 s'estructura en forma d'hèlix α . El fet d'observar densitat en aquesta regió no creiem que obeeix a un efecte estructurador de la unió funcional de la proteïna al DNA sinó que es deu a un efecte de l'empaquetament del cristall (Figura 4.32).

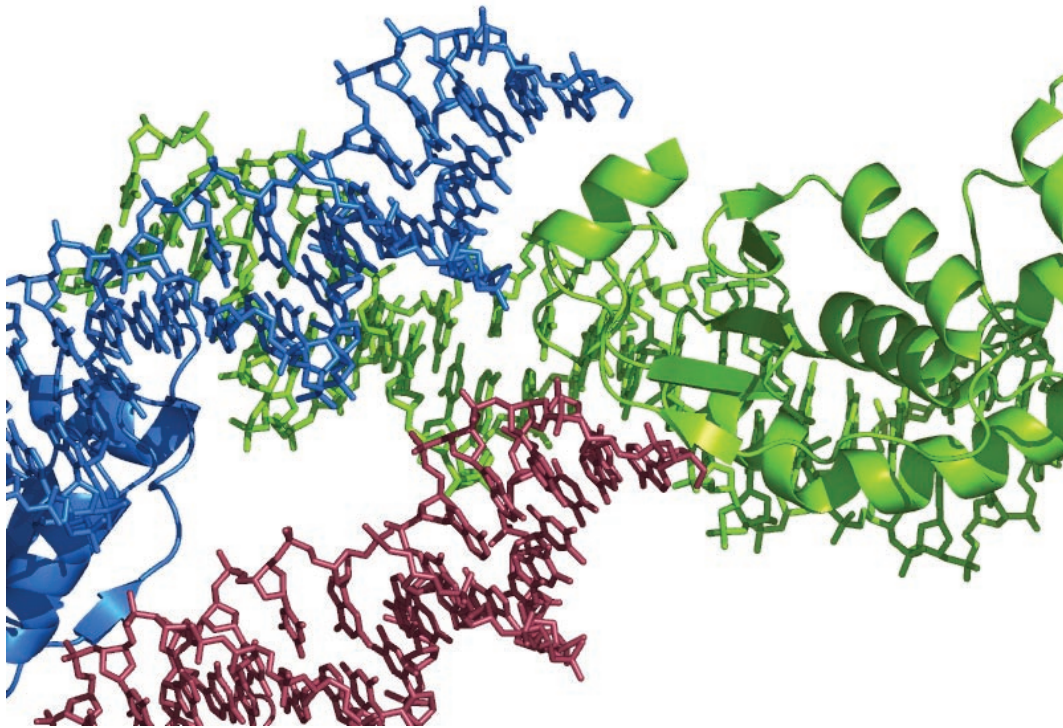


Figura 4.32: Representació en forma de *sticks* (DNA) i cintes (proteïna) per observar l'hèlix $\alpha 4$ atrapada entre dues molècules de DNA veïnes, restringint la seva mobilitat.

Si observem l'altre monòmer que forma el complex, el monòmer B, veiem que la densitat electrònica finalitza en el residu Lys114, tal i com ens trobàvem en els sis monòmers presents en els cristalls de proteïna no unida.

4.5.6. Unió al DNA

La unió entre p4 i el DNA es fa al llarg de l'eix longitudinal de les dues molècules implicades, contactant per una sola cara i produint-se, en la molècula de DNA, una torsió de 70° (valor obtingut utilitzant el programa CURVES (Lavery *et al*, 1988)). En aquesta unió es poden distingir dues zones, en la primera està implicada l'extrem amino terminal de la proteïna, formant el *loop* de reconeixement, i una segona zona on les interaccions es realitzen en la regió de dimerització de la proteïna (Figura 4.33).

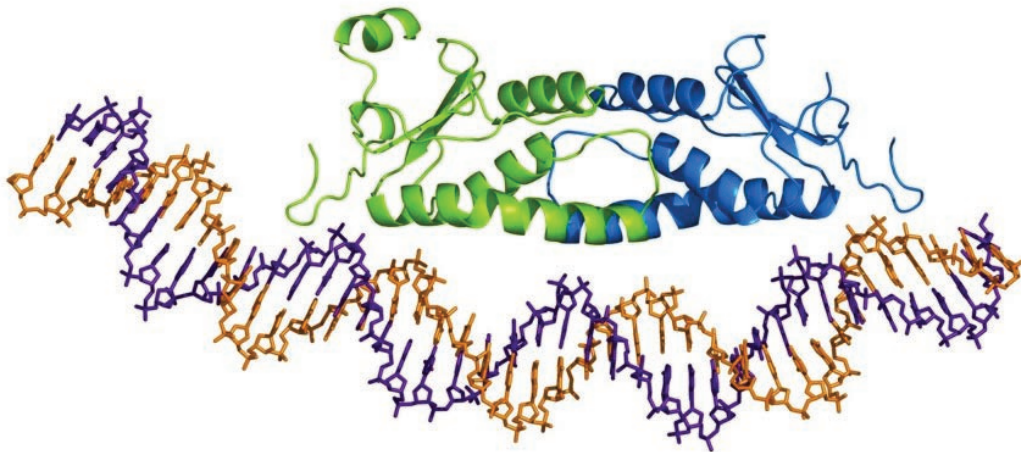


Figura 4.33: Esquema general del complex format per un dímer de proteïna p4 (representat en cintes) i una molècula de DNA de 41 parells de bases (representat en varetes). Es pot observar una gran torsió (70°) en la molècula de DNA.

4.5.7. Zona del *loop* de reconeixement

El *loop* de reconeixement està format per 5 residus (de Thr4 a Ile8) situats en l'extrem N-terminal de la molècula de p4. Com ja hem vist anteriorment en la Figura 4.19, aquesta regió, malgrat la seva aparença flexible, és rígida degut a una gran quantitat d'interaccions que estableixen aquesta conformació. Conformació que també s'observa en la proteïna lliure, no unida al DNA.

Aquesta estructura s'inserta en el solc major del DNA, i dos residus, Gln5 i Arg6, estableixen interaccions de tipus pont d'hidrogen amb les bases del DNA. Els contactes comentats a continuació fan referència al monòmer A del dímer de p4. Els petits canvis en el monòmer B es comentaran oportunament.

El residu Gln5 estableix un pont d'hidrogen entre el seu àtom N^E2 i l'àtom O4 de la base nitrogenada del nucleòtid T32' (Figura 4.34).

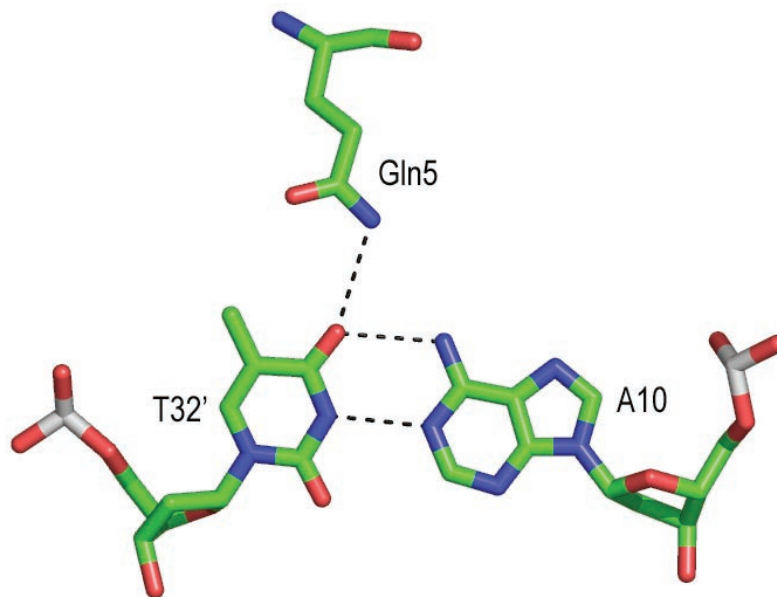


Figura 4.34: Detall de la interacció entre l'àtom NE2 del residu Gln5 i l'àtom O4 de la base nitrogenada del nucleòtid T32'

El residu Arg6 interacciona amb la base G30'. Els àtoms específics implicats en la interacció són Arg6:NH2 amb G30':O6 i Arg6:NE amb G30':N7 (Figura 4.35).

Aquest doble reconeixement de les bases del DNA constitueixen el reconeixement específic de la seqüència de DNA per part de la proteïna. En la regió del *loop* de reconeixement trobem altres interaccions amb el DNA de manera inespecífica o indirecta, contactant amb elements del DNA independents de la seqüència (Figura 4.36).

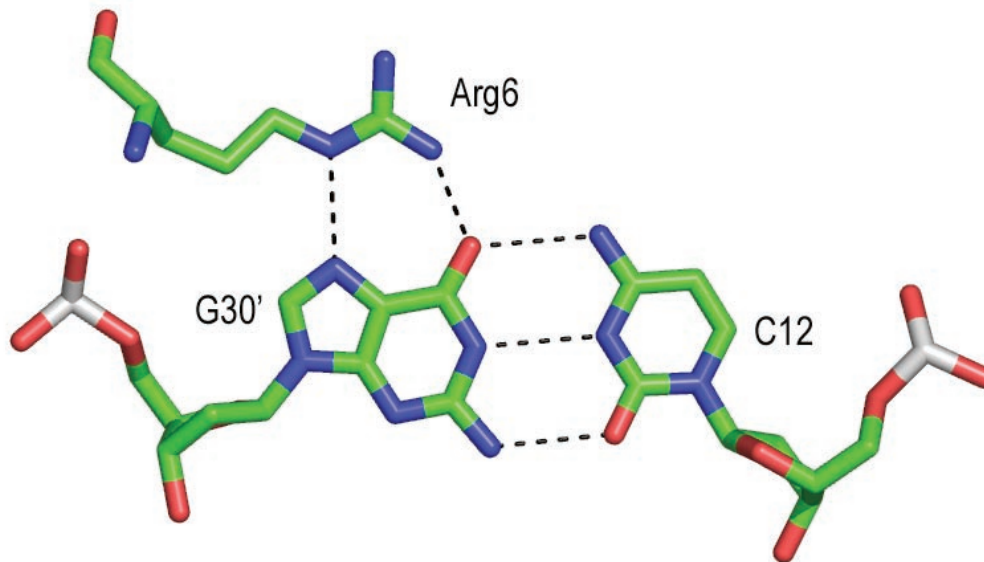


Figura 4.35: Detall de la interacció entre el residu Arg6, mitjançant la seva cadena lateral i la base nitrogenada del nucleòtid G30'.

El primer residu que trobem contactant amb el DNA és Thr4. Aquest residu, implicat també en el manteniment de la conformació del *loop* de reconeixement, estableix un pont d'hidrogen amb un àtom d'oxigen del grup fosfat del residu G30'. Aquest grup fosfat també interacciona amb el residu His10 però a través d'un altre àtom d'oxigen. Hi ha dos residus més que contacten amb els fosfats del DNA. Aquests són el residu Tyr33 i el residu Lys36, que contacten els oxígens dels grups fosfats dels nucleòtids T17 i A18 respectivament.

Aquests grups fosfats queden fortament fixats, com ho demostra els valors més petits dels factors de temperatura (Figura 4.37).

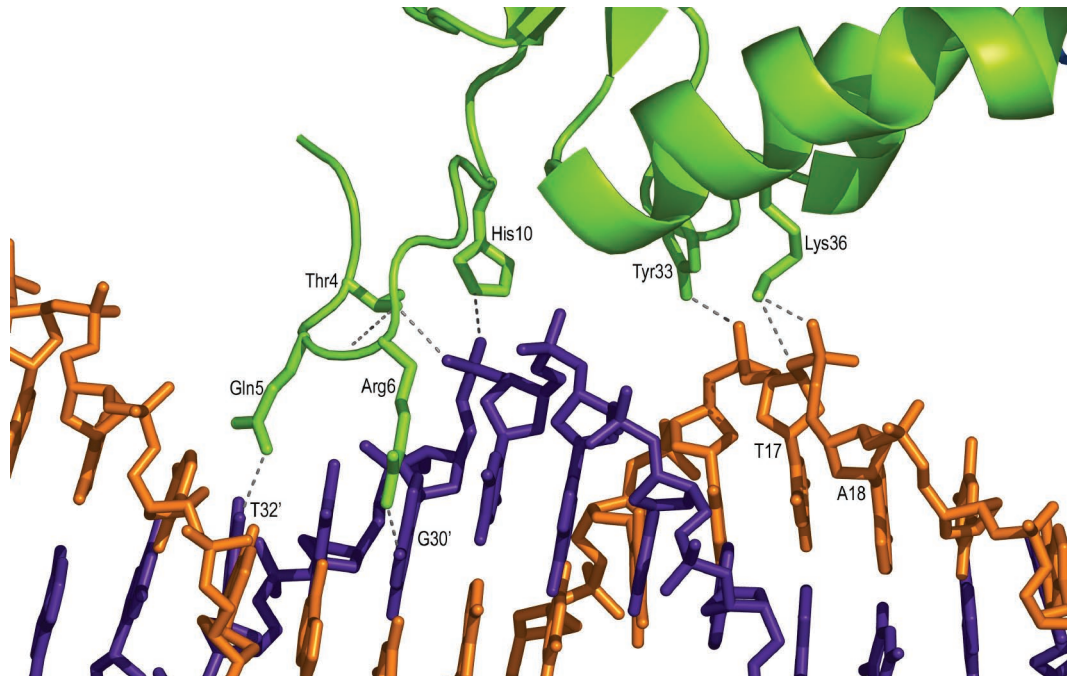


Figura 4.36: Esquema general de les interaccions específiques de seqüència que realitzen els residus del *loop* de reconeixement i les no específiques de seqüència, que es les realitzen varis residus mitjançant contactes amb els fosfats.

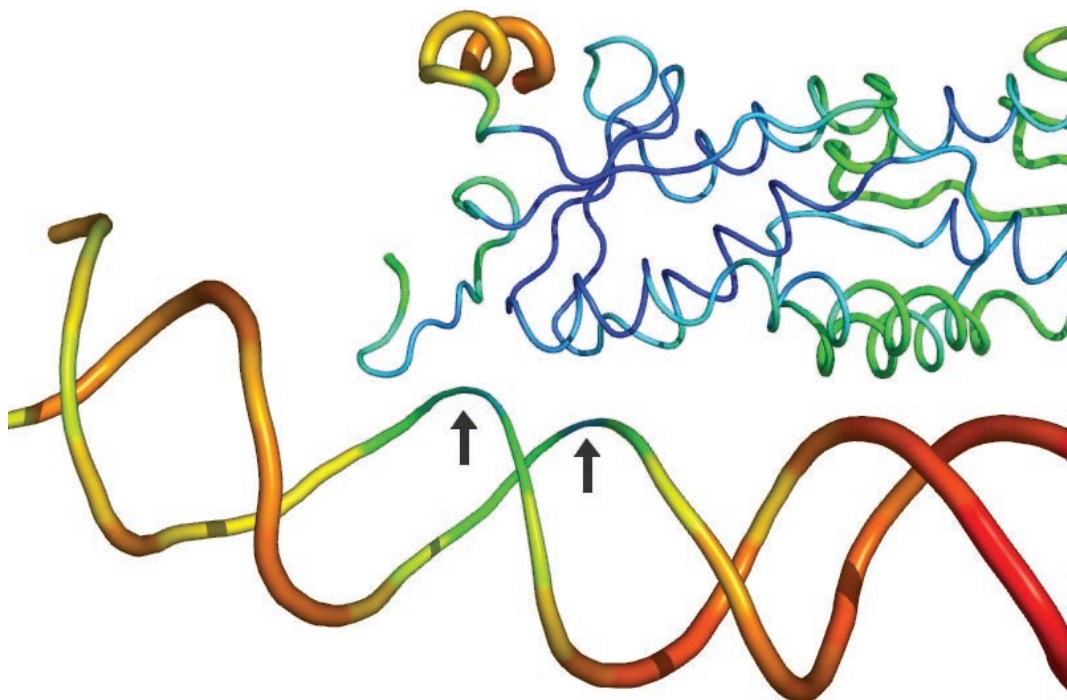


Figura 4.37: Representació en forma de tub dels factors de temperatura. Com més ample i vermell, més alt és el valor del factor de temperatura. Les fletxes senyalen els punts de contacte amb la proteïna, observant-se una reducció dràstica dels factors de temperatura.

Si ens fixem en els grups fosfat que interaccionen amb la proteïna, aquests són els que defineixen el solc menor en un punt determinat de l'oligonucleòtid.

4.5.7.1. Canvis en la interacció del monòmer B

En el monòmer B hi ha dos petits canvis. El primer implica al residu Gln5, on la densitat electrònica observada és molt pobra en la cadena lateral. Això ens fa dubtar a l'hora de decidir si en el monòmer B, el residu Gln5 realitza la mateixa interacció que en el monòmer A. Una possible causa pot ser l'asimetria que presenta el complex. La distància que existeix entre la interacció del residu Gln5 i l'extrem 3' lliure és diferent en els dos extrems, concretament 9 parells de bases en el monòmer A i 1 sol parell de bases en el monòmer B i, a més queden empaquetats de maneres molt diferents.

L'extrem més proper al monòmer A, com hem explicat anteriorment, s'empaquetaria contra la regió que conté $\alpha 4$ d'un monòmer A simètric, mentre que l'extrem proper al monòmer B contactaria, en l'eix binari cristal·logràfic, amb una molècula de DNA. La pobresa observada en la densitat electrònica pot ser deguda a la interacció del residu Gln5 del monòmer B amb una molècula veïna de DNA (Figura 4.38). Així el residu Gln5 contactaria amb 2 nucleòtids alternativament (T40 i T1' d'un complex veí), i per tant dues conformacions, implicant que les ocupàncies dels àtoms de la cadena lateral del residu Gln5 es trobessin disminuïdes, reduint la densitat electrònica en cadascuna.

Una altra causa podria ser una inestabilitat en el contacte entre les molècules de DNA en l'eix binari de la cel·la unitat (Figura 4.39). El contacte no és l'habitual en el DNA, ja que és de tipus 5'-5' i 3'-3', i els contactes entre les bases es veuen alterats.

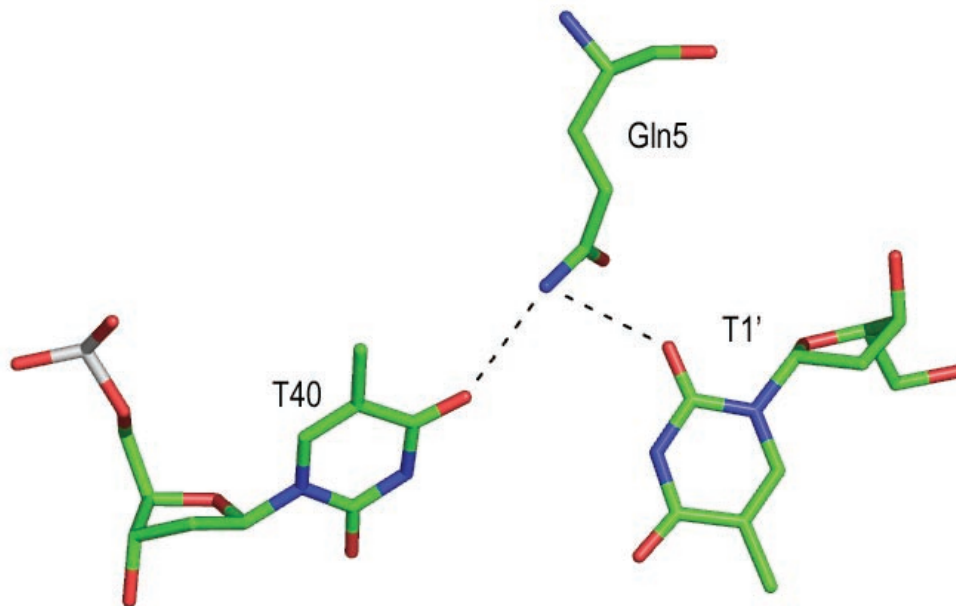


Figura 4.38: Possibles contactes del residu Gln5 del monòmer B, mostrant-se una conformació promitjada que contactaria amb els dos nucleòtids alhora.

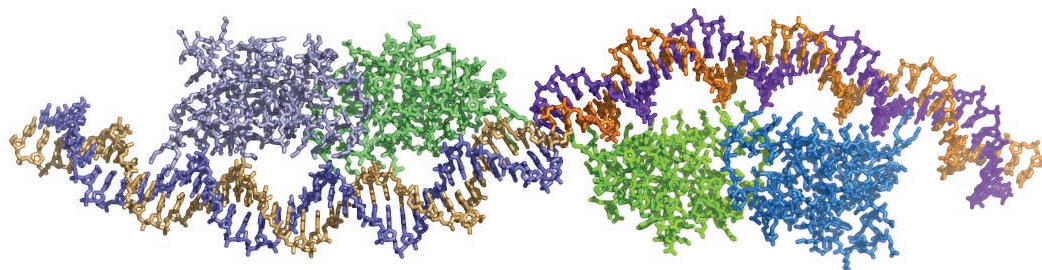


Figura 4.39: Esquema de les molècules de proteïna i DNA contingudes en una cel·la unitat. Al centre hi ha l'eix binari que relaciona els dos complexos

Un altre canvi present en la interacció del monòmer B és el contacte que realitza el residu Lys76 contactant amb un fosfat del DNA (Figura 4.40).

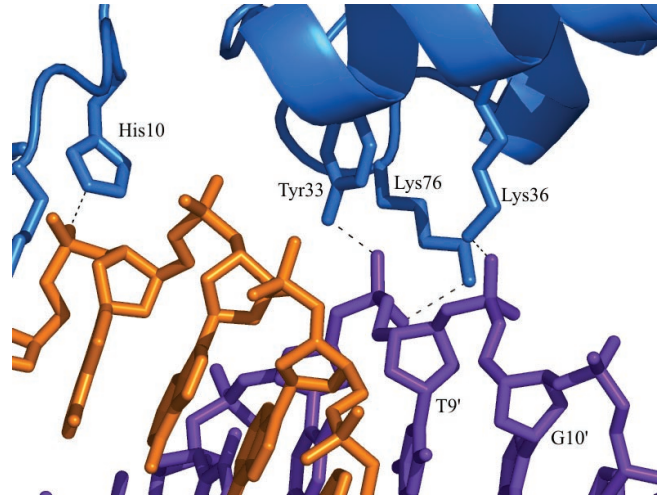


Figura 4.40: Detall de la interacció del residu Lys76 del monòmer B

El residu Lys76 contacta amb el grup fosfat del nucleòtid T9', a l'igual que el residu Tyr33, mentre que el residu Lys36 contacta amb el grup fosfat del nucleòtid G10'. Aquestes interaccions, a excepció del residu Lys76, són exactament iguals que les observades en el monòmer A. En el monòmer A, el residu Lys76 podria establir també aquesta interacció, però el trobem en una conformació replegada cap a la proteïna, interaccionant amb el grup carbonil del residu Glu73 (Figura 4.41).

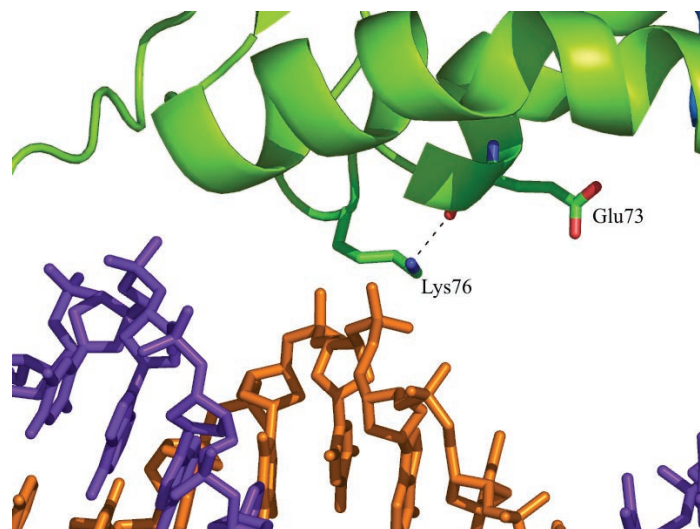


Figura 4.41: Detall de la interacció del residu Lys76 del monòmer A amb el grup carbonil del residu Glu73

4.5.8. Zona de dimerització

La segona regió de contacte entre la proteïna i el DNA es troba en la zona de dimerització de la proteïna. Les interaccions que hi trobem són de tipus inespecífic, ja que els contactes amb el DNA es realitzen, novament, amb els grups fosfat (Figura 4.42).

Troblem dos residus implicats en les interaccions. Aquests residus són Lys51 i Arg54. En trobar-se a la regió de dimerització, els dos residus esmentats de cada monòmer interaccionen alhora.

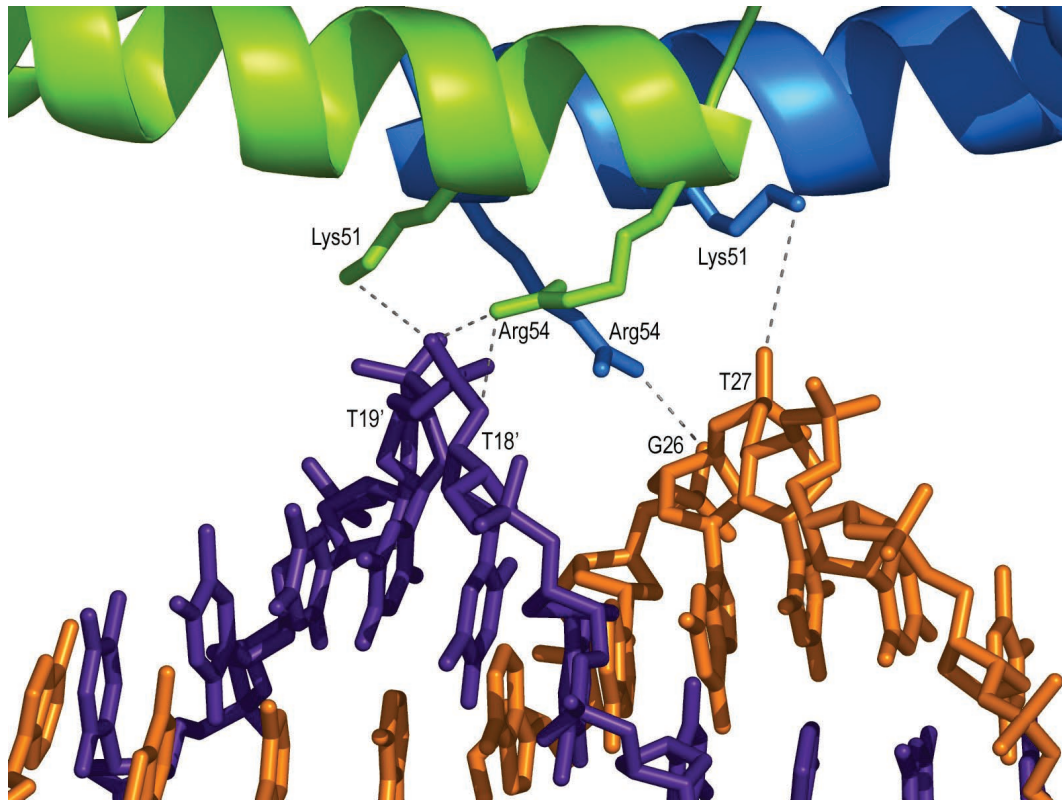


Figura 4.42: Esquema general de la interacció de la proteïna p4 amb el DNA en la zona de dimerització de la proteïna

El residu Lys51 interacciona mitjançant el seu àtom NZ amb un àtom d'oxigen del grup fosfat del nucleòtid T19', en el cas del monòmer A, o del nucleòtid T27,

en el cas del B.

El residu Arg54 actua de manera similar al Lys51 i contacta amb els grups fosfat dels nucleòtids T19' i G26, respectivament pels monòmers A i B.

De la mateixa manera que en l'anterior regió d'interacció proteïna-DNA, vàrem observar els factors de temperatura dels fosfats contactats per la proteïna. En aquest cas no es veu un descens en els factors de temperatura, probablement per una unió més dèbil que en l'altre cas (Figura 4.43).

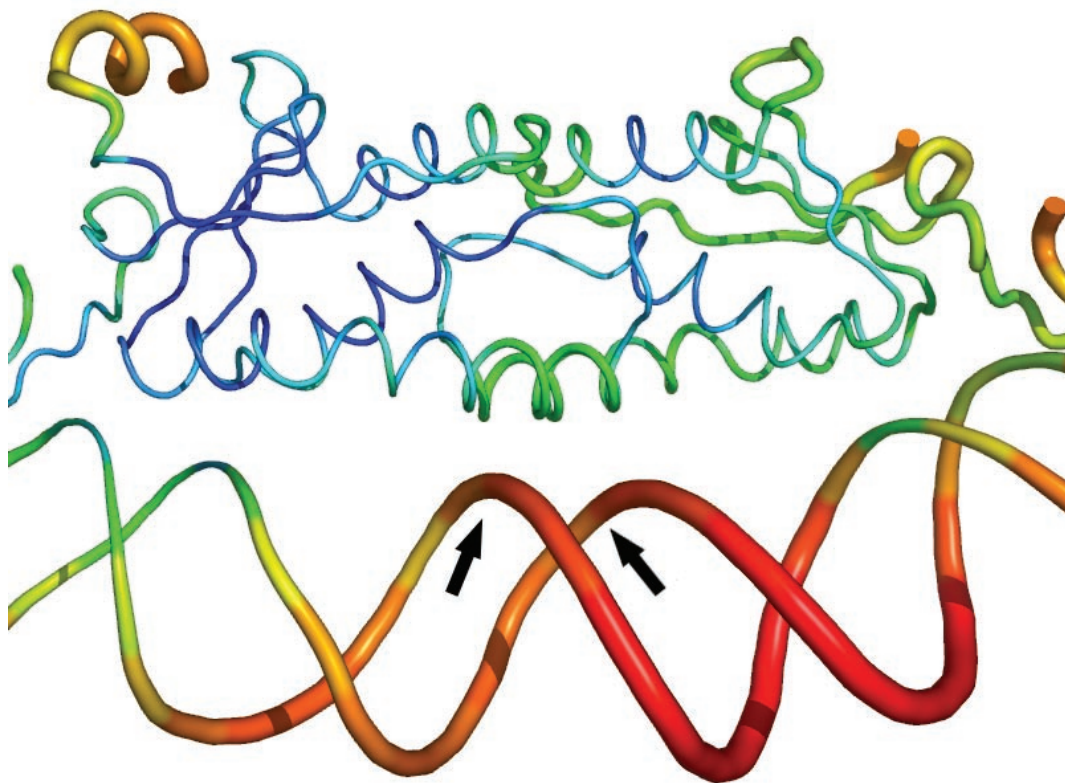


Figura 4.43: Representació en forma de tub dels factors de temperatura. Les fletxes assenyalen els punts de contacte amb la proteïna. En aquest cas no s'observa una diferència apreciable entre les zones que contacten amb la proteïna i les que no ho fan.

4.5.9. Validació de l'estructura: el diagrama de Ramachandran

L'anàlisi del diagrama de Ramachandran (Ramachandran, 1968) ens presenta un resultat semblant a l'observat en la proteïna nativa (Figura 4.44).

Veiem novament que hi ha un residu, el Thr22, que es troba fora de les zones favorables del Ramachandran, per raons que ja s'han explicat en el punt 4.3.10.

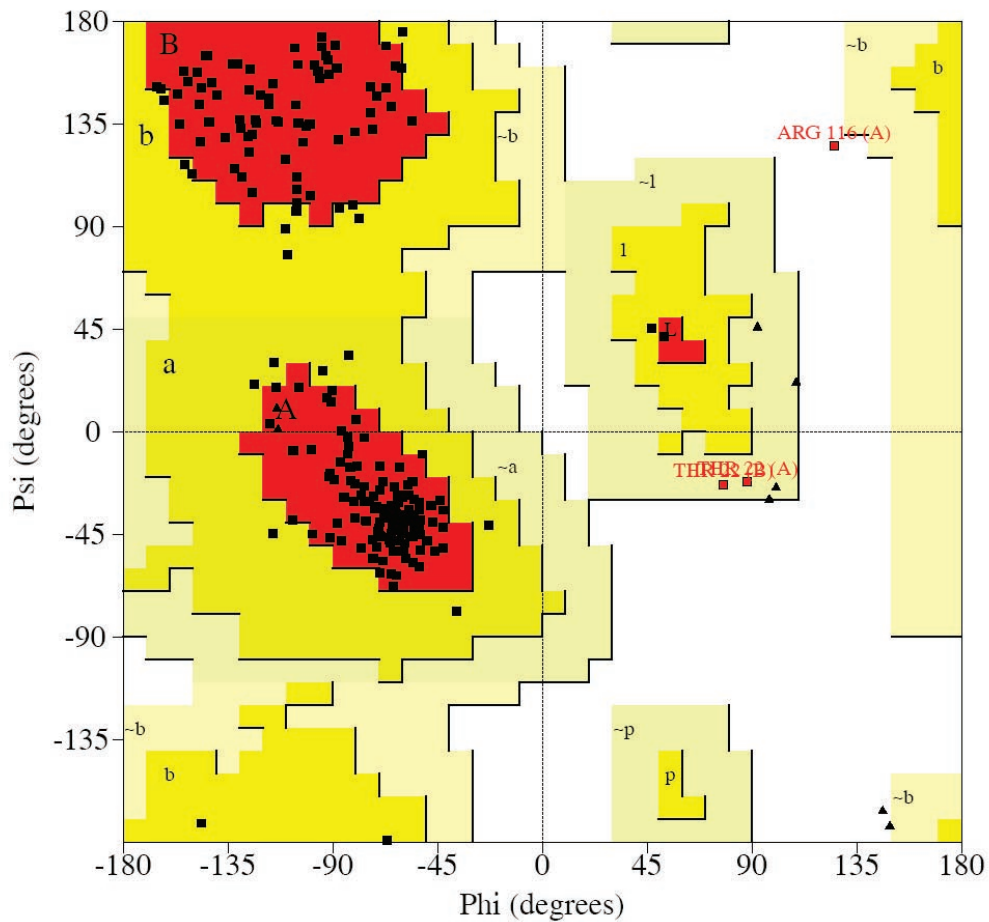


Figura 4.44: Diagrama de Ramachandran de la proteïna p4 en complex amb DNA. Com en el cas de la forma no unida de la proteïna p4, trobem el residu Thr22 fora de les zones favorables, així com el residu Arg116

També veiem que queda fora el residu Arg116 del monòmer A. El fet que aquest residu es trobi en una conformació d'alta energia pot venir de la situació forçada de l'hèlix α_4 , que es troba atrapada en aquesta conformació per efecte de l'empaquetament enmig de dues molècules de DNA.

Els altres residus es troben situats en les zones favorables del diagrama de Ramachandran.

4.5.10. Estructura del DNA

El DNA pot presentar diferents conformacions (Figura 4.45) depenent, essencialment, del sentit de gir de l'hèlix –dextrògira en la majoria dels casos– i de les mides dels solcs majors i menors. La conformació majoritària és la B; A i Z són molt menys nombroses i existeixen algunes altres conformacions molt minoritàries.

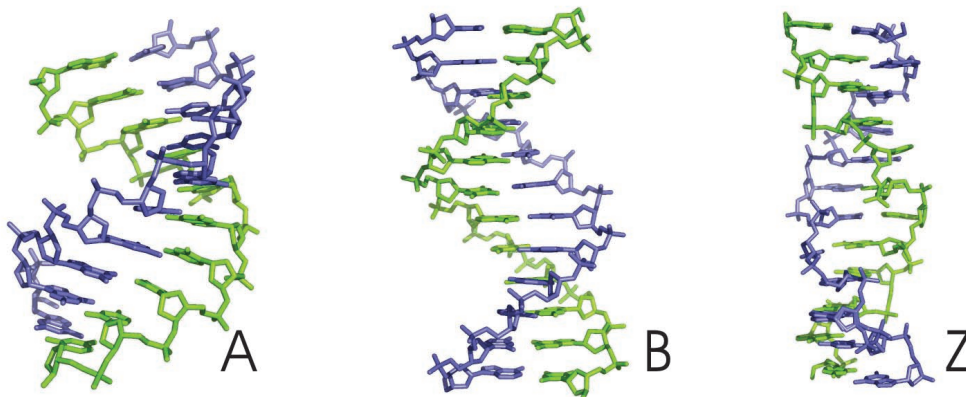


Figura 4.45: Esquema de les conformacions majoritàries del DNA

La conformació B, presenta una hèlix dextrògira i una clara diferenciació en els seus solcs majors i menors. A la conformació A, al igual que a la B, l'hèlix gira cap a la dreta, però aquesta és més ampla i curta, quedant el solc major més estret i profund i el menor més ample i superficial. La conformació Z és la més distant morfològicament, ja que la seva hèlix gira cap a l'esquerra amb un solc major pla i un solc menor estret i profund.

El fragment de DNA que forma el complex presenta una conformació de B-DNA. Aquesta seqüència, rica en parells A-T, té una certa tendència intrínseca a corbar-se, fins a 50° (Rojo, 1990). La unió del dímer de p4 provoca un augment en la curvatura fins als 70° observats en l'estructura cristal·lina. El càlcul de la curvatura es va realitzar amb el programa CURVES (Lavery *et al.*, 1988).

Aquesta curvatura és estabilitzada per la interacció dels grups fosfats amb la proteïna p4 i ve acompanyada d'una reducció de l'amplada del solc menor (9 Å quan el valor ideal de la conformació B de 11.5 Å; Figura 4.46). Aquesta circumstància es dona en els tres punts on el solc menor contacta amb la proteïna, actuant aquesta a mode de pinça.

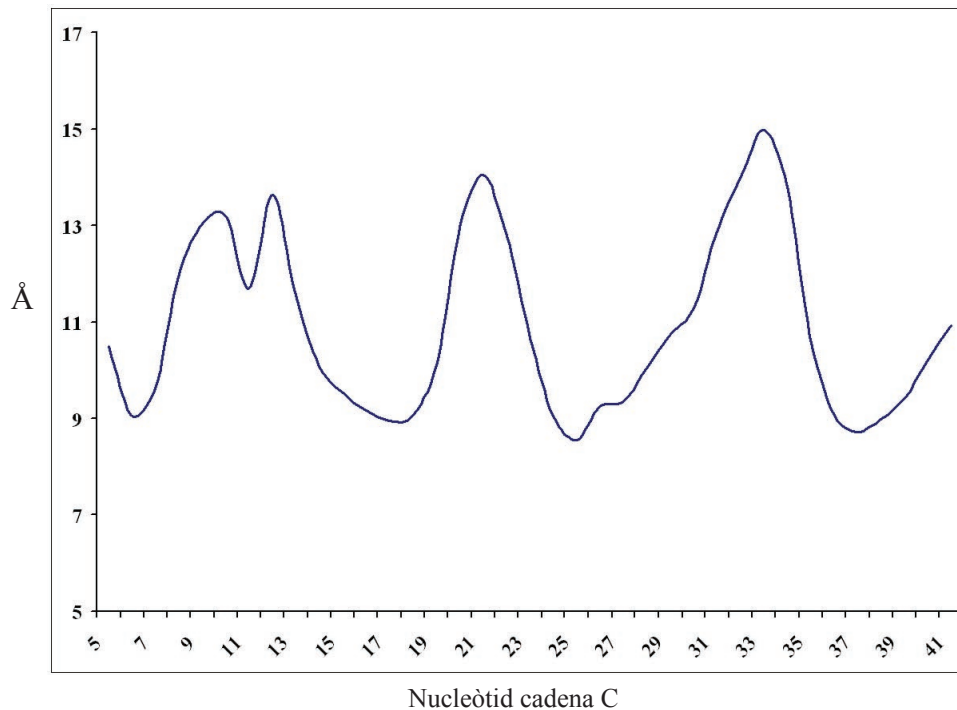


Figura 4.46: Gràfica de l'amplada del solc menor al llarg de tota la molècula de DNA. L'amplada varia molt, havent-hi 3 mínims que corresponen amb les zones de contacte entre la proteïna i els grups fosfat del DNA.

Per altra banda, el solc menor, en les zones oposades al contacte amb la proteïna, augmenta la seva amplada, respecte dels valors ideals (15 Å). En el solc major no s'observa cap correspondència clara entre l'amplada i la curvatura.

Com en altres complexos, el DNA es corba cap a la proteïna augmentant la superfície de contacte i possibilitant interaccions amb la mateixa.

4.5.11. Canvis en la proteïna a l'unir-se al DNA

Els canvis que pateix la proteïna a l'unir-se al DNA són mínims. Superposant els dos models corresponents a les formes unides i no unides, veiem que l'estructura bàsica es conserva clarament, obtenint un valor de r.m.s.d. pels àtoms de carboni- α de 0.665 Å, un valor lleugerament superior al que hi havia entre els 6 monòmers de la unitat asimètrica del cristall de la forma no unida de p4 (punt 4.3.10).

El major canvi en l'estructura secundària el trobem, tal i com hem explicat anteriorment, en l'extrem C-terminal del monòmer A on, per efecte de l'empaquetament del cristall, hi ha una estabilització d'aquesta regió i això ens permet observar la seva densitat electrònica.

Si observem els canvis en l'interfície proteïna-proteïna veiem que hi ha un lleuger canvi, augmentant la superfície de contacte de 808 Å² a 820 Å². Aquest canvi és degut a una petita rotació d'una molècula respecte de l'altra, per acomodar-se al seu substrat, el DNA (Figura 4.47).

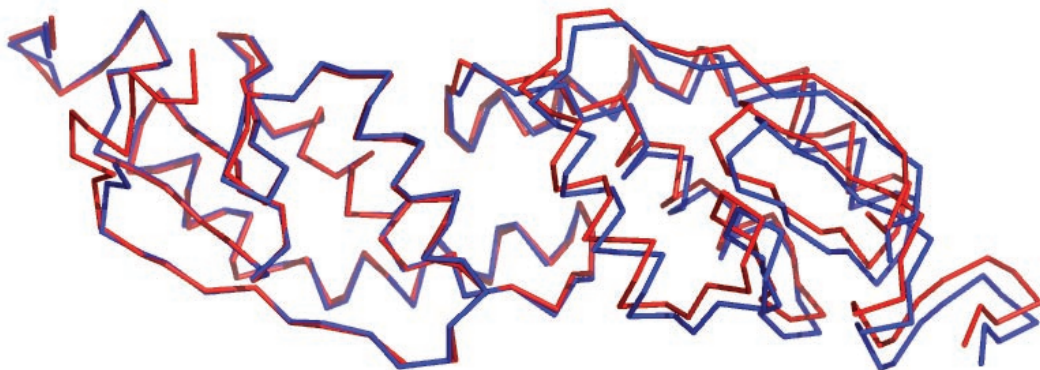


Figura 4.47: Superposició d'un dímer no unit a DNA (blau) i el dímer unit a DNA (vermell). Hi ha una lleugera rotació en la zona de dimerització per tal d'adaptar-se millor al DNA