



Departament d'Enginyeria Química



UNIVERSITAT POLITÈCNICA DE CATALUNYA

TESI DOCTORAL

**Estudis estructurals de complexos
de les proteïnes HMG amb DNA**

Raquel Sánchez-Giraldo

Tesi presentada per obtenir el títol de Doctora per la Universitat Politècnica de Catalunya

Directores: Núria Saperas Plana i J. Lourdes Campos López

Programa de doctorat: Polímers i Biopolímers

Barcelona, Setembre 2014

Al meu pare i al Marc

Freude ist die Essenz des Erfolgs

La satisfacció és l'essència de l'èxit

(Yogi Tee)

Agraïments

En primer lloc voldria donar les gràcies a les meves dues directores, la Dra. Núria Saperas i la Dra. Lourdes Campos, per deixar-me formar part uns anys més d'aquesta petita família científica que és el grup MACROM i donar-me l'oportunitat de formar-me com a investigadora realitzant la tesi doctoral en ell.

Li agraeixo especialment a la Núria la seva gran ajuda durant el desenvolupament d'aquest treball, el temps dedicat en la supervisió d'aquesta memòria i transmetre'm la seva motivació tant per la investigació com per la docència i sobretot pel treball ben fet. A la Lourdes voldria agrair-li els seus consells i ajuda incondicional en la part de cristal·lografia, difracció de raigs X i resolució de l'estructura, donar-li les gràcies per la seva amabilitat i disponibilitat tan a nivell laboral com personal i finalment per contribuir activament a què deixés de fumar. També agraeixo a les dues haver-me donat ànims durant la continua i esgotadora lluita per l'obtenció de resultats.

Voldria mostrar el meu sincer agraïment al Dr. Joan Antoni Subirana per les seves aportacions a aquest treball i per ser l'impulsor de les múltiples, enriquidores i fructíferes col·laboracions d'aquesta tesi. És una sort poder aprendre d'ell sobre el camp de la cristal·lografia.

I wish to express my gratitude to Dr. Mair Churchill for giving me the opportunity to be part of her group at the University of Colorado and work with HMGBs. I would also like to thank her for her valuable suggestions, guidance and encouragement during my period there and the preparation of the paper. This work wouldn't have been possible without her help. I would like to acknowledge the members of Churchill's lab: Chrissy, Sarah, Prince, Kim, Wallace, Luke and Jean for their support and advice, special thanks to Chris for his teachings and help with my experiments. I am also grateful to all the people that contributed to make my experience in Colorado unforgettable and make me feel like at home, most especially to Prince, Judy, Suraj and CP. Finalmente, me gustaría darle las gracias a Graciela por haberme acogido en su casa, haberme ayudado en todo lo posible y más, por los hikes a las montañas y por su amistad.

També agrair l'ajuda rebuda pel Dr. Joan Pous i per la Sònia Rodríguez a la Plataforma Cristal·logràfica del Parc Científic de Barcelona i al Dr. Oscar Domènech del Grup de Biofísica de Biomembranes de la UB. M'agradaria donar especialment les gràcies al Dr. Joaquim Roca del IBMB-CSIC per la seva grandíssima ajuda per a portar a terme els assajos sobre la topologia del DNA, la seva amabilitat i l'haver-me ajudat a endinsar-me en aquest camp tan interessant.

M'agradaria agrair molt sincerament a tots els meus companys i excompanys de laboratori la seva amistat, l'haver fet tan i tan amè el temps compartit i l'haver estat sempre units: a la Mireia per les rialles al laboratori i els bons moments en les barbacoes i demés, a la Marta per les experiències compartides tant fora com dins del laboratori, a l'Elsa per tots els anys viscuts juntes al lab superant els moments dolents i gaudint dels millors i pels bons records acumulats fora d'ell, al Paco pels moments plegats i la seva ajuda inigualable en la resolució d'estructures, a la Cinthia pel seu suport i per aportar sempre alegria i positivitat, al David per les converses transcendents, pels seus ànims en l'obtenció de cristalls i pel vídeo, al Carles (membre adoptiu) per la seva espontaneïtat, per la seva amabilitat i les bones estones passades. Finalment voldria donar mil gràcies a la Mari, per ser la millor de les companyies tant al laboratori com fora, estar sempre disponible per ajudar i animar-me, per la seva generositat, la seva alegria i el seu punt de bogeria. El temps compartit juntes ha estat immillorable i desitjo gaudir de la teva amistat sempre.

Voldria fer extensiu el meu agraïment als membres i exmembres del Departament d'Enginyeria Química (Dra. Lourdes Urpí, Dra. Lourdes Franco, Dr. Alfonso Rodríguez, Dra. M^a Teresa Casas, Dra. M^a Pilar Almajano, Dr. David Zanuy, Elena, Sara, Yolanda, Gustavo i tants altres) per totes les estones compartides a l'hora de dinar, els debats sobre l'ordre del dia, les converses al passadís i els molts moments de desconexió i riures.

També agrair-li al Hussein la seva paciència, amabilitat i simpatia en el suport informàtic aportat durant la realització d'aquest treball, a la Carol Ramos l'ajuda aportada en tantes ocasions i a la Irene Pérez pel suport i indispensable ajuda en la realització de gestions i tràmits.

No em voldria descuidar de tots els amics que tinc la sort de tenir (amics de la universitat, amigues de l'institut i diferents persones que he anat coneixent al llarg de la meva vida). Gràcies pel vostre suport, pels moments inoblidables viscuts plegats i per la vostra amistat. Especialment voldria agrair l'amistat de l'Alba, que més que una amiga és una germana, per ajudar-me tant a poder superar els mals moments, per estar al meu costat sempre que ho he necessitat, per conèixer-me massa i per la infinitat de bons moments juntes.

També voldria donar les gràcies als diferents membres de la meva família (tietes, cosins, cosines i els seus nens) així com a la mare i al germà del Juli pels moments de desconexió en les estones compartides. Especialment a la meva mare i als meus germans per animar-me a acabar aquesta tesis, per donar-me força en els mals moments, pel seu suport incondicional, per romandre sempre units i per estar-hi sempre. També a les meves nebodetes Júlia i Jana per l'energia positiva que em transmeten quan estic amb elles.

Finalment voldria mostrar el meu agraïment de forma especial al Juli per haver confiat tant en mi i haver-me donat sempre ànims durant la realització d'aquesta tesi. Moltes gràcies per aguantar-me en els moments de desànim i per saber com fer-me agafar forces. Gràcies per fer que la vida al teu costat sigui millor i més fàcil. Estic molt contenta d'haver iniciat aquesta nova etapa al pis amb tu i desitjo tenir-te sempre al meu costat.

Aquesta tesi s'ha realitzat gràcies al suport del CUR del DIUE de la Generalitat de Catalunya i del Fons Social Europeu (beca FI-DGR2011) i també gràcies a la beca FPI-UPC (que es va gaudir durant l'inici de la tesi).

Abstract

High Mobility Group proteins (HMG) are architectural factors involved in modulating chromatin structure and influencing a myriad of cellular processes such as transcription, replication and DNA repair. Altered levels of expression and mislocalization of these proteins have several pathological outcomes such as cancer and inflammatory processes.

Understanding how the different HMG proteins recognize DNA and which changes can induce in it is still a challenge, and will be of great importance to interpret their activities *in vivo*. For this purpose, the aim of this work is to obtain the atomic structure of HMG proteins in complex with DNA using X-ray crystallography.

In this work, we have studied two HMG families: HMGA and HMGB. Both bind to the minor groove of the DNA but are distinguished by their binding motifs: AT-hook motif (HMGAA) and HMG-box motif (HMGBB). AT-hooks are short positively charged flexible segments that bind preferentially to AT rich regions whereas HMG-box domains contain ~75 amino acids and have a characteristic L-shaped fold consisting of three α -helices and bind to DNA with limited or no sequence specificity.

We have expressed and purified the HMGA1a protein, the fragment HMGA1b(1-79), the fragments Box A, Box B and Box A+B of the HMGB1 and the NHP6A (yeast HMGB). We have performed crystallographic assays of the complexes of AT rich oligonucleotides with these proteins (as well as with synthetic peptides corresponding to the AT-hook motif). We have studied the packing of some of the obtained crystals.

As a major contribution of this thesis, we report the crystal structure of the HMGB1 Box A domain bound to the oligonucleotide ATATCGATAT at a resolution of 2 Å. It demonstrates a new mode of DNA recognition for HMG-box proteins. Two Box A domains act together underwinding and kinking the DNA by 85°. In this unusual configuration the Phe37 of both domains stack together and intercalate the same CG base step. This duplex shows the greater distortion in just a base pair step that we have evidence for complexes with HMG-box domains.

Resum

Les proteïnes HMG (*High Mobility Group*) són considerades com a factors arquitectònics del DNA que modifiquen l'estructura de la cromatina afectant així diversos processos cel·lulars com la transcripció, la replicació i la reparació del DNA. L'alteració en els nivells d'expressió normals d'aquestes proteïnes i la seva deslocalització estan relacionades amb nombrosos processos patològics com ara el càncer o la inflamació.

Per a poder interpretar l'activitat de les proteïnes HMG *in vivo* és molt important entendre els mecanismes de reconeixement en la unió al DNA que presenten les diferents famílies d'HMG així com els canvis que poden provocar en aquest, fet que representa encara un repte. Amb aquesta finalitat, en aquest treball ens hem proposat obtenir l'estructura atòmica dels complexos de diverses proteïnes HMG amb DNA mitjançant cristal·lografia de raigs X.

Hem treballat amb dues famílies de proteïnes HMG: les HMGA i les HMGB. Ambdues s'uneixen al solc menor del DNA però es diferencien pels seus motius d'unió: motiu AT-hook (les HMGA) i motiu HMG-box (les HMGB). Els *AT-hooks* són cadenes curtes bàsiques i flexibles que s'uneixen preferentment a regions riques en ATs, mentre que els dominis HMG-box, compostos per uns 75 aminoàcids, presenten una forma de L constituïda per tres hèlixs alfa i una especificitat de seqüència d'unió al DNA reduïda o inexistent.

Durant el transcurs d'aquest treball s'han expressat i purificat la proteïna HMGA1a, el fragment HMGA1b(1-79), els fragments corresponents al Box A, Box B i Box A+B de l'HMGB1 i la proteïna NHP6A (HMGB de llevat) i s'han dut a terme assajos cristal·logràfics dels complexos d'oligonucleòtids rics en ATs amb aquestes proteïnes (així com amb pèptids sintètics corresponents als motius *AT-hook*). S'ha pogut estudiar l'empaquetament d'alguns dels cristalls obtinguts d'aquests complexos.

L'aportació més important d'aquesta tesi és la resolució de l'estructura tridimensional d'un complex del domini Box A de l'HMGB1 amb l'oligonucleòtid ATATCGATAT a una resolució de 2 Å. Aquesta demostra un nou mode de reconeixement de DNA lineal per part dels dominis HMG-box. En l'estructura, dos dominis Box A s'uneixen a un mateix dúplex de DNA provocant un desenrotllament i un doblament d'aquest de 85°. En aquest mode d'unió inusual, les Phe37 d'ambdós dominis s'intercalen en el mateix pas CG. Aquesta estructura presenta un dúplex amb la major distorsió en un únic pas de la que tenim constància per a un complex del domini HMG-box.

Sumari

AGRAÏMENTS	VII
ABSTRACT	XI
RESUM	XIII
SUMARI	XV
LLISTAT DE FIGURES I TAULES	XIX
LLISTAT D'ABREVIATURES I SÍMBOLS	XXVII
1. INTRODUCCIÓ	1
1.1. LES PROTEÏNES HMG	1
1.1.1. LES PROTEÏNES HMGA.....	5
1.1.2. LES PROTEÏNES HMGB	10
1.2. L'ESTRUCTURA DEL DNA	19
1.2.1. ESTRUCTURA QUÍMICA	19
1.2.2. PARÀMETRES GEOMÈTRICS DELS PARELLS DE BASES	23
1.2.3. ANGLES DE TORSIÓ DE L'ESQUELET	25
1.2.4. SOLC MENOR I SOLC MAJOR	26
1.2.5. TOPOLOGIA DEL DNA CIRCULAR	27
1.2.6. ALTRES ESTRUCTURES DEL DNA	29
1.2.7. INTERÈS DE LES SEQÜÈNCIES ATs	33
2. PLANTEJAMENT I OBJECTIUS	35
3. MATERIALS I MÈTODES	37
3.1. PROTEÏNES UTILITZADES	38
3.2. OBTENCIÓ DE LES PROTEÏNES HMGA	39
3.2.1. MUTAGÈNESI DIRIGIDA	39
3.2.2. EXPRESSIÓ EN <i>E. COLI</i>	43
3.2.3. TRACTAMENT DEL SEDIMENT BACTERIÀ.....	46
3.2.4. PURIFICACIÓ PER CROMATOGRAFIA	47
3.3. OBTENCIÓ DE LES PROTEÏNES HMGB	51
3.3.1. EXPRESSIÓ EN <i>E. COLI</i>	51
3.3.2. TRACTAMENT DEL SEDIMENT BACTERIÀ.....	54
3.3.3. PURIFICACIÓ DE LES PROTEÏNES DE FUSIÓ GST-HMGB.....	54
3.3.4. CROMATOGRAFIA DE BESCANVI CATIÒNIC I D'EXCLUSIÓ MOLECULAR	57

3.4. OBTENCIÓ DE LA PROTEÏNA NHP6A.....	59
3.4.1. EXPRESSIÓ EN <i>E. COLI</i>	59
3.4.2. TRACTAMENT DEL SEDIMENT BACTERIÀ	61
3.4.3. PRECIPITACIÓ AMB SULFAT D'AMONI	61
3.4.4. PURIFICACIÓ PER CROMATOGRÀFIA	63
3.5. ALTRES TÈCNiques PER A L'ANÀLISI DE PROTEÏNES	65
3.5.1. CONCENTRACIÓ PER ULTRAFILTRACIÓ	65
3.5.2. DETERMINACIÓ DE LA CONCENTRACIÓ.....	67
3.5.3. DIÀLISI	68
3.6. TÈCNiques ELECTROFORÈTIQUES PER A PROTEÏNES	69
3.6.1. ELECTROFORESI EN GELS DE POLIACRILAMIDA-SDS (SDS-PAGE)	69
3.6.2. ELECTROFORESI DE CRISTALLS DE PROTEÏNES I PÈPTIDS.....	73
3.6.3. TINCIÓ AMB PLATA	75
3.7. PREPARACIÓ DE LES MOSTRES D'OLIGONUCLEÒTIDS.....	77
3.8. CRISTAL·LITZACIÓ DELS COMPLEXOS.....	79
3.8.1. FONAMENTS DE LA CRISTAL·LITZACIÓ	79
3.8.2. METODOLOGIES DE CRISTAL·LITZACIÓ.....	81
3.8.3. OPTIMITZACIÓ DE CRISTALLS	84
3.9. DIFRACCIÓ DE RAIGS X I DETERMINACIÓ DE L'ESTRUCTURA	87
3.9.1. CRISTALLS I SIMETRIA	87
3.9.2. DIFRACCIÓ DE RAIGS X.....	89
3.9.3. EL PROBLEMA DE LES FASES.....	93
3.9.4. ANÀLISI DELS DIAGRAMES DE DIFRACCIÓ	94
3.9.5. RESOLUCIÓ DE L'ESTRUCTURA	96
3.9.6. AFINAMENT.....	99
3.9.7. VALIDACIÓ.....	100
3.10. ANÀLISI DE DNA MITJANÇANT ELECTROFORESI	103
3.10.1. GELS EMSA (ASSAIG DEL CANVI EN LA MOBILITAT ELECTROFORÈTICA)	104
3.10.2. ELECTROFORESI EN GELS D'AGAROSA	106
3.10.3. OBTENCIÓ D'UNA DISTRIBUCIÓ DE TOPOISÒMERS	109
3.10.4. GELS BIDIMENSIONALS	112
3.10.5. ASSAJOS D'INTRODUCCIÓ DE SUPERENROTLLAMENT	114
3.11. OBTENCIÓ I PURIFICACIÓ DE DNA PLASMÍDIC	118
3.12. MICROSCÒPIA DE FORÇA ATÒMICA (AFM)	119
4. RESULTATS I DISCUSSIÓ	121
4.1. OBTENCIÓ DE PROTEÏNES.....	121
4.1.1. OBTENCIÓ DE LES PROTEÏNES HMGA.....	121
4.1.1.1. Expressió i purificació de la proteïna HMGA1a.....	122
4.1.1.2. Purificació de la proteïna HMGA1b.....	126

4.1.1.3.	Obtenció de la proteïna HMGA1b(1-79): mutagènesi, expressió i purificació	128
4.1.2.	OBTENCIÓ DE LES PROTEÏNES HMGB	135
4.1.2.1.	Expressió i purificació de la proteïna HMGB1 Box AS	136
4.1.2.2.	Expressió i purificació de la proteïna HMGB1 Box BS.....	140
4.1.2.3.	Expressió i purificació de la proteïna HMGB1 Box ABS.....	144
4.1.2.4.	Expressió i purificació de la proteïna NHP6A.....	148
4.1.3.	RESUM DE LES PROTEÏNES OBTINGUDES PER ALS ASSAJOS CRISTAL·LOGRÀFICS	153
4.2.	ASSAJOS CRISTAL·LOGRÀFICS	155
4.2.1.	COMPLEXOS HMGA-DNA I PÈPTIDS <i>AT-HOOK</i> -DNA.....	156
4.2.2.	COMPLEXOS HMGB1 BOX AS-DNA.....	166
4.2.3.	COMPLEXOS HMGB1 BOX BS-DNA.....	179
4.2.4.	COMPLEXOS HMGB1 BOX ABS-DNA	191
4.2.5.	COMPLEXOS NHP6A-DNA	194
4.3.	ANÀLISI DELS DIAGRAMES DE DIFRACCIÓ	201
4.3.1.	OLIGONUCLEÒTID AATAATTATT I PÈPTID PRGRP.....	203
4.3.2.	PROTEÏNES HMGB1 BOX AS I BOX BS AMB DNA	206
4.4.	RESOLUCIÓ DE L'ESTRUCTURA DEL COMPLEX HMGB1 BOX AS-DNA.....	213
4.4.1.	ANÀLISI DELS DIAGRAMES DE DIFRACCIÓ.....	213
4.4.2.	CONSIDERACIONS PRÈVIES AL REEMPLAÇAMENT MOLECULAR	215
4.4.3.	REEMPLAÇAMENT MOLECULAR	217
4.4.4.	AFINAMENT I VALIDACIÓ.....	220
4.5.	ESTRUCTURA DEL COMPLEX HMGB1 BOX AS-DNA	223
4.5.1.	ESTRUCTURA GENERAL DEL CRISTALL BOX AS-ATATCGATAT	223
4.5.2.	INTERACCIONS PROTEÏNA-DNA	226
4.5.3.	ESTRUCTURA DEL DNA.....	228
4.5.4.	COMPARACIÓ AMB ALTRES BOX AS	231
4.5.5.	CONCLUSIONS	234
4.6.	ESTUDI DE LA INTRODUCCIÓ DE <i>SUPERCOILING</i> NEGATIU.....	235
4.7.	ESTUDIS DE LA TRANSICIÓ ESTRUCTURAL DE DNAs CIRCULARS AMB INSERTS ATs	239
4.7.1.	ELECTROFORESIS BIDIMENSIONALS	240
4.7.2.	MICROSCÒPIA DE FORÇA ATÒMICA (AFM).....	244
5.	DISCUSSIÓ GENERAL.....	247
6.	CONCLUSIONS.....	253
7.	BIBLIOGRAFIA	257
7.1.	REFERÈNCIES BIBLIOGRÀFIQUES	257
7.2.	BIBLIOGRAFIA COMPLEMENTÀRIA.....	269

ANNEXOS.....	271
A. SEQÜÈNCIA, MASSA MOLECULAR I COEF. D'EXTINCIÓ DE LES PROTEÏNES UTILITZADES.....	271
B. MEDIS DE CULTIU	273
C. INDUCCIÓ AMB IPTG.....	275
D. TAMPONS COMERCIALS DE CRISTAL·LITZACIÓ.....	277
E. LA UNITAT CRISTAL·LINA.....	285
F. FUNCIÓ DE PATTERSON I FUNCIÓ D'AUTOROTACIÓ	287
G. AMINOÀCIDS I ENLLAÇ PEPTÍDIC	291

Llistat de figures i taules

Figures

Fig. 1.1. Classificació i característiques de les proteïnes HMGs humanes.	2
Fig. 1.2. Diferents nivells d'expressió de les HMGs i els seus efectes en el desenvolupament i en el càncer.	3
Fig. 1.3. Diferents substrats d'unió de les proteïnes HMGs.	4
Fig. 1.4. Els gens <i>HMGA1</i> i <i>HMGA2</i> que donen lloc als quatre membres de les HMGAs [11].	5
Fig. 1.5. Esquema de les tres vies d'actuació de les proteïnes HMGA.	6
Fig. 1.6. Seqüències aminoacídiques de les proteïnes HMGA humanes [17].	7
Fig. 1.7. Interacció dels <i>AT-hooks</i> amb el DNA (estructures de RMN i de raigs X).	8
Fig. 1.8. Seqüències aminoacídiques de les tres HMGBs humanes.	10
Fig. 1.9. Funcions de l'HMGB1 dins i fora de la cèl·lula [28].	12
Fig. 1.10. Esquema dels tres mecanismes d'actuació de les proteïnes HMGB.	13
Fig. 1.11. Esquema de l'estructura de la proteïna humana l'HMGB1.	14
Fig. 1.12. Seqüències aminoacídiques dels dominis HMG-box d'algunes proteïnes.	15
Fig. 1.13. Interaccions proteïna-DNA de les proteïnes HMGB1 (Box A i Box B), HMGBD i NHP6A.	16
Fig. 1.14. Model de l'estructura del TFAM (HMGB del mitocondri).	17
Fig. 1.15. Enllaç fosfodièster entre dos nucleòtids (esquerra) i bases nitrogenades del DNA (dreta).	19
Fig. 1.16. Esquema de l'estructura secundària del DNA.	20
Fig. 1.17. Models de les formes B, A i Z del DNA.	21
Fig. 1.18. Plegament del sucre (<i>sugar pucker</i>).	22
Fig. 1.19. Conformacions <i>syn</i> i <i>anti</i> de l'enllaç glicosídic en la forma B i Z del DNA.	22
Fig. 1.20. Paràmetres geomètrics pels parells de bases.	24
Fig. 1.21. Angles de torsió de l'esquelet en àcids nucleics i roda conformacional dels angles de torsió.	25
Fig. 1.22. Grups exposats en els costats del solc menor i major.	26
Fig. 1.23. Relacions entre el nombre d'enllaç (Lk), el gir (Tw) i la torsió (Wr) per a un DNA de 260 pb [55].	28
Fig. 1.24. Formació d'un cruciform a partir d'una seqüència desenrotllada contenint.	29
Fig. 1.25. Formació d'unions de Holliday durant la recombinació homòloga.	30
Fig. 1.26. Comparació del B-DNA amb el PX-DNA i esquema de la seva formació en un DNA circular [68].	31
Fig. 1.27. Plasmidis amb inserts de 22 pb i 200 pb (pUC-19 22A i pUC-19 220A) i electroforesis bidimensionals per a l'observació de la transició estructural [68].	32
Fig. 1.28. Imatges d'AFM de les estructures tipus manuelles amb diferents longituds del segment cohesiu [68].	32
Fig. 1.29. Aparellament Hoogsteen [80].	33
Fig. 3.1. Esquema del procés d'obtenció de l'estructura dels complexos proteïna-DNA.	37
Fig. 3.2. Esquema de la mutagènesi dirigida.	39
Fig. 3.3. Esquema del plasmidi pET-3a.	43
Fig. 3.4. Detalls de la cromatografia de bescanvi catiònic utilitzada i dibuix esquemàtic del principi de separació d'aquesta columna.	48
Fig. 3.5. Detalls de la cromatografia d'exclusió molecular utilitzada i dibuix esquemàtic del principi de separació d'aquesta columna.	50
Fig. 3.6. Esquema del plasmidi pGEX-2T que codifica la proteïna de fusió GST-HMGB1 (esquerra) i dels fragments de la proteïna HMGB1 amb els que hem treballat (dreta).	51
Fig. 3.7. Esquema de la purificació de la proteïna HMGB amb una columna de glutatió sefarosa.	55

Fig. 3.8. Detalls de la cromatografia d'afinitat Glutatió Sefarosa i dibuix esquemàtic del principi de separació d'aquesta columna.....	55
Fig. 3.9. Detalls de la cromatografia de bescanvi catiònic per a les proteïnes HMGB i dibuix esquemàtic del principi de separació d'aquesta columna.....	57
Fig. 3.10. Detalls de la cromatografia d'exclusió molecular per a les proteïnes HMGB i dibuix esquemàtic del principi de separació d'aquesta columna.	58
Fig. 3.11. Esquema del plasmidi pET-11a que codifica el gen de la proteïna de llevat NHP6A.....	59
Fig. 3.12. Detalls de la cromatografia de bescanvi aniònic utilitzada per a la proteïna NHP6A i dibuix esquemàtic del principi de separació d'aquesta columna.	63
Fig. 3.13. Detalls de la cromatografia de bescanvi catiònic utilitzada per a la proteïna NHP6A i dibuix esquemàtic del principi de separació d'aquesta columna.	64
Fig. 3.14. Detalls de la cromatografia d'exclusió molecular utilitzada per a la proteïna NHP6A i dibuix esquemàtic del principi de separació d'aquesta columna.	64
Fig. 3.15. Els diferents concentradors fets servir al llarg d'aquest treball.	65
Fig. 3.16. Imatge d'un Nanodrop.....	67
Fig. 3.17. Diagrama del procés de cristal·lització d'un cristall [92].....	80
Fig. 3.18. Esquema de les dues modalitat de cristal·lització per difusió de vapor.	82
Fig. 3.19. Imatges d'una placa Linbro (a), de diferents plaques d'alt rendiment (b) i d'una càpsula (c).	82
Fig. 3.20. Fotografies del robots fets servir per a la realització d'assajos cristal·logràfics.	83
Fig. 3.21. Esquema d'un <i>loop</i> (esquerra) i detall amb un cristall (dreta).....	84
Fig. 3.22. Exemple d'una quadrícula de condicions per a l'optimització de cristalls.	85
Fig. 3.23. Diferents tècniques de sembrat: <i>macroseeding</i> , <i>microseeding</i> i <i>streak seeding</i>	86
Fig. 3.24. Dibuix de la unitat asimètrica, la cel·la unitària, la xarxa i el cristall.....	88
Fig. 3.25. Llei de Bragg ($\text{sen } \theta = n\lambda / 2d$).....	88
Fig. 3.26. Esfera d'Ewald.....	89
Fig. 3.27. Imatge d'un difractor mètre.	90
Fig. 3.28. Imatge aèria de l'ERSF (<i>European Synchrotron Radiaton Facility</i>) a Grenoble (esquerra) i del sincrotró ALBA (dreta) a Cerdanyola del Vallès.....	90
Fig. 3.29. Esquema d'un sincrotró i detall d'una <i>beamline</i> del sincrotró ERSF.....	91
Fig. 3.30. Imatge de difracció d'un cristall.	93
Fig. 3.31. Esquema dels diferents passos per a l'obtenció de l'estructura per difracció de raigs X.....	94
Fig. 3.32. Exemple d'un mapa de densitat electrònica (blau) amb el mapa de diferències (verd i vermell) sobre el traçat d'un model estructural proteïna-DNA.....	99
Fig. 3.33. Diagrama de <i>Ramachandran</i> per a les alanines i residus tipus alanina [113].....	101
Fig. 3.34. Sortida del programa <i>Molprobability</i>	102
Fig. 3.35. Esquema d'un assaig de canvi en la mobilitat electroforètica.	104
Fig. 3.36. Esquema de l'electroforesi horitzontal i imatge d'una cubeta per a aquest tipus d'electroforesi.	107
Fig. 3.37. Acció de la topoisomerasa I.....	109
Fig. 3.38. Relació entre la mobilitat electroforètica i el grau de superenrotllament de DNAs circulars [117][115]...	109
Fig. 3.39. Esquema de les reaccions de relaxació de la topoisomerasa I en presència de diferents concentracions d'intercalador [122].....	110
Fig. 3.40. Electroforesis bidimensionals amb i sense transició estructural amb alta.....	113
Fig. 3.41. Electroforesis bidimensionals de la transició de la forma B a la Z del DNA [124].....	113
Fig. 3.42. Esquema d'un assaig d'introducció de <i>supercoiling</i> [125].....	115
Fig. 3.43. Distribucions de topoisòmers analitzada en un gel d'agarosa contenint cloroquina.	117
Fig. 3.44. Esquema del procés de la Midi-Prep.	118

Fig. 3.45. Funcionament de l'AFM.....	119
Fig. 4.1. Esquema de les proteïnes HMGA1a, HMGA1b i HMGA1b(1-79).....	121
Fig. 4.2. Gel de poliacrilamida-SDS de la inducció i de l'extracte enriquit de l'HMGA1a.....	123
Fig. 4.3. Cromatograma i gels de poliacrilamida-SDS corresponents a la columna MACRO-PREP 25S (Tris/NaCl pH=7) per a la proteïna HMGA1a.	125
Fig. 4.4. Cromatograma i gel SDS-PAGE corresponents a la columna CM52 (Tris/NaCl pH=7) per a la proteïna HMGA1b.....	127
Fig. 4.5. Gels SDS-PAGE corresponents a la columna de repurificació CM52 (Tris/NaCl pH=7).....	128
Fig. 4.6. Cromatograma de la seqüenciació del DNA plasmídic mutat.....	129
Fig. 4.7. Gel de poliacrilamida-SDS de la inducció i de l'extracte enriquit de la proteïna HMGA1b(1-79).	130
Fig. 4.8. Gel de poliacrilamida-SDS per a comparar la forma HMGA1b(1-79) obtinguda per mutagènesi dirigida amb l'HMGA1b original.....	131
Fig. 4.9. Cromatograma i gel SDS-PAGE corresponents a la columna CM52 (Tris/NaCl pH=7) per a la proteïna HMGA1b(1-79).....	132
Fig. 4.10. Cromatograma i gel SDS-PAGE corresponents a la columna Superdex 75 per a l'HMGA1b(1-79).	133
Fig. 4.11. Gel SDS-PAGE de les fraccions pures i concentrades de la proteïna HMGA1b(1-79) obtingudes després de la repurificació per exclusió molecular.	134
Fig. 4.12. Esquema de la proteïna HMGB1 (i els seus fragments) i de la proteïna NHP6A.....	135
Fig. 4.13. Gel SDS-PAGE de les mostres analitzades durant la inducció de la proteïna HMGB1 Box AS.	136
Fig. 4.14. Gel SDS-PAGE corresponent a la purificació GST per a l'HMGB1 Box AS.....	137
Fig. 4.15. Cromatograma i gel d' SDS corresponents a la columna HiTrap SP FF per a l'HMGB1 Box AS.....	138
Fig. 4.16. Cromatograma i gel SDS-PAGE corresponents a la columna Superdex 75 per a l'HMGB1 Box AS.....	139
Fig. 4.17. Gel SDS-PAGE de les fraccions pures i concentrades obtingudes per a l'HMGB1 Box AS.....	140
Fig. 4.18. Gel SDS-PAGE de les mostres analitzades durant la inducció per a la proteïna HMGB1 Box BS.	141
Fig. 4.19. Gel SDS-PAGE corresponent a la purificació GST per a l'HMGB1 Box BS.....	141
Fig. 4.20. Cromatograma i gel SDS-PAGE corresponents a la columna HiTrap SP FF per a l'HMGB1 Box BS. ...	142
Fig. 4.21. Cromatograma i gel SDS-PAGE corresponents a la columna Superdex 75 per a l'HMGB1 Box BS.....	143
Fig. 4.22. Gel SDS-PAGE de les fraccions pures i concentrades obtingudes per a l'HMGB1 Box BS.	144
Fig. 4.23. Gel SDS-PAGE de les mostres analitzades durant la inducció de la proteïna HMGB1 Box ABS.	144
Fig. 4.24. Gel SDS-PAGE corresponent a la purificació GST per a l'HMGB1 Box ABS.....	145
Fig. 4.25. Cromatograma i gel SDS-PAGE corresponents a la columna HiTrap SP per a l'HMGB1 Box ABS.	146
Fig. 4.26. Cromatograma i gel SDS-PAGE corresponents a la columna Superdex 75 per a l'HMGB1 Box ABS....	147
Fig. 4.27. Gel SDS-PAGE de les fraccions pures i concentrades obtingudes per a l'HMGB1 Box ABS.....	147
Fig. 4.28. Gel SDS-PAGE de les mostres analitzades durant la inducció de la proteïna NHP6A.	148
Fig. 4.29. Gel SDS-PAGE corresponent a la precipitació amb sulfat d'amoni de l'NHP6A.....	149
Fig. 4.30. Gel SDS-PAGE corresponent a la columna DEAE per a l'NHP6A.	149
Fig. 4.31. Cromatograma i gel SDS-PAGE corresponents a la columna HiTrap SP FF per a l'NHP6A.	150
Fig. 4.32. Cromatograma i gel SDS-PAGE corresponents a la columna Superdex 75 per a l'NHP6A.....	151
Fig. 4.33. Gel SDS-PAGE de les fraccions pures i concentrades obtingudes de l'NHP6A.	152
Fig. 4.34. Seqüència aminoacídica dels <i>AT-hooks</i> de la proteïna HMGA1a.....	156
Fig. 4.35. Imatges, condicions de cristal·lització i diagrames de difracció dels millors resultats obtinguts en assajos previs [128]. a) HMGA1a truncada + CGCATTAATGCG. b) Pèptid PRGRP + AATAATTATT.	157
Fig. 4.36. Assaigs EMSA per a la formació de complexos entre l'HMGA1b(1-79) i diferents oligonucleòtids.....	158
Fig. 4.37. Imatge, condicions de cristal·lització i diagrama de difracció d'un cristall corresponent a la proteïna HMGA1a truncada i l'oligonucleòtid CGCATTAATGCG (gota R3, cristall R3-1).	160

Fig. 4.38. Imatge, condicions de cristal·lització i diagrama de difracció d'un cristall corresponent al pèptid KRPRGRP (<i>AT-hook</i> II) i l'oligonucleòtid CGCATTAATGCG (gota R24, cristall R24-1).	161
Fig. 4.39. Imatge, condicions de cristal·lització i diagrama de difracció d'un cristall corresponent al pèptid RKPRGRPKK (<i>AT-hook</i> III) i l'oligonucleòtid CGCATTAATGCG (gota R90, cristall R90-1).	162
Fig. 4.40. a. Gel EMSA de l'oligonucleòtid CGCATTAATGCG. b. Condicions de cristal·lització i diagrama de difracció d'un cristall corresponent al pèptid KRPRGRP (<i>AT-hook</i> II) i aquest oligonucleòtid (cristall R51-1).	162
Fig. 4.41. Imatge, condicions de cristal·lització i diagrama de difracció d'un cristall corresponent al pèptid PRGRP i l'oligonucleòtid AATAATTATT (gota R107, cristall R107-1).	163
Fig. 4.42. Imatge, condicions de cristal·lització i diagrama de difracció d'un cristall corresponent al pèptid PRGRP i l'oligonucleòtid AATAATTATT (gota 45, cristall 45-2).	164
Fig. 4.43. Gels de poliacrilamida dels cristalls de la proteïna HMGA1a i del pèptid <i>AT-hook</i> III.	165
Fig. 4.44. Variants amb extrems no roms de l'oligonucleòtid AATAAATTTATT.	166
Fig. 4.45. Assaigs EMSA per a la formació de complexos entre la Box AS i diferents oligonucleòtids.	167
Fig. 4.46. Imatge i condicions de cristal·lització de la gota RS10-B8 corresponent a la proteïna Box AS i a l'oligonucleòtid GTTATTAATAAC.	170
Fig. 4.47. Imatge i condicions de cristal·lització de la gota RS13-H4 corresponent a la proteïna Box AS i a l'oligonucleòtid AATAAATTTATT.	171
Fig. 4.48. Imatge i condicions de cristal·lització de la gota RS13-C4 corresponent a la proteïna Box AS i a l'oligonucleòtid AATAAATTTATT.	171
Fig. 4.49. Imatge i condicions de cristal·lització de la gota AS19 corresponent a la proteïna Box AS i a l'oligonucleòtid AATAAATTTATT.	171
Fig. 4.50. Imatge, condicions de cristal·lització i diagrama de difracció d'un cristall corresponent al domini Box AS i l'oligonucleòtid AATAAATTTATT (gota AS20, cristall AS20-1).	172
Fig. 4.51. Imatge i condicions de cristal·lització de la gota AS31 corresponent a la proteïna Box AS i a l'oligonucleòtid ATATCGATAT.	173
Fig. 4.52. Imatge, condicions de cristal·lització i diagrama de difracció d'un cristall corresponent al domini Box AS i l'oligonucleòtid ATATCGATAT (gota AS32, cristall AS32-1).	174
Fig. 4.53. Imatge, condicions de cristal·lització i diagrama de difracció d'un cristall corresponent al domini Box AS i l'oligonucleòtid ATATCGATAT (gota AS32, cristall AS32-3).	174
Fig. 4.54. Imatge, condicions de cristal·lització i diagrama de difracció d'un cristall corresponent al domini Box AS i l'oligonucleòtid ATATCGATAT (gota AS67, cristall AS67-1).	175
Fig. 4.55. Imatge, condicions de cristal·lització i diagrama de difracció d'un cristall corresponent al domini Box AS i l'oligonucleòtid ATATCGATAT (gota AS35, cristall AS35-2).	176
Fig. 4.56. Imatge, condicions de cristal·lització i diagrama de difracció d'un cristall corresponents al domini Box AS i l'oligonucleòtid ATATCGATAT (gota AS36, cristall AS36-1).	176
Fig. 4.57. Gel poliacrilamida-SDS dels cristalls de la proteïna Box AS.	177
Fig. 4.58. Assaigs EMSA per a la formació de complexos entre la Box BS i diferents oligonucleòtids.	179
Fig. 4.59. Imatge i condicions de cristal·lització de la gota RS18-A11 corresponent a la proteïna Box BS i a l'oligonucleòtid AATAAATTTATT.	182
Fig. 4.60. Imatge i condicions de cristal·lització de la gota RS15-B4 corresponent a la proteïna Box BS i a l'oligonucleòtid AATAAATTTATT.	182
Fig. 4.61. Imatge, condicions de cristal·lització i diagrama de difracció d'un cristall corresponent al domini Box BS i l'oligonucleòtid AATAAATTTATT (gota D4 de la placa RS04, cristall RS04-D4-3).	183
Fig. 4.62. Evolució dels cristalls d'una gota i les seves condicions de cristal·lització.	183
Fig. 4.63. Imatge, condicions de cristal·lització i diagrama de difracció d'un cristall corresponent al domini Box BS i l'oligonucleòtid AATAAATTTATT (gota D5 de la placa RS05, cristall RS05-D5-5).	184

Fig. 4.64. Imatge, condicions de cristal·lització i diagrama de difracció d'un cristall corresponent al domini Box BS i l'oligonucleòtid AATAAATTTATT (gota A6 de la placa RS07, cristall RS07-A6-1).....	184
Fig. 4.65. Imatge, condicions de cristal·lització i diagrama de difracció d'un cristall corresponent al domini Box BS i l'oligonucleòtid AATAAATTTATT (gota A4 de la placa RS04, cristall RS04-A4-2).....	185
Fig. 4.66. Imatge, condicions de cristal·lització i diagrama de difracció d'un cristall corresponents al domini Box BS i l'oligonucleòtid AATAAATTTATT (gota BS4, cristall BS4-1).....	186
Fig. 4.67. Imatge, condicions de cristal·lització i diagrama de difracció d'un cristall corresponent al domini Box BS i l'oligonucleòtid AATAAATTTATT (gota B5 de la placa RS20, cristall RS20-B5-2).....	187
Fig. 4.68. Imatge, condicions de cristal·lització i diagrama de difracció d'un cristall corresponent al domini Box BS i l'oligonucleòtid AATAAATTTATT (gota BS77, cristall BS77-1).....	187
Fig. 4.69. Imatge, condicions de cristal·lització i diagrama de difracció d'un cristall corresponent al domini Box AS i l'oligonucleòtid AATAAATTTATTT (gota BS67, cristall BS67-1).....	188
Fig. 4.70. Imatge i condicions de cristal·lització de la gota BS104 corresponent a la proteïna Box BS i l'oligonucleòtid AATAATCGCGATTATT.....	189
Fig. 4.71. Gel poliacrilamida-SDS d'alguns cristalls de la proteïna Box BS.....	190
Fig. 4.72. Assaigs EMSA per a la formació de complexos entre la Box ABS i els oligonucleòtids sub21 i sub23... ..	191
Fig. 4.73. Imatge i condicions de cristal·lització de la gota ABS34 corresponent a la proteïna Box ABS i a l'oligonucleòtid sub21.....	193
Fig. 4.74. Assaigs EMSA per a la formació de complexos entre l'NHP6A i els diferents oligonucleòtids.....	194
Fig. 4.75. Imatge i condicions de cristal·lització de la gota RS23-H12-d1 corresponent a la proteïna NHP6A i l'oligonucleòtid AATAAATTTATT.....	197
Fig. 4.76. Imatge i condicions de cristal·lització de la gota RS23-H08-d1 corresponent a la proteïna NHP6A i l'oligonucleòtid AATAAATTTATT.....	197
Fig. 4.77. Imatge i condicions de cristal·lització de la gota RS24-C5 corresponent a la proteïna NHP6A i a l'oligonucleòtid AATAAATTTATT.....	198
Fig. 4.78. Imatge i condicions de cristal·lització de la gota NH19 corresponent a la proteïna NHP6A i a l'oligonucleòtid AATAAATTTATT.....	198
Fig. 4.79. Imatge i condicions de cristal·lització de la gota NH28 corresponent a la proteïna NHP6A i a l'oligonucleòtid AATAAATTTATT.....	198
Fig. 4.80. Imatge i condicions de cristal·lització de la gota NH33 corresponent a la proteïna NHP6A i a l'oligonucleòtid AATAATCGCGATTATT.....	199
Fig. 4.81. Imatge i condicions de cristal·lització de la gota NH37 corresponent a la proteïna NHP6A i a l'oligonucleòtid ATATATATATATAT.....	199
Fig. 4.82. Diagrames de difracció i dades del cristall 45-2.....	203
Fig. 4.83. Model de l'estructura corresponent al cristall 45-2 que conté l'oligonucleòtid AATAATTATT.....	204
Fig. 4.84. Model de la subcel·la de l'estructura corresponent al cristall 45-2 que conté l'oligonucleòtid AATAATTATT.....	205
Fig. 4.85. Comparació del model construït amb les dades del cristall 45-2.....	205
Fig. 4.86. Model de l'estructura corresponent al cristall AS20 que conté l'oligonucleòtid AATAAATTTATT.....	208
Fig. 4.87. Probabilitats de <i>Matthews</i> per al coeficient de <i>Matthews</i> (V_m) i per al contingut en solvent.....	209
Fig. 4.88. Model de l'estructura corresponent al cristall RS20-B5-2 que conté l'oligonucleòtid AATAAATTTATT.....	211
Fig. 4.89. Algunes gràfiques dels paràmetres d'avaluació obtingudes al principi de l'escalat.....	214
Fig. 4.90. Sortida de l'escalat del programa <i>HKL2000</i> mostrant els valors estadístics obtinguts.....	214
Fig. 4.91. Mapa de Patterson per al cristall AS32-3.....	216
Fig. 4.92. <i>Self-rotation function</i> per al cristall AS32-3.....	217

Fig. 4.93. Models del dúplex de DNA i de la proteïna fets servir per al reemplaçament molecular.....	218
Fig. 4.94. Contingut de la unitat asimètrica.	219
Fig. 4.95. Sortida del programa <i>Molprobit</i> [114] obtinguda per al model final del cristall AS32-3.	221
Fig. 4.96. <i>Ramachandran</i> per al model final.	222
Fig. 4.97. Organització dins de la cel·la del complex Box AS-DNA.....	223
Fig. 4.98. Interacció de dos dominis Box AS per a doblegar el DNA.....	224
Fig. 4.99. Mapes de densitat electrònica de l'estructura.	225
Fig. 4.100. Interaccions entre els dos dominis Box AS a través d'una cadena d'aigües.	226
Fig. 4.101. Interaccions entre la proteïna i el DNA.	227
Fig. 4.102. Comparació amb altres oligonucleòtids interaccionant amb HMG-boxes.....	230
Fig. 4.103. Comparació entre diferents dominis Box A i les conformacions de les fenilalanines 37.	232
Fig. 4.104. Comparació entre la forma reduïda i l'oxidada del domini Box A.	233
Fig. 4.105. Assaig d'introducció de <i>supercoiling</i> a alta força iònica (~140 mM).	235
Fig. 4.106. Assaig d'introducció de <i>supercoiling</i> a baixa força iònica (~50 mM).	236
Fig. 4.107. Assaig d'introducció de <i>supercoiling</i> a baixa força iònica (~50 mM) i 0,6 µg/ml cloroquina.	237
Fig. 4.108. Assaig d'introducció de <i>supercoiling</i> a alta força iònica (~140 mM) per l'NHP6A.	238
Fig. 4.109. Esquema de la formació d'un fragment del tipus PX-DNA en un DNA circular [68].	239
Fig. 4.110. Esquema dels plasmidis amb inserts ATs que s'han fet servir en aquest treball.....	239
Fig. 4.111. Distribució de topoisòmers per a pSTAT-CEN2 (control) (esquerra) i pSTAT-CEN2-S2 (dreta).	240
Fig. 4.112. Electroforesi bidimensional per a una distribució de topoisòmers de pSTAT-CEN2-S2 sense (a) i amb cloroquina (b) a la primera dimensió.	241
Fig. 4.113. Electroforesis bidimensionals per pSTAT-CEN2 (control) i pSTAT-CEN2-S2 (2 inserts) sense (a, b) i amb cloroquina (c, d) a la primera dimensió.....	242
Fig. 4.114. Distribucions de topoisòmers (a, b) i electroforesis bidimensionals (b, c).....	243
Fig. 4.115. Imatges d'AFM per als plasmidis pSTAT-CEN2 (a), pSTAT-CEN2-S1 (b) i pSTAT-CEN2-S2 (c).	245

Taules

Taula 1.1. Nomenclatura revisada per a les proteïnes HMG [6].	1
Taula 1.2. Nomenclatura revisada per a les proteïnes HMGA humanes [12].	5
Taula 1.3. Classificació de les proteïnes amb motius HMG-box segons la seva especificitat d'unió al DNA.	11
Taula 1.4. Comparació entre les formes B, A i Z del DNA [55].	21
Taula 1.5. Valors promitjos dels paràmetres geomètrics dels parells de bases per a cristalls de la forma A i B.	24
Taula 3.1. Resum dels principals passos per a l'obtenció de les diferents proteïnes.....	38
Taula 3.2. Disseny dels <i>primers</i>	40
Taula 3.3. Components de la barreja de reacció.	41
Taula 3.4. Detalls dels cicles tèrmics.	41
Taula 3.5. Especificacions dels concentradors utilitzats i proteïna concentrada.	66
Taula 3.6. Composició dels gels de poliacrilamida-SDS.....	71
Taula 3.7. Composició del tampó de cubeta per a electroforesi en gels de poliacrilamida-SDS.	71
Taula 3.8. Composició del tampó de mostres per a electroforesi en gels de poliacrilamida-SDS.	72
Taula 3.9. Composició de la solució de tinció per a gels de poliacrilamida-SDS.	72
Taula 3.10. Composició de la solució destenyidora per a gels de poliacrilamida-SDS.	72
Taula 3.11. Composició dels gels de poliacrilamida-SDS (Colorado).	73

Taula 3.12. Composició dels gels AU-PAGE per a cristalls de pèptids.....	74
Taula 3.13. Composició del tampó de mostres per a electroforesi en gels de poliacrilamida-àcid acètic-urea.....	74
Taula 3.14. Composició de la solució de fixació amb àcid pícric i metanol.....	75
Taula 3.15. Protocol per a la tinció amb plata amb el <i>SilverQuest Silver Staining kit</i> (Invitrogen).	76
Taula 3.16. Coeficients d'extinció a 260 nm, 25°C, pH neutre i cadena senzilla de DNA [91].....	77
Taula 3.17. Coeficients d'extinció i concentracions dels oligonucleòtids preparats.	78
Taula 3.18. Factors que afecten la solubilitat del complex.	80
Taula 3.19. Concentracions d'agarosa o poliacrilamida utilitzades per a la separació electroforètica de fragments lineals de DNA de diferents mides [117].	103
Taula 3.20. Composició dels gels EMSA per a les proteïnes HMGA.....	105
Taula 3.21. Composició dels gels EMSA per a les proteïnes HMGB.....	105
Taula 3.22. Reaccions d'incubació per a un gel EMSA en el cas de les proteïnes HMGB.	106
Taula 3.23. Tampons de cubeta per a l'electroforesi en gels d'agarosa [121]......	108
Taula 3.24. Tampó de mostres per a gels d'agarosa [121]......	108
Taula 3.25. Reaccions d'incubació per a l'obtenció d'una distribució de topoisòmers.....	111
Taula 3.26. Condicions de les electroforesis bidimensionals.	114
Taula 3.27. Reaccions d'incubació per a un gel <i>supercoiling</i>	116
Taula 4.1. Fraccions pures i concentrades de les proteïnes HMGA i HMGB pels assajos cristal·logràfics.	153
Taula 4.2. Resum dels assajos cristal·logràfics per a les proteïnes HMGA i els pèptids <i>AT-hook</i> amb DNA.....	159
Taula 4.3. Condicions de cristal·lització de la proteïna HMGD [129][39][44]......	168
Taula 4.4. Resum dels assajos cristal·logràfics per a la proteïna HMGB1 Box AS.....	169
Taula 4.5. Resum dels assajos cristal·logràfics per a la proteïna HMGB1 Box BS.....	180
Taula 4.6. Resum dels assajos cristal·logràfics per a la proteïna HMGB1 Box ABS.....	192
Taula 4.7. Resum dels assajos cristal·logràfics per a la proteïna NHP6A.....	195
Taula 4.8. Resum dels cristalls amb millors diagrames de difracció de les proteïnes HMG.....	202
Taula 4.9. Empaquetament dels diferents cristalls de les proteïnes HMGB1.....	207
Taula 4.10. Dades de l'anàlisi dels diagrames de difracció i refinament.....	215
Taula 4.11. Dades obtingudes durant l'afinament del cristall AS32-3.....	221
Taula 4.12. Paràmetres geomètrics dels parells de bases i amplada dels solcs del dúplex doblegat.....	228
Taula 4.13. Paràmetres geomètrics de l'esquelet sucre-fosfat del dúplex doblegat.....	229
Taula 4.14. Valors dels r.m.s.d. per a les superposicions del Box A d'aquesta estructura amb d'altres.....	232

Llistat d'abreviatures i símbols

Ac: acetat (CH_3COO^-)

Amp: antibiòtic ampicil·lina

Amp^r: resistència a l'antibiòtic ampicil·lina

Aprox.: aproximadament

AT: adenina i timina

AT-hooks: motius d'unió al solc menor del DNA de les proteïnes HMGA (preferentment en seqüències riques en AT)

AU-PAGE: gel electroforètic de poliacrilamida-àcid acètic-urea

Box AS: fragment de la proteïna HMGB1 que conté el motiu HMG-box anomenat Box A (residus 7-80 de l'HMGB1)

Box BS: fragment de la proteïna HMGB1 que conté el motiu HMG-box anomenat Box B, (residus 92-164 de l'HMGB1)

Box ABS: fragment de la proteïna HMGB1 que conté els dos motius HMG-box de la proteïna (Box A i Box B) (residus 7-164 de l'HMGB1)

BrEt: bromur d'etidi

Cac: cacodilat (dimetil-arsenat)

Cam: antibiòtic cloramfenicol

CHES: àcid 2-(ciclohexilamino)etanosulfònic

CM: carboximetil cel·lulosa

Cp.: càpsula

C-terminal: extrem carboxil terminal d'un pèptid o d'una proteïna

Cv: volum empaquetat d'una columna (de l'anglès *column volume*)

DEAE: dietilaminoetil

DMSO: dimetilsulfòxid

DNA: àcid desoxiribonucleic

DNasa: desoxiribonucleasa

DO: densitat òptica

DTT: ditiotreitòl

ERSF: *European Synchrotron Radiation Facilities*

Eq.: equació

Fig.: figura

FPLC: *Fast Protein Liquid Chromatography*

Fr.: fracció

FT: *flow through* (nom que s'assigna al volum eluït que es recull mentre s'introdueix la mostra de partida a una columna cromatogràfica)

GST: glutatió S-transferasa

HEPES: N-[2-hidroxietil]piperazina-N'[2-àcid etanosulfònic]

HMG: *High Mobility Group*

HMG-box: motiu d'unió al solc menor del DNA de les proteïnes HMGB

IPTG: isopropil- β -D-tiogalactopirànòsid

Isop.: isopropanol (2-propanol)

Lac: lactosa

LB: medi Luria-Bertani

MAD: Dispersió Anòmala Múltiple (de l'anglès *Multi-wavelength Anomalous Diffraction*)

MALDI-TOF: tècnica d'ionització suau utilitzada en espectroscòpia de masses (de l'anglès *Matrix-Assisted Laser Desorption/Ionization Time-Of-Flight*)

MES: àcid 2-morfolinoetanosulfònic

MIR: Reemplaçament Isomorf Múltiple (de l'anglès *Multiple Isomorphous Replacement*)

MP: mostra de partida (nom que s'assigna a la mostra que s'introdueix a la columna cromatogràfica)

MPD: 2-metil-2,4 pentanodiol

MR: Reemplaçament Molecular (de l'anglès *Molecular Replacement*)

MWCO: massa molecular límit (de l'anglès *molecular weight cut off*)

NLS: *Nuclear Localisation Signal domain*

N-terminal: extrem amino terminal d'un pèptid o d'una proteïna

O.N.: tota la nit (de l'anglès *overnight*)

p.: pàgina

pb: parell de bases

PCB: Parc Científic de Barcelona

PEG: polietilenglicol

PEG MME: polietilenglicol monometil eter

pET: *plasmid for Expression by T7 RNA polymerase*

PFC: Projecte Final de Carrera

PSA: persulfat amònic

Recombinació V(D)J (*Variable, Diversity and Joining*): mecanisme de recombinació genètica que es dona en vertebrats pels quals se selecciona i s'assembla a l'atzar segments de gens que codifiquen proteïnes específiques amb papers importants en el sistema immunitari

r.m.s.d: mitjana quadràtica de les distàncies entre àtoms equivalents de dues molècules superposades, normalment proteïnes (de l'anglès *root-mean-squared deviation*)

rpm: revolucions per minut

SDS: dodecil sulfat de sodi

SDS-PAGE: gel poliacrilamida-SDS

SOB: medi de cultiu molt ric en nutrients desenvolupat per Douglas Hanahan al 1983 que és una adaptació del medi d'ús comú LB (de l'anglès *Super Optimal Broth*)

SOC: medi de cultiu SOB amb glucosa afegida (de l'anglès *Super Optimal broth with Catabolite repression*)

TCA: àcid tricloroacètic

TEMED: N,N,N',N'-tetrametilendiamina

Tm: temperatura de fusió

Tris: tris(hidroximetil)aminometà

Tritó X-100: Èter octilfenòlic del decaetilenglicol

u.a.: unitat asimètrica

UV: ultraviolat/da

v.: veure

1. Introducció

1.1. Les proteïnes HMG

L'estructura dels nucleosomes és modificada i organitzada per les histones i per d'altres proteïnes cromosòmiques no-histones, entre les quals les proteïnes HMG (*high mobility group*) representen el major grup. Les proteïnes HMG van ser descobertes fa 40 anys i van rebre el seu nom degut a la seva alta mobilitat electroforètica en gels de poliacrilamida [1]. Aquestes proteïnes actuen com a factors arquitectònics (*architectural factors*) modulant l'estructura de la cromatina i del nucleosoma. D'aquesta manera intervenen en diverses funcions cel·lulars, com ara la transcripció, la replicació i la reparació del DNA, afectant així l'expressió de nombrosos gens. Les HMGs pertanyen a una xarxa de proteïnes associades a la cromatina en la qual la seva implicació en processos relacionats amb el DNA té lloc de forma dinàmica [2]-[5].

Les proteïnes HMG, caracteritzades per tenir masses moleculars menors de 30 kDa i un alt contingut en aminoàcids carregats positivament, es van classificar en tres famílies segons els seus motius d'unió al DNA: HMGA (motiu AT-hook), HMGB (motiu HMG-box) i HMGN (domini d'unió al Nucleosoma). La nomenclatura d'aquestes proteïnes canòniques va ser adoptada l'any 2001 quan es va veure que d'altres proteïnes nuclears de diferents organismes contenien també aquests dominis [6]. Així es van anomenar HMGA, HMGB i HMGN a les proteïnes HMG canòniques prèviament anomenades HMG-I/Y, HMG-1/2 i HMG-14/17, respectivament, i proteïnes amb motiu HMG a la resta per a evitar-ne confusions (v. taula 1.1).

Taula 1.1. Nomenclatura revisada per a les proteïnes HMG [6].

Motius HMG	Proteïnes amb motius HMG	Proteïnes HMG canòniques	
		Nomenclatura antiga	Nova nomenclatura
<i>AT-hook</i>	Proteïnes <i>AT-hook</i>	HMG-I, HMG-Y, HMG-C	HMGA1,2...n
HMG-box	Proteïnes HMG-box	HMG-1, HMG-2	HMGB1,2...n
Domini d'unió al nucleosoma (NDB)	Proteïnes NBD	HMG-14, HMG-17	HMGN1,2...n

A la fig. 1.1 es mostren les tres famílies de proteïnes HMG juntament amb les seves característiques principals. Les HMGs participen en molts processos biològics comuns: estan implicades en el desenvolupament embrionari, regulen la transcripció i modulen la reparació del DNA. No obstant això, cada família HMG es distingeix de les altres pels seus motius d'unió al DNA característics, pels seus substrats d'unió preferents, pels canvis que produeixen en aquests substrats i pels diferents conjunts de processos cel·lulars en els que influeixen. Així les característiques concretes de cada família HMG determinen la seva forma de modificar l'estructura de la cromatina [5].

Les proteïnes HMGA contenen els motius *AT-hook*, que són segments de 7 a 9 aminoàcids no estructurats en solució que s'uneixen al solc menor del DNA preferentment en seqüències riques en ATs. La família de les HMGA comprèn quatre membres codificats per dos gens diferents (HMGA1a, HMGA1b, HMGA1c i HMGA2). Les HMGB presenten el domini HMG-box, d'uns 75 aminoàcids, que s'uneix també al solc menor del DNA però en aquest cas amb especificitat de seqüència limitada o inexistent. Els tres membres de la família HMGB estan codificats per gens individuals (HMGB1, HMGB2, HMGB3). Finalment, les proteïnes HMGN s'uneixen específicament al nucli del nucleosoma (octàmer d'histones + 146 pb) a través del seu domini d'unió al nucleosoma (NBD) de 30 aminoàcids de mida i que conté la seqüència consens RRSARLSA. El motiu NBD està completament conservat en totes les variants d'HMGN. Són conegudes cinc variants per a la família HMGN (HMGN1, HMGN2, HMGN3a, HMGN3b, HMGN4). Les HMGN només s'han trobat en vertebrats. Posteriorment es va incloure la proteïna NSBP1, de major massa molecular, reanomenada HMGN5 [4].

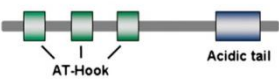
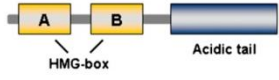
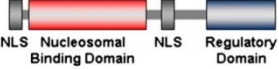
Família	Estructura	Mida	Membres	Interacció amb el DNA
HMGA		95-178 aa	HMGA1a, HMGA1b, HMGA1c, HMGA2	S'uneixen al solc menor del DNA en regions riques en seqüències AT
HMGB		199-214 aa	HMGB1, HMGB2, HMGB3	S'uneixen al solc menor del DNA amb especificitat de seqüència limitada o inexistent
HMGN		98-406 aa	HMGN1, HMGN2, HMGN3a, HMGN3b, HMGN4, HMGN5	S'uneixen al nucleosoma entre les voltes del DNA i l'octàmer d'histones

Fig. 1.1. Classificació i característiques de les proteïnes HMGs humanes.

A la figura es detallen algunes de les característiques principals de les tres famílies de proteïnes HMG [4].

Totes les proteïnes HMG tendeixen a reconèixer preferiblement particularitats locals de la cromatina més que seqüències específiques del DNA. Així, les HMGA han mostrat preferència per l'heterocromatina i diferents dominis subnuclears. Les HMGB es troben distribuïdes en tot el nucli de les cèl·lules, mentre que les HMGN es localitzen en els nucleosomes. Malgrat aquests llocs d'unió preferents, les HMGs es troben també repartides al llarg de tota la fibra de cromatina. Això podria estar relacionat amb l'abundància d'aquestes proteïnes. Les proteïnes HMG són les més abundants de les proteïnes d'unió a la cromatina no-histones en cèl·lules no diferenciades i en proliferació, en canvi presenten diferents nivells d'expressió durant la diferenciació i es troben pràcticament absents en cèl·lules diferenciades. Les HMGA són les menys abundants de les tres famílies, trobant-se una molècula d'HMGA per cada 1000 nucleosomes. La quantitat d'HMGN és d'una molècula per cada 100 nucleosomes mentre que la d'HMGB és aproximadament d'una molècula per cada 10-15 nucleosomes [4].

La sobreexpressió de les proteïnes HMG és una característica coneguda en cèl·lules tumorals. A més a més, ratolins genèticament modificats i cèl·lules amb nivells d'expressió d'HMG alterats indiquen que l'expressió de les proteïnes HMG està regulada en el desenvolupament i que els canvis en els nivells d'HMGs alteren el fenotip cel·lular i poden conduir al desenvolupament d'anomalies i malalties (v. fig. 1.2).

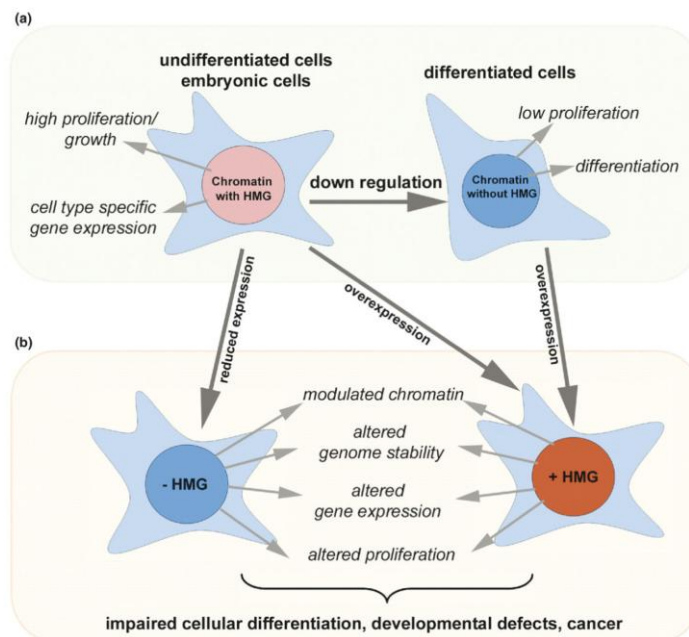


Fig. 1.2. Diferents nivells d'expressió de les HMGs i els seus efectes en el desenvolupament i en el càncer.

a) En condicions de desenvolupament normal, l'expressió de les proteïnes HMG està relacionada amb la diferenciació. b) La desregulació de l'expressió de les HMGs desemboca en una situació anòmla. Nivells augmentats o disminuïts causen alteracions en el perfil de transcripció cel·lular i poden conduir al desenvolupament de defectes, malalties i càncer [7].

Per tal d'interpretar l'activitat *in vivo* de les proteïnes HMG cal conèixer els substrats d'unió d'aquestes proteïnes *in vitro* i entendre els seus mecanismes de reconeixement, fet que representa encara un repte. Se sap que les proteïnes HMGA i HMGB s'uneixen a DNA doblegat, a cruciforms i DNA superenrotllat (v. fig. 1.3). Les proteïnes HMGA difereixen de les HMGB en la seva preferència d'unió a seqüències riques en ATs i en la seva habilitat per unir-se a la superfície del nucleosoma. Les HMGBs no s'uneixen eficientment als nucleosomes, i prefereixen els cromatosomes. La proteïna HMGN es diferencia de les dos altres famílies d'HMGs per la seva habilitat de reconèixer específicament l'estructura del nucleosoma pròpiament i d'assemblar-se com un homodímer a la superfície del mateix unint-se tant al dímer d'histones H2A-H2B com al DNA del nucleosoma [8].

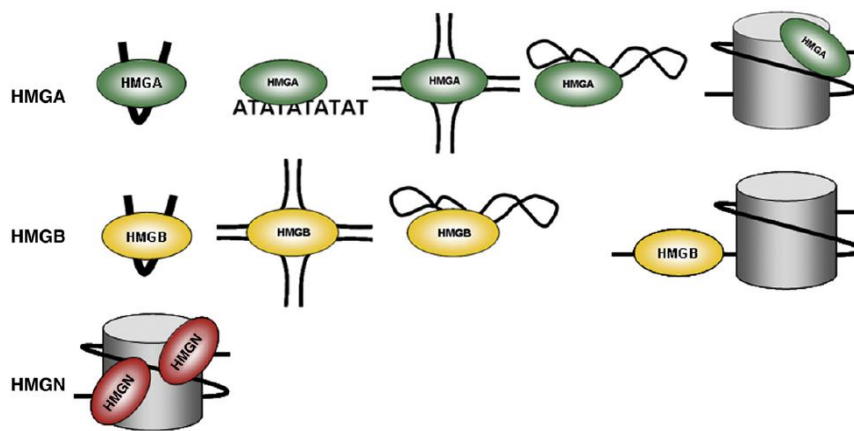


Fig. 1.3. Diferents substrats d'unió de les proteïnes HMGs.

Imatge extreta de [4].

Cada cop s'ha vist més clar que les proteïnes HMG i la histona H1 juguen uns rols similars i alhora oposats dins de la cromatina. L'accessibilitat del DNA dins la cromatina depèn, en un primer nivell, de la localització i mobilitat dels nucleosomes i de les modificacions postraduccionalment del nucli d'histones. Les modulacions estructurals de l'estructura de la cromatina per part de la histona H1 i les HMGs representaria un segon nivell [2]. Les proteïnes HMG i la histona H1 competeixen en la unió al nucleosoma. H1 i HMGN competeixen directament i un increment en els nivells d'HMGN incrementa la mobilitat de l'H1 [9]. En cultius cel·lulars, HMGBs i HMGA també competeixen amb la histona H1 i debiliten la seva unió a la cromatina; però HMGN, HMGA i HMGB semblen no competir entre elles [10]. Aquesta xarxa d'interaccions dinàmiques i contínues amb el nucleosoma podria donar una flexibilitat a l'estructura de la cromatina i un calibrat fi de la transcripció dels gens en resposta als canvis de l'entorn.

En aquest treball hem estudiat l'estructura de les proteïnes HMGA1a, HMGA1b i HMGB1. Per aquest motiu s'explicarà amb més detall les característiques d'aquestes dues famílies (HMGA i HMGB).

1.1.1. Les proteïnes HMGA

Família de les HMGA

Pel que fa a les HMGA, se sap que als humans existeixen dos gens que donen lloc a dos tipus de proteïnes HMGA, les A1 i les A2 (v. fig. 1.4 i taula 1.2). Dins del primer grup es produeixen a la vegada tres tipus diferents de proteïnes per *splicing* alternatiu: les variants més abundants HMGA1a (107 aa) i HMGA1b (96 aa) i una variant més rara, l'HMGA1c (179 aa). L'HMGA2 (109 aa) és codificada per un altre gen.

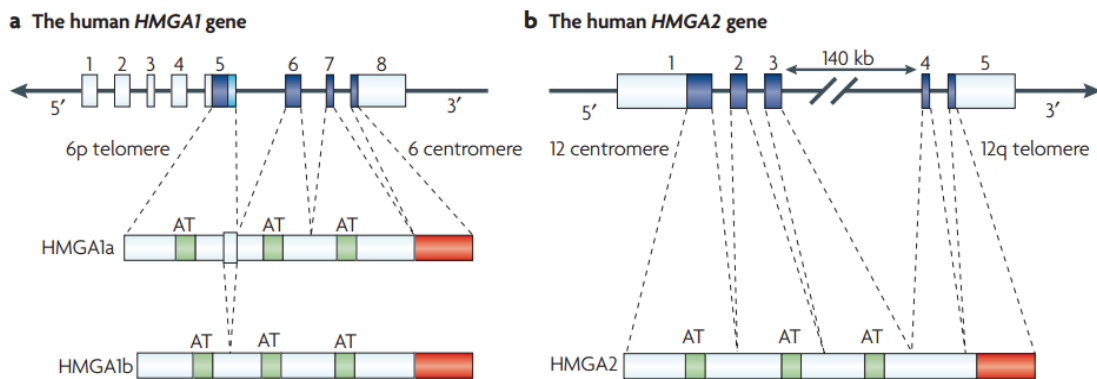


Fig. 1.4. Els gens *HMGA1* i *HMGA2* que donen lloc als quatre membres de les HMGA [11].

Taula 1.2. Nomenclatura revisada per a les proteïnes HMGA humanes [12].

Família de les HMGA humanes	
Nomenclatura antiga	Nova nomenclatura
HMG-I, HMGI o bé HMG I	HMGA1a
HMG-Y, HMGY o bé HMG Y	HMGA1b
HMG-I/R	HMGA1c
HMGI-C o bé HMGIC	HMGA2

Les proteïnes HMGA canòniques són expressades en la majoria de les plantes, els invertebrats i els vertebrats encara que una variant o les dues no es troben en algunes espècies com en el cas del llevat (no tenen HMGA) o el *Xenopus* (no té HMGA1) o bé estan substituïdes per proteïnes relacionades com en la *Drosophila* (proteïna D1, amb 11 *AT-hooks*) [4].

Funció

Les proteïnes HMGA participen en una ampla varietat de processos nuclears actuant com un factor estructural de transcripció que regula l'expressió de nombrosos gens [13]. Per tant, aquesta interacció influeix en diversos processos biològics normals com el creixement, la proliferació, la diferenciació i la mort [14]. Fusco i Fedele [11] resumeixen en tres les diferents formes d'actuar de les proteïnes HMGA:

- Les proteïnes HMGA poden interaccionar directament amb el DNA, modificant la seva conformació i conseqüentment facilitant la unió a un determinat grup de factors de transcripció (v. fig. 1.5 a). Les HMGs interactuen així, tant amb el DNA com amb els factors de transcripció, generant un complex multiproteic. Es coneix que les proteïnes HMGA interactuen d'aquesta manera per a la regulació de molts gens, el cas més estudiat és l'expressió del gen interferó β (IFN- β). L'activació del gen IFN- β es produeix per la unió de diversos factors, on l'HMGA s'uneix doblegant el DNA, i facilitant el contacte d'altres factors (com IRF, ATF, cJun i p50) per activar l'RNA-polimerasa II.
- Les proteïnes HMGA també poden influir en el procés de transcripció a través d'interaccions proteïna-proteïna amb els factors de transcripció modificant la seva conformació i augmentant l'afinitat d'aquests per a interactuar amb el DNA (v. fig. 1.5 b). L'augment de l'activitat de transcripció del factor de resposta al sèrum (SRF) degut a l'acció de l'HMGA1a és un exemple d'aquest mecanisme.
- Per últim, de forma general, les proteïnes HMGA tenen l'habilitat d'alterar l'estructura de la cromatina, unint-se a ella i desplaçant les histones H1. D'aquesta manera activen la transcripció de la cromatina ja que aquesta està descompactada (v. fig. 1.5 c).

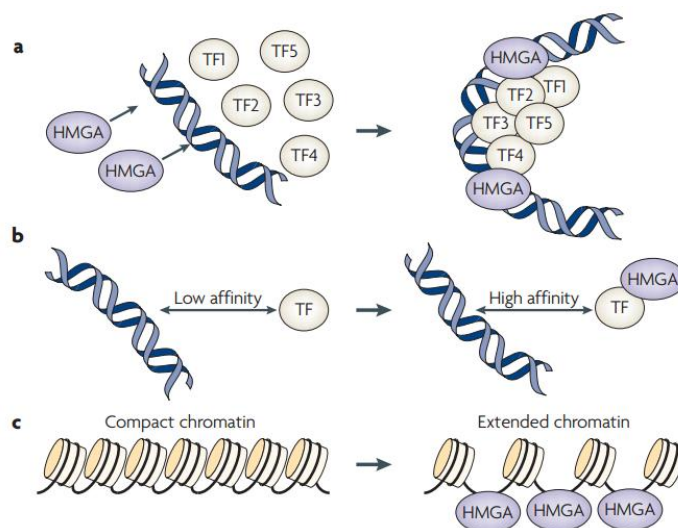


Fig. 1.5. Esquema de les tres vies d'actuació de les proteïnes HMGA.

Imatge extreta de [11]. Veure més informació al text.

Estructura

Les proteïnes HMGA contenen els motius característics *AT-hook*, que s'uneixen al solc menor del DNA preferentment en seqüències riques en AT, i una cua acídica C-terminal (v. la fig. 1.6). Exceptuant l'HMGA1c, les proteïnes HMGA posseeixen tres motius *AT-hook* bàsics i molt conservats, amb una seqüència consens (P)RGRP (menys en el primer *AT-hook* on la prolina és una glicina) flanquejats per fragments flexibles d'aminoàcids. Les proteïnes HMGA en dissolució no posseeixen ni estructura secundària ni terciària definides, presentant una estructura intrínsecament flexible i desestructurada. Se sap, però, que l'element central dels *AT-hook* s'estructura a l'interaccionar amb el DNA induint canvis estructurals sobre aquest. Cada cop es creu més important la rellevància biològica que juguen en la seva activitat funcional les regions desordenades de les proteïnes i les transicions entre una estructura desordenada i una ordenada [12]. És per això que en els últims anys s'està començant a estudiar amb més profunditat aquestes regions desestructurades en les que s'han anomenat proteïnes intrínsecament desordenades (IDP) [15][16].



Fig. 1.6. Seqüències aminoacídiques de les proteïnes HMGA humanes [17].

Els tres dominis *AT-hook* es representen en negreta i tenen càrrega positiva (es ressaltava la zona invariable R-G-R-P). La regió acídica C-terminal acídica apareix ressaltada en vermell. Els asteriscs indiquen la deleció d'11 aminoàcids que presenta l'HMGA1b respecte l'HMGA1a.

Interacció amb el DNA

La seqüència aminoacídica en la zona d'interacció de les proteïnes HMGA amb el DNA és sempre (P)RGRP on la primera prolina pot ser substituïda per una glicina. Estudis d'RMN han permès determinar que el fragment RGR (arginina-glicina-arginina) té una estructura totalment estirada que s'adapta perfectament al solc menor del DNA en seqüències riques en AT [18][19]. Durant la realització d'aquesta tesi ha estat resolta per un membre del nostre grup (MACROM) la primera estructura d'un cristall d'un *AT-hook*, concretament de l'*AT-hook* III en complex amb l'oligonucleòtid CGAATTAATTCG [20]. Aquesta estructura ens ha permès veure amb més detall la interacció d'un motiu *AT-hook* al solc menor d'un DNA ric en ATs. A la fig. 1.7 observem els models de tots dos complexos.

Els estudis d'RMN amb un fragment de la proteïna HMGA i l'oligonucleòtid GGGAAATTCCTC van donar una idea de les diferències en la unió al DNA entre l'*AT-hook* II i el III [19]. La interacció entre l'*AT-hook* i el segment AT pot ser analitzada en tres trams. En primer lloc un nucli RGR que adopta una conformació estesa en la profunditat del solc menor, un parell lisina/arginina a cada extrem del nucli que fan contactes electrostàtics i hidrofòbics amb l'esquelet sucre-fosfat del DNA i finalment (només en el cas de l'*AT-hook* II i no per al III) sis aminoàcids a l'extrem C-terminal del nucli (GSKNKGAA) que interaccionen també amb l'esquelet. Es creu que justament per aquestes últimes interaccions l'*AT-hook* II presenta major afinitat que el III pel DNA [19].

L'estructura de raigs X de l'*AT-hook* III amb un DNA ric en ATs va permetre per una banda veure com aquest modifica la forma B del DNA eixamplant el solc menor en la zona d'unió i provocant un apreciable doblegament local del DNA (uns 20° entre les 5 bases superiors i les 7 inferiors) [20]. D'altra banda s'ha confirmat que la seqüència AATT és un lloc d'unió preferent per als motius *AT-hook* de les proteïnes HMGA. També es van observar interaccions entre el pèptid i el segment AATT, formant-se ponts d'hidrogen entre els grups NH de les cadenes principals de la Gly37* i l'Arg38* amb els oxígens de tres timines, dos d'una cadena i un de la cadena complementària (*numeració de [20]).

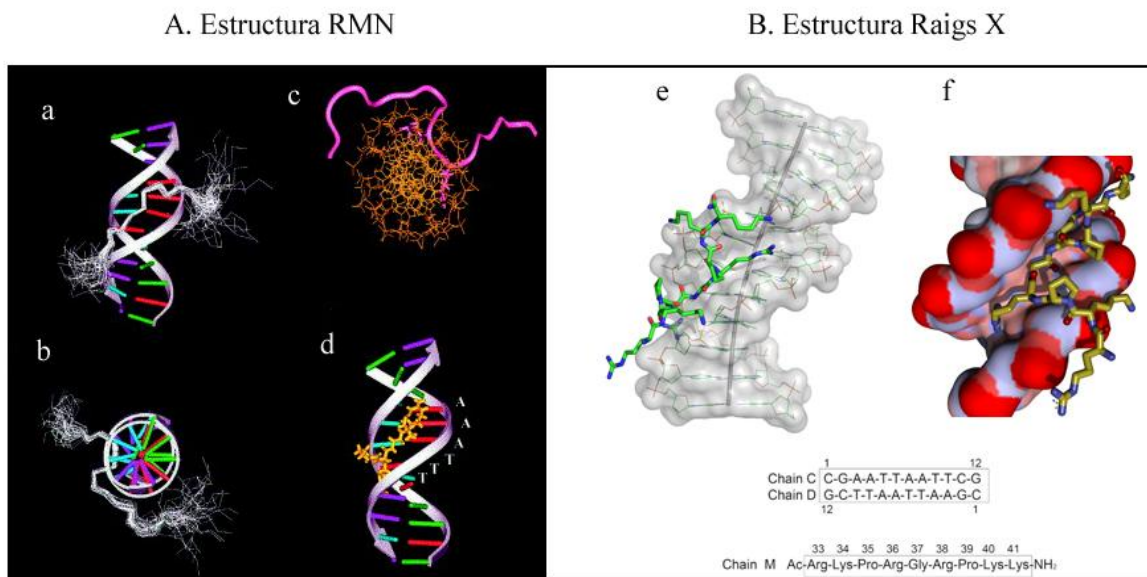


Fig. 1.7. Interacció dels *AT-hooks* amb el DNA (estructures de RMN i de raigs X).

A. Estructures del segon i el tercer *AT-hook* de l'HMGA1a unides a l'oligonucleòtid GGGAAATTCCTC resoltes per RMN. a) Vista lateral del complex del segon *AT-hook* amb el DNA mostrant la superposició del conjunt de 35 models per a l'esquelet del pèptid. b) Vista superior de (a). c) Vista superior del complex mostrant les cadenes laterals dels residus d'arginina de l'*AT-hook* d) Vista lateral del model del tercer *AT-hook* unit al solc menor [12]. B. Estructura del tercer *AT-hook* de l'HMGA1a amb l'oligonucleòtid CGAATTAATTTCG resolta per cristal·lografia de raigs X. e) Vista lateral del complex format per l'*AT-hook* III i l'oligonucleòtid, es mostra l'eix apreciablement doblegat del dúplex de DNA. f) Detall del posicionament d'aquest *AT-hook* dins del solc menor del DNA [20].

HMGA's i càncer

Les proteïnes HMGA participen en nombrosos processos biològics degut a la seva capacitat de reconèixer i alterar l'estructura del DNA i la cromatina. Conseqüentment, les cèl·lules que expressen alts nivells de proteïnes HMGA (com les cèl·lules tumorals) tenen una alteració profunda en l'estructura de la cromatina.

El nivell d'expressió dels gens de la proteïna HMGA és màxim durant el desenvolupament embrionari i les ràpides proliferacions cel·lulars, però és menor (o indetectable) en les cèl·lules adultes i els teixits diferenciats o que no es divideixen [3] (v. fig. 1.2). Per tant, basant-se en les dades disponibles, es creu que les proteïnes HMGA estan directament implicades en la regulació de la proliferació cel·lular, el creixement embrionari i les cèl·lules amb funció mesenquimal.

Després de l'embriogènesi, l'HMGA és reexpressada en cèl·lules transformades i tumorals. Així, en les cèl·lules neoplàsiques transformades el nivell de proteïna HMGA és sovint excepcionalment alt amb una correlació entre l'increment de la concentració i el grau de malignitat o potencial metastàtic [3][11][21]. Aquesta correlació és tan consistent i estesa que ha estat suggerit que l'elevada concentració d'HMGA sigui un marcador de diagnòstic de les transformacions neoplàsiques [22][23]. Hi ha, gairebé, una perfecta correlació entre la sobreexpressió de les proteïnes HMGA i la progressió del tumor en cada tipus de càncer que ha estat estudiat (colorectal, de pròstata, de pit, cervical i pulmonar) [21].

Els factors que regulen l'expressió dels gens HMGA són variats. Existeixen diferents senyals de transducció i una varietat de factors de transcripció que participen en el control de l'expressió *in vivo* de la proteïna HMGA [12].

HMGA's i altres malalties

L'alteració en la regulació de l'expressió de les proteïnes HMGA pot estar relacionada amb d'altres malalties a més a més del càncer [17]. Per exemple, se sap que les proteïnes HMGA contribueixen a la regulació de nombrosos gens relacionats amb la resposta inflamatòria. L'HMGA regula gens relacionats amb la proliferació, la diferenciació, la presentació d'antígens, i el reclutament de cèl·lules inflamatòries com per exemple el gen de la immunoglobulina E (IgE) o de la interleucina 4 (IL-4) [17]. D'altra banda també es va descobrir que malalts de diabetis tipus 2 presentaven nivells baixos d'expressió d'HMGA que afectaven negativament l'expressió del receptor d'insulina i la funció de l'insulina en cèl·lules diana i teixits.

1.1.2. Les proteïnes HMGB

Família de les HMGB

Existeixen tres gens que codifiquen tres variants d'HMGB canòniques humanes que presenten un 80% d'identitat entre elles: HMGB1 (215 aa), HMGB2 (209 aa) i HMGB3 (200 aa) (fig. 1.8). Varien bàsicament en la longitud de les seves cues acídiques C-terminals. En adults, l'HMGB1 s'expressa a tot arreu menys en les neurones, en canvi l'HMGB2 i l'HMGB3 s'expressen totes dues únicament als òrgans limfoïdes i als testicles. Els embrions expressen alts nivells de totes tres mentre que les cèl·lules tumorals presenten sobreexpressió de l'HMGB1 [2].

1	MGKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKCKSERWKTMSAKEKGF	60	P09429	
1	MGKGDPNKPRGKMSSYAFFVQTCREEHKKKHPDSSVNFAEFSKCKSERWKTMSAKEKSKF	60	P26583	
1	MAKGDPKKPKGMSAYAFFVQTCREEHKKKNPEVFPVNFAEFSKCKSERWKTMSGKEKSKF	60	O15347	
	* . : * * * * * : * * * * * : * * * * * * * * * * * * * * * * * * : * * * * * : *			
61	EDMAKADKARYEREMKTYIPPKGETKKKFKDPNAPKRPPSAFFLFCSEYRPKIKGEHPGL	120	P09429	
61	EDMAKSDKARYDREMKNYVPPKGDKKKGGKDPNAPKRPPSAFFLFCSEHRPKIKSEHPGL	120	P26583	
61	DEMAKADKVRYDREMKDYGPAKGGK--KKDPNAPKRPPSGFFLFCSEFRPKIKSTNPGI	118	O15347	
	: : * * * * * : *			
121	SIGDVAKKLGEMWNNTAADDKQPYEKKAALKKEKYEKDIAAYRAKPKDAAKKGVVKAEK	180	P09429	
121	SIGDTAKKLGEMWSEQSAKDKQPYEQKAALKKEKYEKDIAAYRAKPKSEAGKKGPRPTG	180	P26583	
119	SIGDVAKKLGEMWNNLNDSEKQPYITKAALKKEKYEKDVADYKSKGKFDGAK-GPAKVAR	177	O15347	
	* * * * * . *			
181	SKKKKEEEEEDEEEDDEEEDEDEEDEDDEEDEDDEE	215	P09429	HMGB1_HUMAN
181	SKKKNEPEDEEEDEEEDDEEEDDEEEDEDEE-----	209	P26583	HMGB2_HUMAN
178	KKV--E-EEDEEEDEEEDEEEDEE-----	200	O15347	HMGB3_HUMAN
	. *			

Fig. 1.8. Seqüències aminoacídiques de les tres HMGBs humanes.

Alineament de les seqüències aminoacídiques de les proteïnes humanes HMGB1, HMGB2 i HMGB3.

Figura preparada amb l'eina d'alineament de l'UniProt. (*totalment conservat, : conservació d'aminoàcids amb grups amb propietats molt similars, . conservació d'aminoàcids amb grups amb propietats dèbilment similars).

D'altres proteïnes no canòniques presenten també el motiu HMG-box, generalment són factors de transcripció. En totes les proteïnes amb domini HMG-box l'estructura general d'aquest és força més conservada que la seva seqüència aminoacídica. Normalment s'adopta una classificació basant-se en l'especificitat de seqüència per a estudiar les diferents proteïnes amb motius HMG-box (v. taula 1.3). En el primer grup, trobem les proteïnes HMG-box que presenten especificitat de seqüència d'unió al DNA limitada o inexistent, podent presentar dos HMG-boxes com en el cas de les HMGB de mamífer, un únic HMG-box com l'NHP6A (HMGB de llevat) o fins i tot més de dos com el factor UBF. El segon grup és molt divers i consisteix en proteïnes molt menys abundants que presenten un únic HMG-box que reconeix específicament seqüències de DNA (dins de la limitació de formació de punts d'hidrògens específics en el solc menor) [24][25].

Taula 1.3. Classificació de les proteïnes amb motius HMG-box segons la seva especificitat d'unió al DNA.

Proteïnes amb motius HMG-box segons la seva especificitat	
Amb especificitat limitada o inexistent	Amb especificitat de seqüència
<p>HMGB1-3 de mamífer (2 HMG-boxes)</p> <p>Factors de transcripció del mitocondri: TFAM1 i ABF2 (2 HMG-boxes)</p> <p>Drosophila Dorsal Switch protein I (DSP1) (2 HMG-boxes)</p> <p>Factor de transcripció nucleolar UBF (4-6 HMG-boxes)</p> <p>HMG1A/B d'insecte (1 HMG-box)</p> <p>HMGD/Z de <i>Drosophila</i> (1 HMG-box)</p> <p>NHP6A/B de <i>S. cerevisiae</i> (1 HMG-box)</p>	<p>Factors de transcripció LEF1 i TCF (1 HMG-box)</p> <p>Factor TOX (1 HMG-box)</p> <p>Factor de determinació de sexe SRY (1 HMG-box)</p> <p>Factors SOX (1 HMG-box)</p>

Es mostren alguns exemples de proteïna o família de proteïnes classificades segons la seva especificitat d'unió al DNA i s'indiquen el nombre d'HMG-boxes que presenten cadascuna entre parèntesi [26].

En aquest treball hem estudiat fragments de la proteïna HMGB1 de rata, que presenta la mateixa seqüència aminoacídica que la humana exceptuant dos residus de la cua acídica. També s'ha volgut estudiar l'estructura de la proteïna NHP6A del *Saccharomyces cerevisiae*.

Funció

Les proteïnes HMGB tenen dos rols diferents: un dins del nucli i un altre extracel·lular (v. fig. 1.9). Dins del nucli aquestes proteïnes actuen com un factor d'unió a la cromatina produint alteracions estructurals en el DNA i interactuant amb d'altres proteïnes per a facilitar diversos processos cel·lulars. D'aquesta manera participen en la regulació de la transcripció, la recombinació i la reparació del DNA i a més estan implicades en el procés de la transposició [25][27]. A més d'aquest rol nuclear, l'HMGB1 també actua com una molècula de senyal extracel·lular en la inflamació, en la diferenciació cel·lular, en la migració cel·lular i en la metastasi de tumors. L'HMGB1 s'allibera en cèl·lules necròtiques i en cèl·lules inflamatòries, unint-se amb alta afinitat a alguns receptors (com RAGE, *receptor for advanced glycation end products*), mitjançant d'aquesta manera el senyal d'infecció i promovent per tant la inflamació [28][29][30].

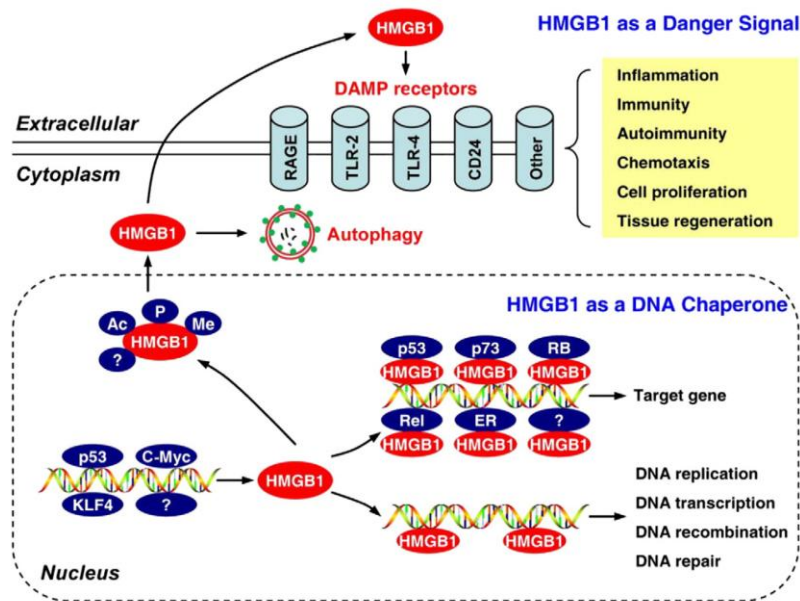


Fig. 1.9. Funcions de l'HMGB1 dins i fora de la cèl·lula [28].

L'HMGB actua dins del nucli com a un factor arquitectònic d'unió a la cromatina (també anomenada xaperona del DNA per aquest motiu) participant en nombrosos processos com la replicació, la recombinació, la transcripció i la reparació del DNA. En el seu rol extracel·lular l'HMGB1 és un senyal de perill en condicions d'inflamació, mitjançant la resposta a la infecció, la immunitat, l'autoimmunitat, la quimiotaxi, la proliferació cel·lular i la regeneració de teixits.

En la seva funció nuclear d'interacció amb la cromatina totes les proteïnes HMGB són capaces de doblegar el DNA. Es creu que aquesta pot ser la seva funció principal en el seu rol nuclear. Tant les proteïnes de seqüència-específica com les d'inexistent especificitat de seqüència poden doblegar el DNA de 60° a 110°. D'aquesta manera aquestes proteïnes podrien promoure la unió d'altres proteïnes i/o la transcripció de nombrosos gens de tres maneres [2]:

- L'HMGB pot interactuar directament amb el nucleosoma, desenrotllant el DNA i facilitant d'aquesta manera l'accessibilitat d'altres complexos moduladors de l'estructura o factors de transcripció (v. fig. 1.10 a).
- L'HMGB1 podria doblegar el DNA promotor, per incrementar l'afinitat d'unió de factors de transcripció com la TBP (*TATA-binding protein*) (v. fig. 1.10 b).
- Les HMGB1 poden establir interaccions proteïna-proteïna amb factors de transcripció específics. Un cop unides, doblegarien el DNA, estabilitzant la unió de la proteïna aparellada i/o promovent el reclutament d'altres proteïnes que interactuarien també (v. fig. 1.10 c).

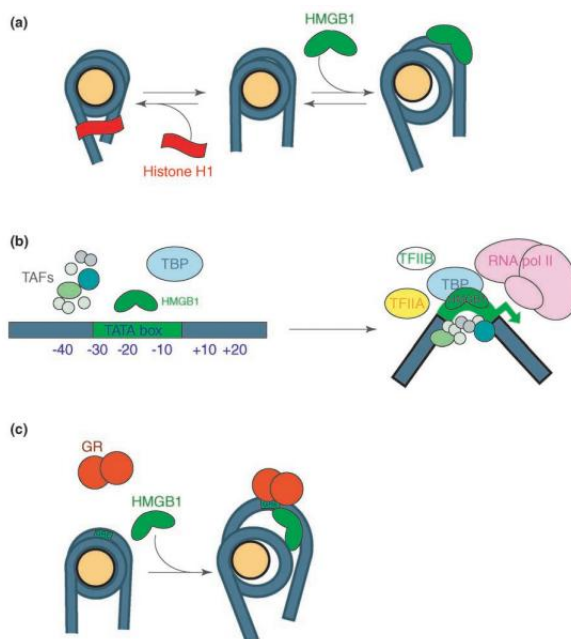


Fig. 1.10. Esquema dels tres mecanismes d'actuació de les proteïnes HMGB.

Imatge extreta de [2]. Veure més informació al text.

A més a més, les HMGBs faciliten la recombinació V(D)J interaccionant amb proteïnes relacionades amb la recombinació de gens i incrementat les ràtios d'escissió. També tenen un rol en la reparació del DNA i estan involucrades en el reconeixement d'estructures del DNA distorsionades o modificades i danyades.

Estructura

Les proteïnes HMGB1-3 de mamífer presenten dos HMG-boxes (anomenats Box A i Box B) i una cua àcida C-terminal (v. fig. 1.11) (revisat a [24][25][31]). Les estructures d'aquests HMG-boxes individuals van ser resoltes per estudis d'RMN [32][33]. El plegament global dels dominis HMG-box, d'uns 75 aminoàcids, està ben conservat i consisteix en tres hèlixs alfa que formen una L. El braç llarg d'aquesta L consisteix en la cadena estesa N-terminal i l'hèlix III, mentre que el braç curt està format per les hèlixs I i II, amb un angle de 80° entre ambdós braços.

El Box A i el Box B de l'HMGB1 de mamífer difereixen entre ells lleugerament en les geometries de l'hèlix I i II respecte la resta de la molècula. El *loop* entre les hèlix I i II és més llarg per a la Box A així com l'hèlix I és més curta i estirada (estant l'hèlix I lleugerament doblegada per al Box B). A més a més el Box A està carregat més positivament en les hèlix I i II que el Box B. L'estructura del Box B és molt semblant als dominis HMG-box de l'NHP6A (*S. cerevisiae*) i HMGD (*Drosophila*) i en menor grau a l'HMG-box de les proteïnes amb especificitat de seqüència SRY i LEF1, sent la Box A més diferent a totes.

Tots dos dominis s'uneixen al solc menor del DNA, però el Box A té una major afinitat per les estructures distorsionades del DNA, com el DNA modificat amb cisplatí o el *four-way junction* (unió de quatre vies). A més el Box B doblega més efectivament el DNA que el Box A [34].

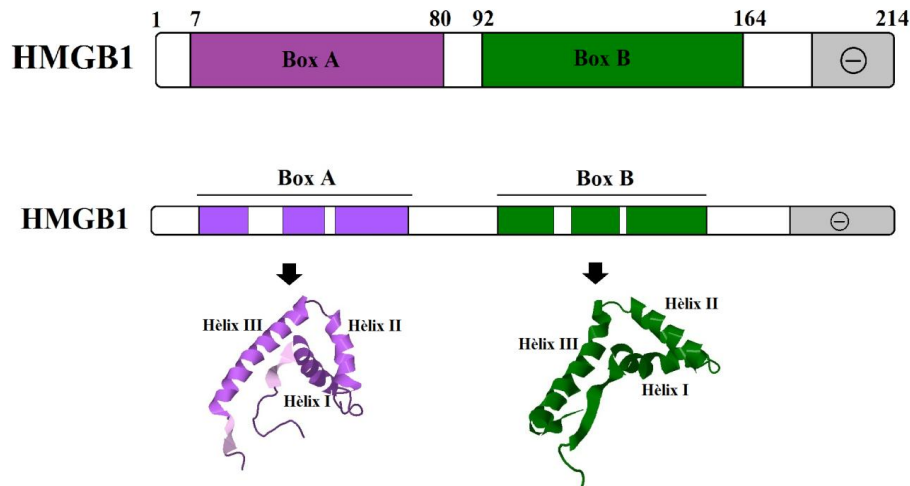


Fig. 1.11. Esquema de l'estructura de la proteïna humana l'HMGB1.

L'HMGB1 presenta dos dominis HMG-box, el Box A (lila) i el Box B (verd), i una cua àcida C-terminal. Les estructures representades corresponen a les estructures individuals dels dos HMG-boxes resoltes per RMN [32][33].

Interacció amb el DNA

HMGB1 i HMGB2 mostren una clara preferència per la unió a DNA no canònic, com per exemple una única cadena de DNA, DNA que conté cruciforms o està doblegat, DNA superenrotllat i Z-DNA [35]. Tot i que els HMG-boxes (A i B) d'aquestes proteïnes poden unir-se individualment al DNA lineal (Box A més fortament que Box B [36]), l'afinitat d'unió millora quan es troben tots dos dominis units. A més a més les proteïnes HMG-boxes tenen com a característica comuna que s'uneixen al solc menor i poden doblegar el DNA lineal de 60° a 110° cap al solc major [37]. Aquesta distorsió tan dramàtica del DNA depèn de la intercalació amb el DNA i la complementarietat de la seva forma [38][39][40].

La majoria de les proteïnes amb especificitat de reconeixement de seqüència limitat o inexistent (HMGB1-3, HMGBD, NHP6A, etc.) presenten dos aminoàcids no polars que s'intercalen al DNA en cadascun dels seus HMG-boxes. El residu corresponent al primer lloc d'intercalació es troba a l'hèlix I (1r en fig. 1.12) mentre que el corresponent al segon es situa al principi de l'hèlix II (2n en fig. 1.12). La intercalació d'aquests dos residus voluminosos i apolars a una distància d'un o dos parells de bases provoca el doblegament del DNA característic de les proteïnes HMG-box. En el cas del Box A de l'HMGB1, però, no es produeix una intercalació en el primer lloc ja que la cadena lateral de l'alanina no arriba a intercalar-se. En el cas de les proteïnes amb especificitat de seqüència d'unió al DNA (SRY i mLEF1 a la fig. 1.12) hi ha un aminoàcid polar al segon lloc d'intercalació que fa contactes amb el DNA però no s'intercala [41][42].

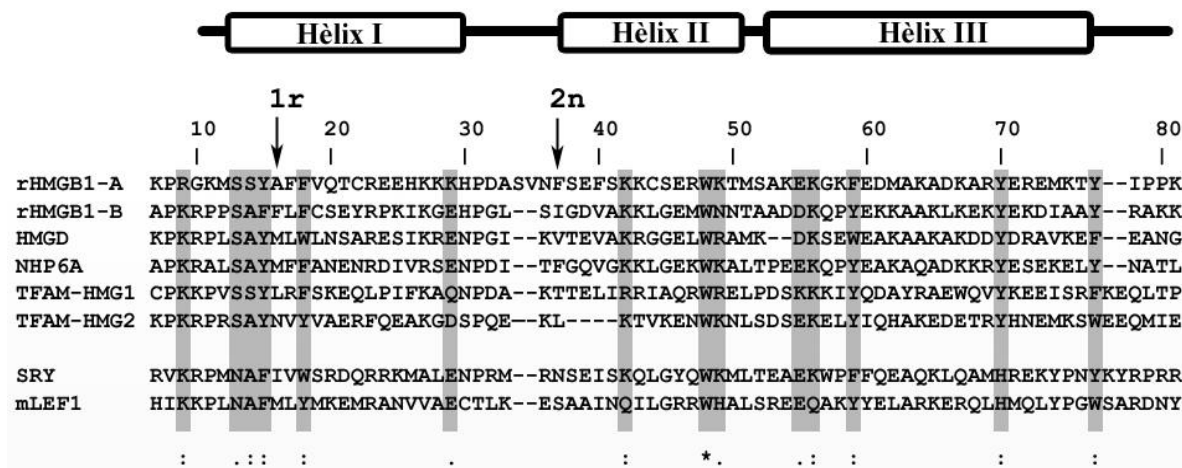


Fig. 1.12. Seqüències aminoacídiques dels dominis HMG-box d'algunes proteïnes.

Alineament de les seqüències aminoacídiques dels dominis HMG-box de les proteïnes HMGB1 (rata) [rHMGB1-A, Box A; -B, Box B], HMGD (*Drosophila*), NHP6A (*S. cerevisiae*), TFAM (mitochondri humà) [TFAM-HMG1, Box A; -HMG2, Box B], SRY (humà), LEF1 (ratolí). Les fletxes indiquen els residus corresponents al primer (1r) i al segon (2n) lloc d'intercalació. Es mostren ombrejats els aminoàcids conservats (*totalment conservat, : conservació d'aminoàcids amb grups amb propietats molt similars, · conservació d'aminoàcids amb grups amb propietats dèbilment similars) (Cluster Omega).

L'estructura de la proteïna HMGB1 amb els seus dos dominis HMG-box units al DNA no ha estat resolta fins ara. Per una banda s'ha vist l'estructura de raigs X del domini Box A unit a un DNA prèviament distorsionat per la unió amb cisplatí (v. fig. 1.13 a) [43]. En aquest cas el DNA presentava un doblegament de $\sim 60^\circ$. D'altra banda, Stott i col·laboradors van produir un mutant consistent en la proteïna de seqüència específica SRY i el Box B de l'HMGB1 unides per un linker (quimera SRY-Box B) i es va estudiar la seva interacció amb un DNA de 16 pb per RMN (v. fig. 1.13 b) [44]. Tots dos HMG-box (SRY i Box B) introduïen un doblegament global al DNA de $\sim 101^\circ$ ($\sim 90^\circ$ cadascun localment). Totes dues estructures van aportar informació de la unió d'aquestes proteïnes amb el DNA i els seus mecanismes de reconeixement. Però resta encara pendent conèixer els canvis que pot induir el Box A en un DNA lineal i les interaccions específiques entre tots dos. A més a més seria molt interessant poder arribar a trobar l'estructura de tots dos dominis interactuant junts per a doblegar un DNA.

El cristall d'un complex amb la proteïna HMGD (v. fig. 1.13 c) [45] i l'estructura de RMN de l'NHP6A unida a un DNA de 15 pb (v. fig. 1.13 d) [46] també han estat resolts. En aquests casos l'angle de doblegament ha estat de $\sim 90^\circ$ i $\sim 110^\circ$ aproximadament. En totes dos casos contribueixen a distorsionar el DNA la intercalació en dos llocs: Met13 en el primer lloc i Val32 i Thr33 en el segon per a l'HMGD, i Tyr28* Met29* en el primer lloc i Phe48* en el segon per a l'NHP6A (*numeració de [46]).

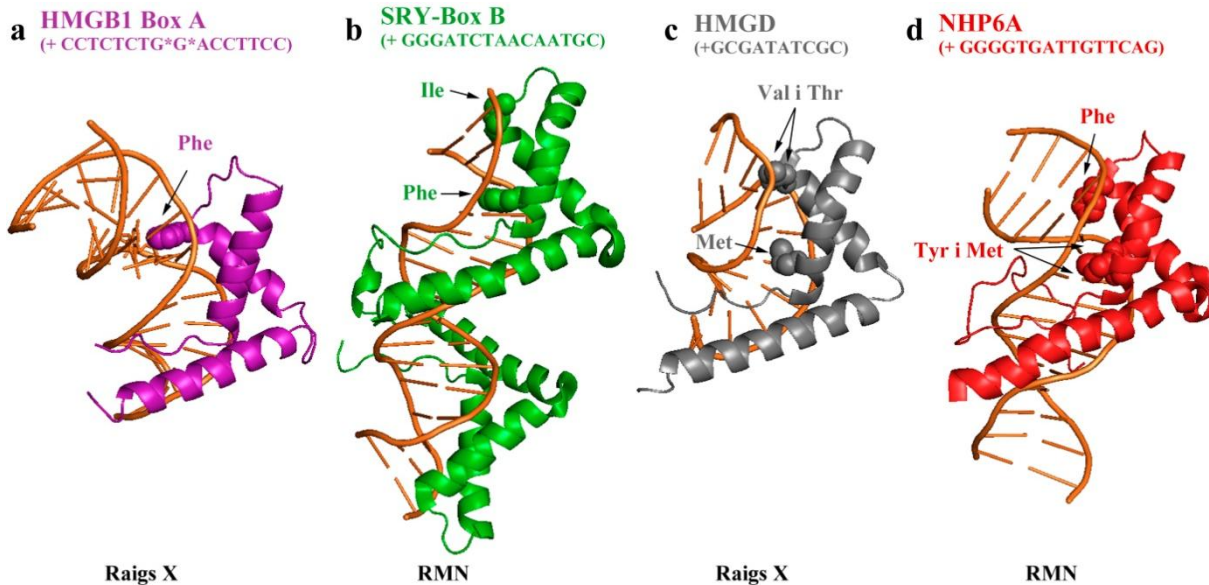


Fig. 1.13. Interaccions proteïna-DNA de les proteïnes HMGB1 (Box A i Box B), HMGD i NHP6A.

Figura realitzada amb els models del PDB corresponents a aquestes proteïnes (a, Box A (1CKT); b, Quimera SRY-Box B (2GZK); c, HMGD (1QRV); d, NHP6A (1J5N)). En la figura s'indica el nom de la proteïna i la seqüència de bases dels oligonucleòtids del corresponent complex. Es ressalten (representats com a esferes) els aminoàcids que participen en la interacció amb el DNA en els llocs d'intercalació 1r i 2n.

L'any 2011 es van resoldre paral·lelament per cristal·lografia de raigs X dues estructures corresponents a l'HMG-box del mitocondri humà, el TFAM [47][48]. Tot i presentar diferències en les seqüències aminoacídiques respecte la proteïna HMGB1 (v. fig. 1.12), el TFAM presenta també dos dominis HMG-box (Box A i Box B). Cada domini s'intercala al DNA provocant un doblegament de 90° i globalment el DNA fa una volta en forma d'U de 180° (v. fig. 1.14).

Aquest tàndem d'HMG-boxes del TFAM presenta una unió cua-cua, mentre que a la quimera SRY-Box B interaccionaven cap-cap (comparar fig. 1.14 i fig. 1.13 b). En el Box A del TFAM es va veure que en el primer lloc d'intercalació es trobava una leucina (apolar) mentre que en el segon una treonina (polar) que fa ponts d'hidrogen amb el DNA, i per tant correspondria a la classificació d'HMG-box amb especificitat de seqüència. No obstant això, en el Box B s'intercalava la leucina del segon lloc d'intercalació i feia ponts d'hidrogen una asparagina (polar) corresponent al primer lloc. Per tant es va veure que el TFAM no seguia la classificació de l'especificitat de seqüència de les HMG-box (v. taula 1.3) mostrant un altre patró de reconeixement (revisat a [31]).

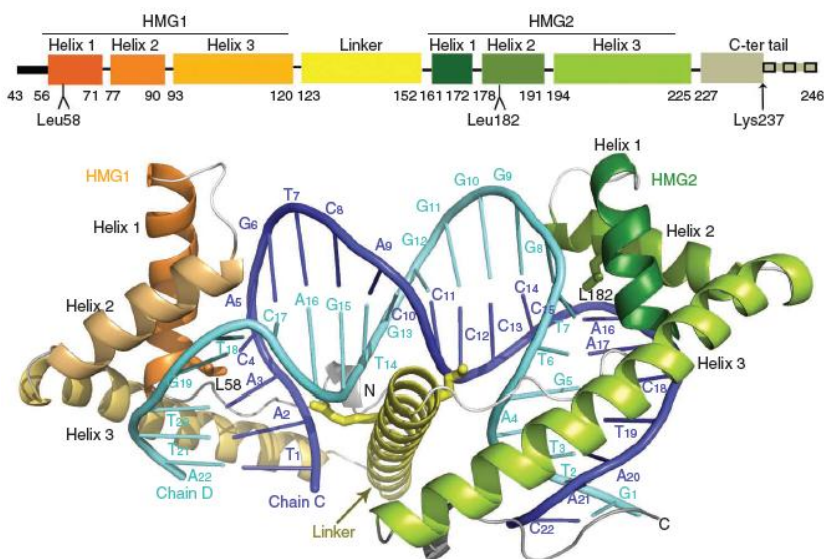


Fig. 1.14. Model de l'estructura del TFAM (HMGB del mitocondri).

Imatge extreta de [48].

HMGB, càncer i inflamació

La proteïna HMGB1 és tant un factor nuclear com una proteïna secretada. Al nucli de la cèl·lula actua com un factor arquitectònic que s'uneix al DNA i promou l'associació d'aquest amb d'altres proteïnes. Fora de la cèl·lula, s'uneix amb molta afinitat a RAGE i és un potent mediador de la inflamació. L'HMGB1 també juga un rol molt important en el desenvolupament del càncer i la metastasi, amb la senyal RAGE-HMGB1 promovent-se arreu en la majoria dels diferents tipus de tumors [28].

a) HMGB i càncer

Les proteïnes HMGB1 són expressades tant en cèl·lules normals com en cèl·lules canceroses. Però la seva sobreexpressió així com la seva localització al citoplasma ha estat observada en la majoria de les cèl·lules tumorals, incloent tumors de l'estroma gastrointestinal [49], carcinomes de colon [50] i carcinomes cervicals [51]. A més a més l'HMGB1 també interactua amb i afavoreix les activitats d'alguns factors de transcripció implicats en el desenvolupament de càncer, com p53, p73, la proteïna del retinoblastoma entre d'altres (v. fig. 1.9) (revisat a [28]). La localització de la proteïna HMGB1 al medi extracel·lular indica la presència de cèl·lules estressades o que han mort i alerta a d'altres cèl·lules mitjançant senyals de perill. Per tant les alteracions en la localització de l'HMGB1, la seva abundància i nivells d'expressió estan associades amb tots els distintius del càncer.

b) HMGB i inflamació

Recentment es va descobrir que no només els agents patògens estranys i els seus productes poden activar el sistema immunològic innat, sinó que també un grup de molècules endògenes desenvolupaven una funció similar. Entre aquestes molècules endògenes (anomenades alarmines) es troba l'HMGB1 [29]. Les alarmines són constituents normals de les cèl·lules però són alliberades al medi extracel·lular en estats d'estrès o dany cel·lular, activant d'aquesta manera el sistema immunològic. Un cop alliberada a l'exterior de les cèl·lules, l'HMGB1 actua com un senyal que avisa a diferents receptors estant així involucrada en condicions d'inflamació tant infecciosos com no infecciosos. Entre aquests rols en models d'inflamació no infecciosos es troben processos com autoimmunitat, càncer, trauma i isquèmia.

1.2. L'estructura del DNA

Watson i Crick [52], basant-se en anàlisis químiques i estudis de difracció de raigs X, van proposar un model de DNA que consisteix en dues cadenes de nucleòtids enrotllades l'una sobre l'altra, formant una hèlix. Aquest model d'estructura de doble hèlix representa un dels descobriments científics més importants del segle XX i la base de la biologia molecular moderna. No només el model satisfieix les propietats físico-químiques conegudes del DNA sinó que també donava una idea de com aquest duia a terme les seves funcions biològiques. Per exemple, la complementarietat de les dues cadenes de la doble hèlix va proporcionar un mecanisme de com podia ser copiat. D'altra banda la seqüència definida de bases al llarg de la cadena de DNA possibilitava la codificació de la informació necessària per a la síntesi d'una determinada proteïna.

En aquest apartat es consideraran les característiques més importants del DNA, tot i fent èmfasi en la seva diversitat estructural.

1.2.1. Estructura química

El DNA és una macromolècula formada per la unió de nucleòtids amb un enllaç fosfodièster. Un nucleòtid és una molècula constituïda per la unió d'una base nitrogenada, un sucre (concretament la 2'-desoxiribosa) i un fosfat. Les bases nitrogenades són anells aromàtics de dos tipus: purines i pirimidines. Al DNA les purines són la guanina (G) i l'adenina (A) i les pirimidines són la timina (T) i la citosina (C).

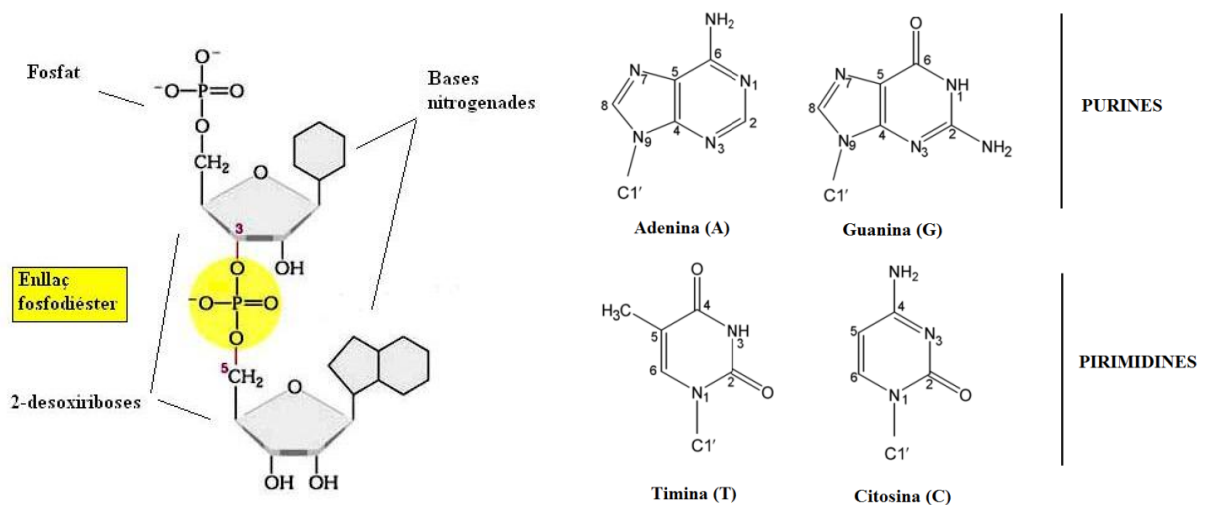


Fig. 1.15. Enllaç fosfodièster entre dos nucleòtids (esquerra) i bases nitrogenades del DNA (dreta).

Podem observar com es forma l'enllaç entre el carboni 3' d'un sucre i el carboni 5' de l'altre.

El model de Watson i Crick [52] consisteix en una doble hèlix dextrogira on les bases nitrogenades se situen a l'interior de l'hèlix i les pentoses i els fosfats formen l'esquelet de la molècula (fig. 1.16). Les cadenes de DNA són complementàries. Sempre queden enfrontades una base nitrogenada purínica amb una pirimidínica formant ponts d'hidrogen entre elles, concretament l'adenina (A) amb la timina (T) amb dos ponts d'hidrogen, i la guanina (G) amb la citosina (C) amb tres (v. fig. 1.16). A més a més, les cadenes són antiparal·leles, les pentoses d'una queden en posició inversa respecte a l'altra. La forma B, que és la que es troba normalment als éssers vius, té 10,4 parells de bases per volta, cada volta té una longitud de 3,4 nm (pas d'hèlix) i el diàmetre de l'hèlix és de 2,0 nm.

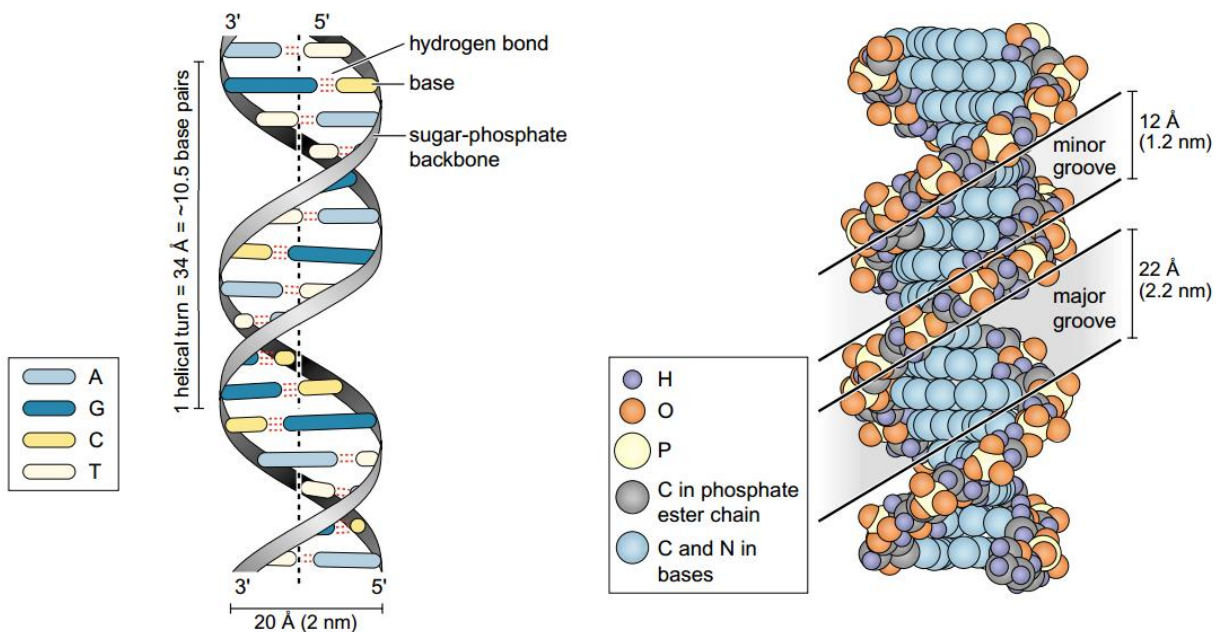


Fig. 1.16. Esquema de l'estructura secundària del DNA.

Observem l'esquelet sucre-fosfat que forma l'hèlix i com les bases nitrogenades queden al seu interior aparellades A-T i C-G. També apreciem el solc menor (*minor groove*) i major (*major groove*).

Imatges extretes de [53].

A part de la forma B del DNA existeixen altres formes, com la forma A i Z, que varien en els paràmetres conformationals de la doble hèlix: pas d'hèlix, separació entre residus, l'angle de gir per residu, el nombre de parells de bases per volta i el tipus d'enllaç glicosídic (veure fig. 1.17 i taula 1.4). Tot i que aquestes formes són les més estudiades s'han trobat per difracció de raigs X d'altres conformacions variants del B-DNA, com la B', la C, la D i la E [54].

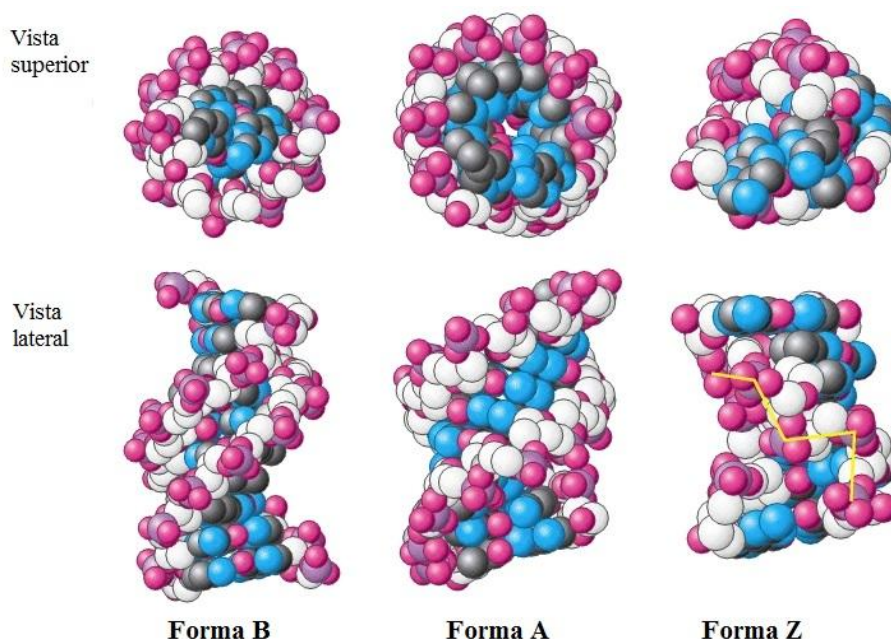


Fig. 1.17. Models de les formes B, A i Z del DNA.

Les formes B i A presenten hèlixs dextrogires i la Z levogira. L'hèlix de la forma B és més llarga i estreta que la forma A. En la forma Z els grups fosfats es col·loquen en forma de ziga-zaga. Imatge adaptada de [55].

La forma B, que s'observa a major humitat, és la forma que correspon a l'estructura promig del DNA en condicions fisiològiques. La forma A, observada en condicions de menor humitat (75% d'humitat relativa) [56], presenta uns solc major més estret i profund que la forma B mentre que el solc major és més ample i superficial. L'àmplia majoria del DNA a la cèl·lula es troba en forma B, però el DNA pot adoptar la forma A en alguns complexos DNA-proteïna. També, la forma A és similar a la forma que adopta l'RNA quan es troba en doble hèlix (A-RNA) [57].

Taula 1.4. Comparació entre les formes B, A i Z del DNA [55].

Tipus d'hèlix			
	Forma B	Forma A	Forma Z
Diàmetre de l'hèlix	23,7 Å	25,5 Å	18,4 Å
Distància entre cada pb	3,4 Å	2,3 Å	3,8 Å
Sentit de gir	dextrogir	dextrogir	levogir
Enllaç glicosídic	<i>anti</i>	<i>anti</i>	alternen <i>anti</i> i <i>syn</i>
Parells de bases per pas d'hèlix	10,4	11	12
Pas d'hèlix	35,4 Å	25,3 Å	45,6 Å
Inclinació de les bases respecte la perpendicular de l'eix de l'hèlix	1°	19°	9°
Solc major	ample i força profund	estret i molt profund	pla
Solc menor	estret i força profund	molt ample i superficial	molt estret i profund

Moltes de les diferències entre la forma B i A del DNA es deuen a diferents plegaments de les seves riboses (*sugar puckers*) com es mostra a la fig. 1.18. En el DNA A, el C-3' es troba fora del pla que formen els altres quatre àtoms de l'anell de la furanosa (una conformació anomenada C-3' *endo*). En el DNA B, el C-2' es troba fora d'aquest pla (una conformació anomenada C-2' *endo*). La conformació C-3' *endo* del DNA A provoca que els parells de bases tinguin una inclinació d'onze graus respecte del pla normal de l'hèlix. Existeixen unes altres quatre conformacions de la ribosa (C-2' *exo*, O-4' *exo*, C-4' *endo* i O-4' *endo*) que són conformacions intermèdies entre les dues conformacions preferents per a la forma A i B.

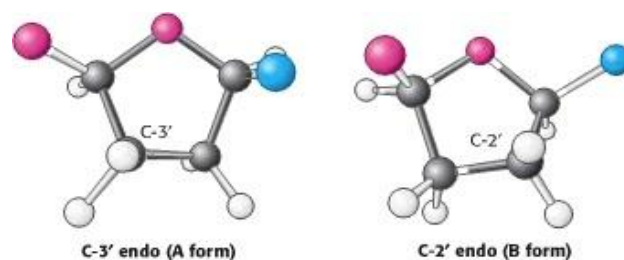


Fig. 1.18. Plegament del sucre (*sugar pucker*).

En el DNA A, el C-3' es troba fora del pla que formen els altres quatre àtoms de l'anell de la furanosa, en canvi en el DNA B, el C-2' es troba fora d'aquest pla. Imatge extreta de [55].

La forma Z del DNA va ser descrita l'any 1979 al resoldre's l'estructura de l'oligonucleòtid d(CGCGCG) [58]. El més sorprenent d'aquesta estructura va ser trobar que la doble hèlix era levogira. A més a més, els fosfats de l'esquelet es col·locaven en forma de ziga-zaga (per això va rebre el nom de forma Z). La forma Z del DNA l'adopten oligonucleòtids de seqüència curta que alternen pirimidines i purines a altes concentracions de sal. Presenta una conformació *anti* de l'enllaç glicosídic en les posicions de les pirimidines i conformació *syn* en les posicions de les purines, a diferència de la forma A i B que presenten tots els enllaços glicosídics en conformació *anti* (fig. 1.19). Són justament aquestes conformacions *syn* les responsables del gir levogir.

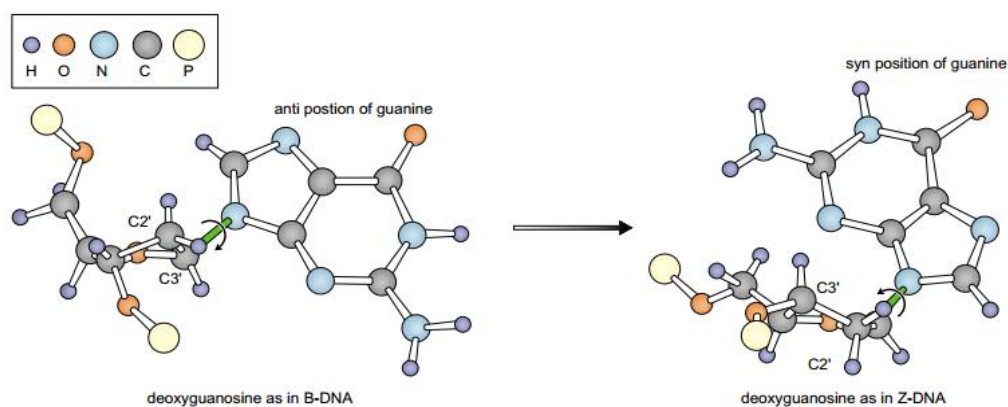


Fig. 1.19. Conformacions *syn* i *anti* de l'enllaç glicosídic en la forma B i Z del DNA.

Imatge extreta de [53].

1.2.2. Paràmetres geomètrics dels parells de bases

Cal tenir en compte que la forma B del DNA és una estructura promig mentre que el DNA real no és perfectament regular. Per exemple, les dues bases d'un parell de bases no es troben exactament en un mateix pla (variació en el *propeller*) i la rotació exacta entre parells de bases (*twist*) no és constant. Així les molècules de DNA no presenten una doble hèlix perfecta, la seva conformació exacta depèn dels parells de bases que es trobin en cada posició i els parells de bases veïns al llarg de la doble hèlix. Aquestes variacions properes al model ideal del B-DNA van ser analitzades a resolució atòmica usant la difracció de raigs X de cristalls únics del dodecàmer d(CGCGAATTCGCG) l'any 1980 [59].

D'aquesta manera, l'estudi de la morfologia de les bases ha esdevingut molt important per a descriure i entendre moltes característiques i deformacions dels àcids nucleics relacionades amb la seqüència. S'han proposat diferents paràmetres rotacionals i translacionals per a descriure aquestes relacions geomètriques entre bases i parells de bases, les quals van ser originàriament definides al 1989 [60]. Les definicions dels paràmetres geomètrics més característics són [61]:

Per a parells de bases individuals:

- El propulsor d'hèlix (*propeller twist* o *propeller*) (π) entre les bases és l'angle diedre entre normals a les bases quan es mira al llarg de l'eix llarg del parell de bases.
- El plegament (*buckle*) (κ) és l'angle diedre entre bases, al llarg del seu eix curt, després de posar el valor del *propeller twist* a 0°.
- La inclinació (*inclination*) (η) és l'angle entre l'eix llarg del parell de bases i un pla perpendicular a l'eix de l'hèlix.
- Els desplaçaments X i Y (*X, Y displacements*) (dx , dy) defineixen les translacions d'un parell de bases dins del seu pla promig en termes de la distància entre el punt mig del parell de bases al llarg de l'eix llarg des de l'eix de l'hèlix.

Pels passos entre parells de bases:

- El gir d'hèlix (*helical twist*) (Ω) és l'angle entre parells de bases successives, mesurat com el canvi en l'orientació dels vectors C1'-C1', anant d'una base a la següent i projectat en l'eix de l'hèlix.
- El balandreig (*roll*) (ρ) és l'angle diedre per a la rotació entre una base respecte a la seva veïna, sobre l'eix llarg del parell de bases. El *tilt* (τ) és el mateix angle però al llarg de l'eix curt.
- L'esllavissament (*slide*) (Dy) és el desplaçament relatiu entre dos parells de bases, mesurat entre els punts mitjos de cada C6-C8 de cada parell de bases de l'eix llarg.

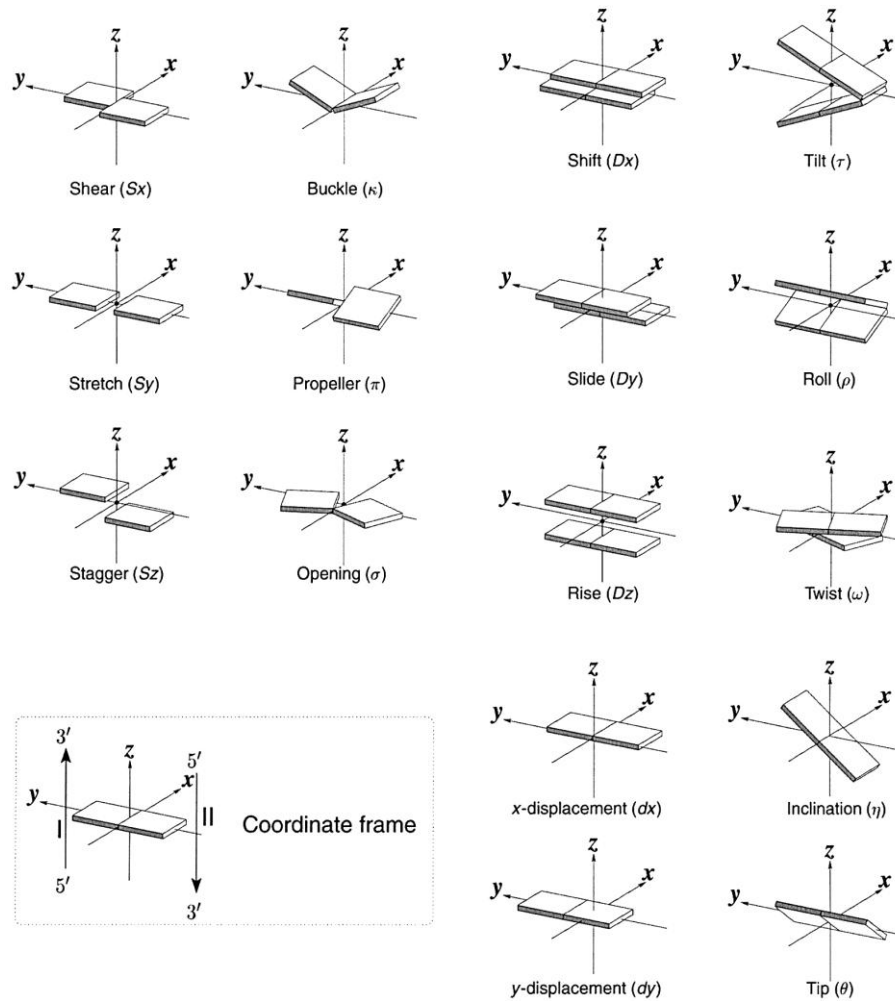


Fig. 1.20. Paràmetres geomètrics pels parells de bases.

Degut a motius il·lustratius, l'*helical twist* (gir d'hèlix) (Ω) és el mateix que el *twist* (gir) (ω) i l'*helical rise* (ascensió d'hèlix) (h) és el mateix que el *rise* (ascensió) (Dz). Imatge extreta de [62].

Taula 1.5. Valors promitjos dels paràmetres geomètrics dels parells de bases per a cristalls de la forma A i B.

	Forma A	Forma B
<i>Shift</i> (Å)	0,00	-0,02
<i>Slide</i> (Å)	-1,53	0,23
<i>Rise</i> (Å)	3,32	3,32
<i>Tilt</i> (°)	0,1	-0,1
<i>Roll</i> (°)	8,0	0,6
<i>Twist</i> (°)	31,1	36,0
<i>X-displacement</i> (Å)	-4,17	0,05
<i>Y-displacement</i> (Å)	0,01	0,02
<i>Helical rise</i> (Å)	2,83	3,29
<i>Inclination</i> (°)	14,7	2,1
<i>Tip</i> (°)	-0,1	0,0
<i>Helical twist</i> (°)	32,5	36,5

Taula adaptada de [63].

1.2.3. Angles de torsió de l'esquelet

L'esquelet fosfodièster d'un oligonucleòtid té sis diferents angles de torsió designats com α , β , γ , δ , ϵ , i ζ a més a més dels cinc angles de torsió interns del sucre ($\nu_0, \nu_1, \nu_2, \nu_3, \nu_4$) i de l'angle de l'enllaç glicosídic (χ) (v. fig. 1.21). Les consideracions estèriques dels angles de l'esquelet del DNA restringeixen els seus valors en rangs discrets i no poden adoptar qualsevol valor entre 0° i 360° . A la fig. 1.21 s'observa una roda conformacional que mostra els valors preferents per a aquests angles de torsió.

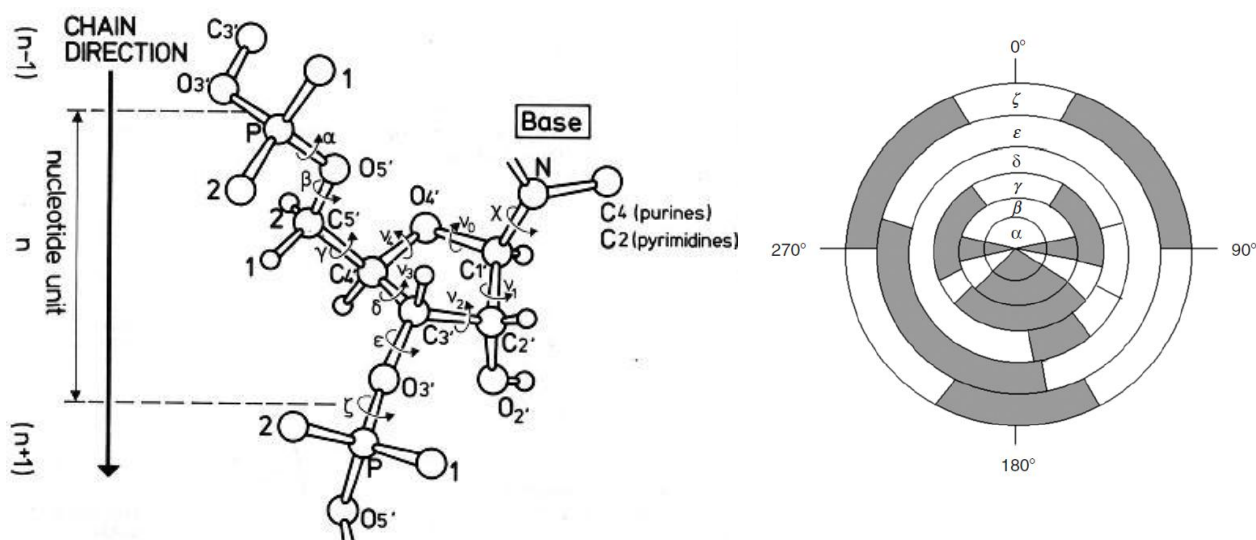


Fig. 1.21. Angles de torsió de l'esquelet en àcids nucleics i roda conformacional dels angles de torsió.

A la roda conformacional es mostren ombrejats els rangs permesos per a aquests valors.

Figures extretes de [<http://x3dna.org/>] i [61], respectivament.

Els angles de torsió èpsilon (ϵ) i zeta (ζ) determinen la posició del fosfat respecte del sucre del residu anterior. Un fosfat pot adoptar la conformació BI i la BII, sent la BI la trobada de forma més comú en dobles hèlixs. No obstant això, alguns grups fosfats canvien a conformació BII temporalment. Aquests residus mostren un equilibri dinàmic entre totes dues conformacions. La diferència entre èpsilon i zeta es fa servir per a caracteritzar la conformació del grup fosfat, sent els valors per a les dues conformacions:

$$\text{BI: } \epsilon - \zeta = -160^\circ \rightarrow +20^\circ$$

$$\text{BII: } \epsilon - \zeta = +20^\circ \rightarrow +200^\circ$$

1.2.4. Solc menor i solc major

Els costats de cada parell de bases estan exposats de tal manera que en el solc menor i en el major es creen uns patrons d'àtoms donadors i acceptors de punts d'hidrogen i interaccions de van der Waals específics per a cada parell de bases (fig. 1.22). En el solc menor, el N3 de l'adenina o de la guanina i el O2 de la timina o de la citosina poden servir d'acceptors de punts d'hidrogen. El grup amino unit al C2 de la guanina pot ser donador de punts d'hidrogen. En canvi al solc major, el N7 de la guanina o l'adenina, el O4 de la timina i el O6 de la guanina són acceptors mentre que els grups amino de la C6 adenina i C4 citosina poden ser donadors.

D'aquesta manera es creen patrons específics de formació de punts d'hidrogen per l'exposició dels costats de les bases. En el cas del solc major aquests patrons permeten distingir les bases A:T de les C:G i fins i tot diferenciar A:T de T:A i C:G de G:C. Per exemple, fent servir la nomenclatura de la fig. 1.22, l'entorn AADH (G:C) és molt diferent de l'ADAM (A:T). Aquest codi de grups químics en el solc major especifica la identitat del parell de bases. Això és important perquè permet a les proteïnes identificar de manera inequívoca una seqüència de DNA. El solc menor, en canvi, no presenta una informació química tan rica. El codi per a A:T i per a T:A és el mateix (AHA) i passa igual per a C:G i G:C. Així, des del punt de vista de la formació de punts d'hidrogen, els parells de bases A:T i C:G semblen diferents, no obstant això, els parells de bases A:T i T:A (així com C:G i G:C) no són fàcilment diferenciables [53].

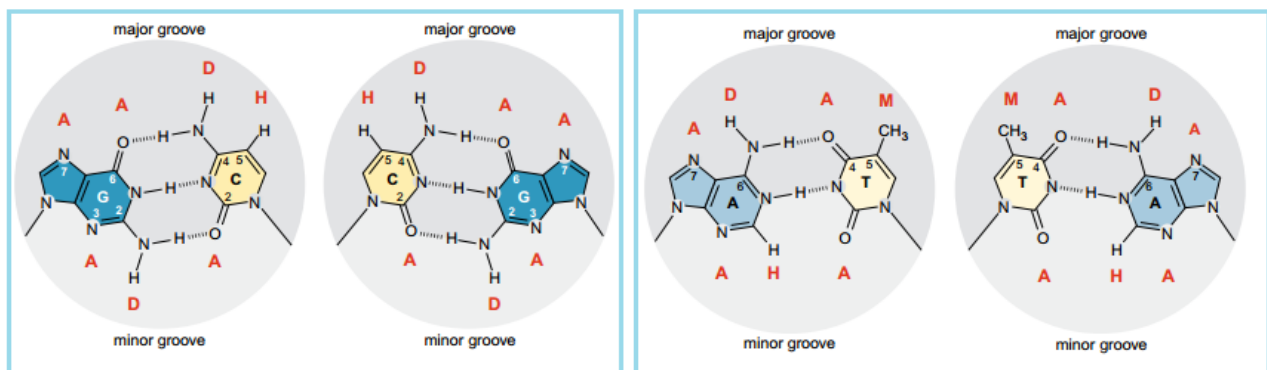


Fig. 1.22. Grups exposats en els costats del solc menor i major.

Les lletres identifiquen els acceptors de punts d'hidrogen (A), els donadors de punts d'hidrogen (D), els hidrògens no polars (H) i els grups metil (M). Imatge extreta de [53].

1.2.5. Topologia del DNA circular

Les dobles hèlixs de DNA es poden plegar sobre si mateixes per a formar estructures terciàries creades mitjançant superenrotllament. Es poden considerar DNA circulars (molècules de DNA covalentment tancades) per a entendre aquest fenomen més fàcilment, però també és aplicable a molècules de DNA obligades a girar d'altres maneres. En realitat, la major part de les molècules de DNA de la cèl·lula presenten superenrotllament.

Considerem un DNA lineal en forma B de 260 pb (v. fig. 1.23). Com en una molècula de DNA circular relaxada cada volta d'hèlix suposa 10,4 parells de bases, aquesta molècula de DNA fa 25 voltes ($260/10,4$). Si s'uneixen els extrems de la molècula obtenim un DNA circular relaxat (fig. 1.23, B). Es pot formar un DNA circular diferent si el dúplex es desenrotlla dues voltes abans d'unir els extrems (fig. 1.23, C). Hi ha dues conformacions possibles un cop lligats els extrems. El DNA es pot plegar en una estructura que conté 23 voltes d'hèlix i un llaç desenrotllat (fig. 1.23, D) o bé la doble hèlix es pot plegar per a entrecreuar-se amb si mateixa (fig. 1.23, E). Aquesta classe d'entrecreuaments s'anomenen superenrotllaments (de l'anglès *supercoiling*). En aquest cas, és una estructura superenrotllada amb 25 voltes d'hèlix B i dues voltes de superhèlix dextrogira (o voltes de superhèlix negatives). Una molècula de DNA superenrotllada és més compacta que una molècula de DNA relaxat de la mateixa llargària. Per tant, el DNA superenrotllat es mou més de pressa que el relaxat quan s'analitza mitjançant electroforesi. El desenrotllament de l'hèlix provocarà superenrotllament en molècules covalentment tancades.

Una propietat topològica d'una molècula circular de DNA és el seu nombre d'enllaç (Lk , de l'anglès *linking number*), que és igual al nombre de voltes que una cadena de DNA creua en sentit dextrogir l'eix de l'hèlix quan aquesta es projecta en un pla. Per al DNA relaxat de la fig. 1.23 B, $Lk=25$. Per a la molècula parcialment desenrotllada (D) i per a la superenrotllada (E), $Lk=23$, ja que abans d'unir els extrems s'han desenrotllat dues voltes. Les molècules que només es diferencien en el nombre d'enllaç són isòmers topològics o topoisòmers. Els topoisòmers de DNA es poden transformar entre si simplement si es talla una o les dues cadenes del DNA i després es tornen a unir.

El DNA desenrotllat i el superenrotllat de la figura (D i E) són idèntics topològicament, però diferents geomètricament. Tenen el mateix valor de Lk , però són diferents en el gir (*twist*, Tw) i la torsió (*writhe*, Wr). Encara que les definicions rigoroses de gir i torsió són complexos, el gir és una mesura de l'enrotllament helicoidal de les cadenes de DNA entre si, mentre que la torsió és una mesura de l'enrotllament dels eixos de la doble hèlix, que s'anomena superenrotllament. A un enrotllament dextrogir se li assigna valor negatiu (superenrotllament negatiu) i al levogir, un nombre positiu (superenrotllament positiu).

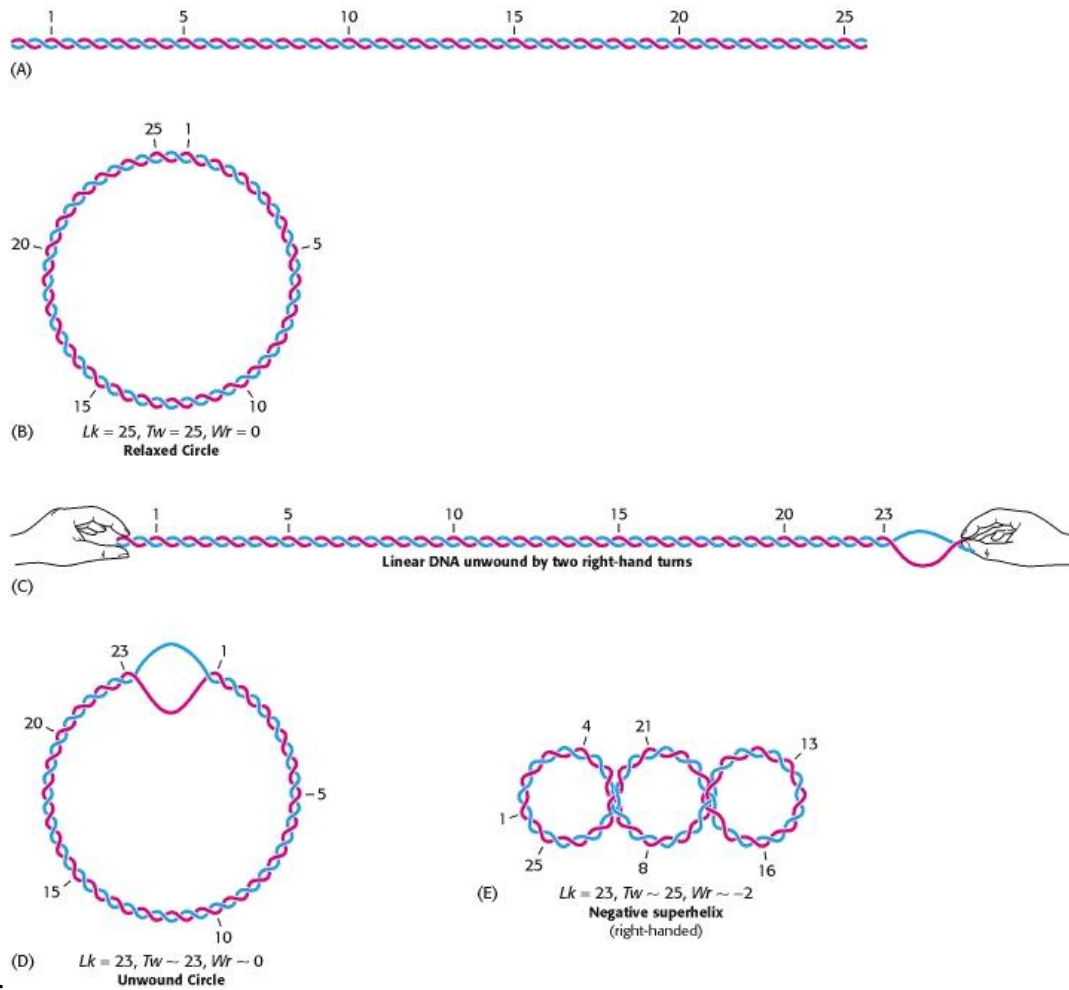


Fig. 1.23. Relacions entre el nombre d'enllaç (Lk), el gir (Tw) i la torsió (Wr) per a un DNA de 260 pb [55].

Per a qualsevol molècula de DNA es compleix que la suma del gir (Tw) i de la torsió (Wr) és igual al nombre d'enllaç (Lk) [57]: $Lk = Tw + Wr$.

Això implica que per a una molècula de DNA circular tancada, sent Lk invariant, qualsevol canvi en el gir (Tw) del DNA ha de ser acompanyat per un canvi igual i oposat en la torsió (Wr), i a l'inrevés. A més a més, qualsevol canvi en el nombre d'enllaç es manifesta geomètricament com un canvi en el gir i/o la torsió: $\Delta Lk = \Delta Tw + \Delta Wr$.

Els enzims capaços d'alterar el nombre d'enllaç del DNA són les topoisomereses. El procés consta de tres etapes: l'escissió d'una o de les dues cadenes del DNA, el pas d'un segment de DNA a través d'aquest lloc d'escissió i el tancament del tall en el DNA. Les topoisomereses del tipus I catalitzen el relaxament del DNA superenrotllat (tallant una cadena), mentre que les del tipus II afegeixen superenrotllaments negatius (tallant les dues). Totes dues tenen un paper molt important en la replicació, la transcripció i la recombinació del DNA.

1.2.6. Altres estructures del DNA

Cruciforms i unions de Holliday

A part de l'habilitat del DNA per adoptar diferents formes de la seva hèlix (formes A, B i Z, entre d'altres), es poden trobar d'altres estructures més complexes que parteixen de la doble hèlix. Algunes d'aquestes estructures són estabilitzades gràcies al superenrotllament del DNA, com és el cas del cruciforms. Els cruciforms sorgeixen com una conseqüència d'un canvi en l'aparellament de les bases dins de les pròpies cadenes de la doble hèlix del DNA i consisteix en una estructura en forma de creu amb dues tiges (o braços), dos *loops* (també anomenats *hairpins*) i una *four-way junction* (unió de quatre vies) (v. fig. 1.24) [57][64]. La formació d'aquestes estructures requereix la presència de repeticions invertides de la seqüència (palíndroms) de la doble hèlix del DNA, és a dir, una seqüència seguida immediatament (o una mica després) de la mateixa seqüència però en el sentit oposat. La formació de cruciforms en el B-DNA lineal és termodinàmicament desfavorable però aquest tipus d'estructura es pot observar experimentalment en DNA circular amb superenrotllament negatiu. Així, l'energia lliure desfavorable associada amb el supercoiling negatiu és dissipada amb la formació de cruciforms.

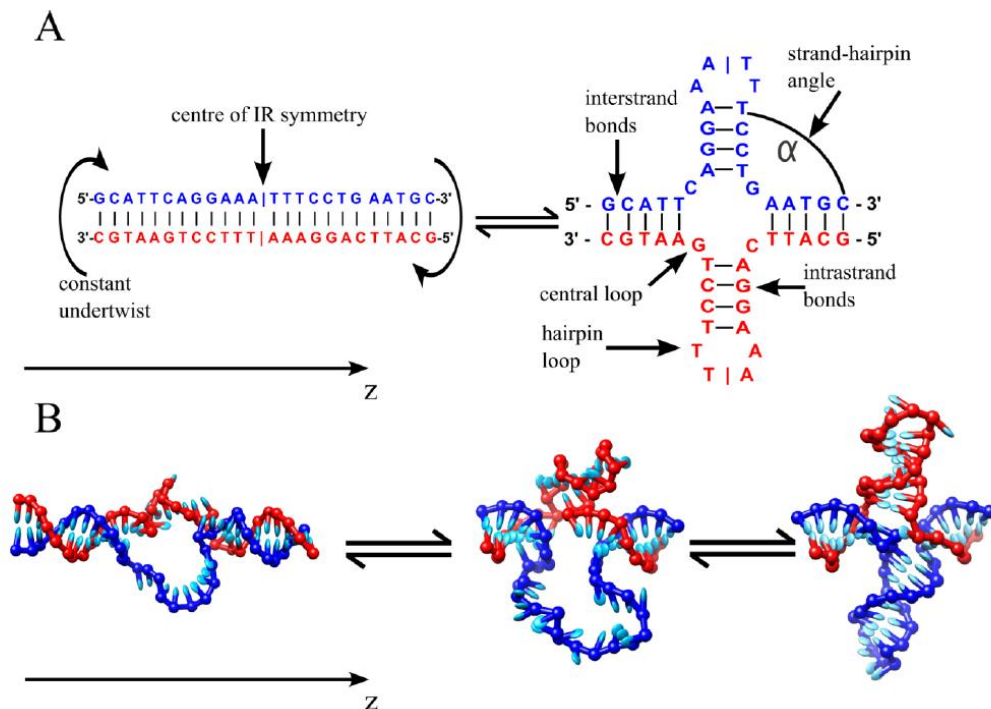


Fig. 1.24. Formació d'un cruciform a partir d'una seqüència desenrotllada contenint una repetició invertida (IR).

A) Representació esquemàtica de la formació d'un cruciform. B) Estructures d'una seqüència de DNA amb una bombolla, formació d'un únic braç i cruciform completament format. Imatge extreta de [65].

Durant la recombinació es formen unes estructures intermèdies en forma de creu anomenades unions de Holliday. En aquest procés dues dobles hèlixs homologues s'alineen, i com a resultat de l'escissió d'una cadena de cada hèlix, l'aparellament de bases té lloc amb la cadena intacta de l'altre hèlix. La formació d'enllaços de les cadenes de DNA resulta en la formació d'una unió de Holliday, que és similar a un cruciform però sense els dos *loops*. En la recombinació, la successiva escissió i formació d'enllaços de les altres cadenes, genera molècules de DNA recombinants. De forma general aquest tipus d'estructures, equivalents a la unió de Holliday en la recombinació, són considerades estructures *four-way junctions*.

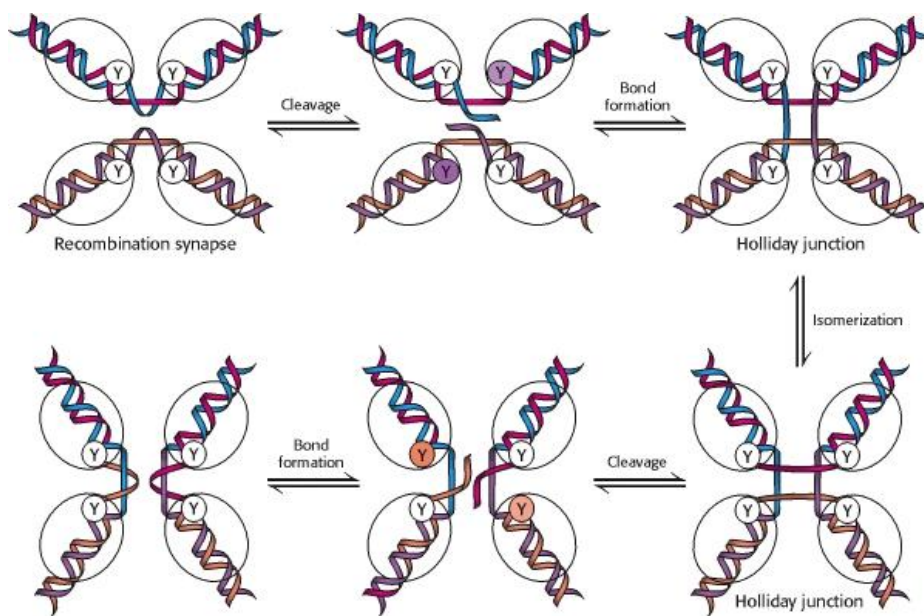


Fig. 1.25. Formació d'unionis de Holliday durant la recombinació homòloga.

Imatge extreta de [55].

PX-DNA

Els DNAs que presenten un alt superenrotllament (estat d'alt nivell energètic) poden formar estructures diferents a la tradicional forma B del DNA. Se sap que les molècules amb superenrotllament poden experimentar transicions estructurals relacionades amb la seqüència. Per exemple, les seqüències palindròmiques poden resultar en la extrusió de cruciforms [66] i les seqüències alternants purina-pirimidina poden produir la forma Z del DNA [67]. Aquestes transicions estructurals relaxen el DNA i simultàniament produeixen una característica estructural distintiva.

Fa pocs anys, el Dr. Seeman i els seus col·laboradors van descriure la transició estructural que pot experimentar un DNA amb superenrotllament com a conseqüència de l'homologia de seqüència [68]. Aquest tipus de transició en segments de DNA homòlegs provoca la formació d'una estructura anomenada PX-DNA (de l'anglès *Paranemic Crossover DNA*). El PX-DNA

presenta unions de Holliday de forma generalitzada. Concretament, el PX-DNA pot ser descrit com un DNA de quatre cadenes en el qual dos dobles hèlixs (en blau i vermell a la fig. 1.26 A) estan enllaçades per intercanvi recíproc de les cadenes de la mateixa polaritat a cada punt on les cadenes es troben [69]. Per tal que tingui lloc aquest tipus d'estructura és necessari, com a mínim, una homologia de seqüència entre la doble hèlix vermella i la blava en les mitges voltes anomenades H de la figura (homologia PX). Per suposat, l'homologia total de seqüència podria produir també aquesta estructura.

El pas d'hèlix de les estructures PX-DNA és aproximadament el doble en comparació amb la forma B del DNA, ja que cada cadena s'ha d'enrotllar al voltant de les dues hèlixs per arribar a fer una volta sencera. Així, el seu *twist* s'espera que sigui de promig la meitat que el del DNA convencional. En un DNA circular, les molècules resultants d'aquesta transició estructural presenten una forma similar a unes manuelles (*dumbbell-like structure*) (amb un segment PX central i un *loop* a cada extrem), on la longitud del fragment PX és funció de la localització i del grau d'homologia de seqüència (v. fig. 1.26 B).

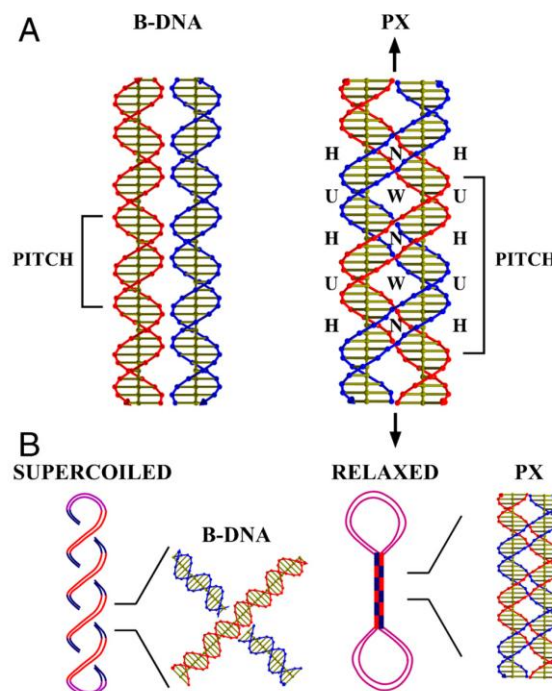


Fig. 1.26. Comparació del B-DNA amb el PX-DNA i esquema de la seva formació en un DNA circular [68].

A) Comparació del B-DNA amb el motiu PX. B) Comparació d'un DNA circular superenrotllat amb la seva forma relaxada per la formació d'un fragment PX, provocant una estructura en forma de manella (*dumbbell-like structure*).

La formació i l'estudi de les estructures tipus PX-DNA s'ha realitzat tant per anàlisis electroforètiques en gels bidimensionals de plasmidis com per microscòpia de força atòmica (AFM) [68].

La transició estructural d'un DNA circular amb alt superenrotllament negatiu a un de relaxat i presentant un fragment de la forma PX es pot visualitzar com un salt en el perfil de taques en un gel bidimensional (v. fig. 1.27).

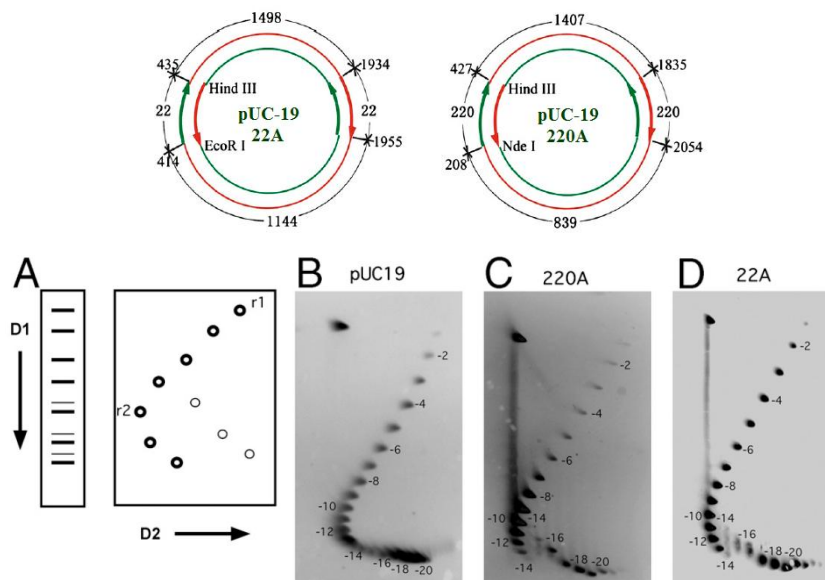


Fig. 1.27. Plasmidis amb inserts de 22 pb i 200 pb (pUC-19 22A i pUC-19 220A) i electroforesis bidimensionals per a l'observació de la transició estructural [68].

A) Esquema d'una electroforesi bidimensional amb transició estructural. B) Plasmidi pUC19 (control). C) Plasmidi amb dos inserts en costats oposats de 22 pb cadascun. D) Plasmidi amb dos inserts en costats oposats de 220 pb cadascun. En l'electroforesi bidimensional del plasmidi control (pUC-19) (sense cap insert) podem veure un perfil de taques continu (B) en canvi, degut a la transició estructural en la primera dimensió observem una discontinuïtat en el perfil de taques per als plasmidis amb els inserts de 22 i 220 pb (C i D).

D'altra banda, quan les molècules que presenten aquesta transició estructural s'observen per AFM es pot apreciar la forma de manuelles que adopten els DNAs circular al formar-se una estructura tipus PX-DNA al segment cohesiu central (v. fig. 1.28).

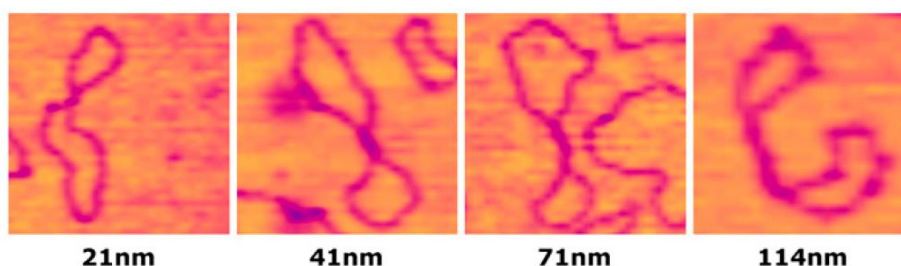


Fig. 1.28. Imatges d'AFM de les estructures tipus manuelles amb diferents longituds del segment cohesiu [68].

Imatges d'AFM obtingudes de molècules amb transició estructural corresponents al plasmidi pUC-19 220A (extret del gel bidimensional de la fig. 1.27 C). La característica comuna de les estructures en forma de manuelles (*dumbbell-like structure*) és la formació d'un segment central que connecta els *loops* dels extrems de la molècula.

1.2.7. Interès de les seqüències ATs

L'objectiu d'aquest treball es centra en la interacció de les proteïnes HMG justament amb seqüències riques en els parells de bases A i T. Però aquest tipus de seqüències són també de gran interès per si mateixes [70].

Una de les característiques del genoma humà, i en general de tots els organismes eucariotes, és la gran quantitat de DNA amb funció desconeguda. Concretament, entre un 90% i un 97% del genoma està format per seqüències més o menys repetitives que no són traduïdes a proteïnes. Entre aquestes abunden les seqüències riques en parells de bases ATs [71][72][73]. Un exemple extrem de DNA ric en ATs és el de l'ameba *Dictyostelium discoideum*, que posseeix un contingut en ATs d'un 78% [74]. D'altra banda algunes seqüències ATs, sense ser codificants, juguen un paper molt important en el metabolisme del DNA. Aquest alt contingut en ATs pot jugar un rol estructural en la funció del genoma.

Per aquestes raons existeix un interès considerable en l'estudi del polimorfisme que presenten aquestes seqüències riques en ATs [75][76]. Amb la comprensió de l'estructura d'aquestes seqüències es pretén poder entendre millor la seva funció. S'han descrit amb anterioritat diverses formes del DNA per aquestes seqüències riques en ATs com són la forma B [75], la forma B modificada [77] i la forma D [78] (aquesta forma es troba a humitats molt baixes i té 8 parells de bases per volta).

També s'ha descobert en el nostre laboratori que les seqüències alternants ATs (o $(AT)_n$) poden adoptar l'aparellament Hoogsteen, que no s'havia trobat abans en una estructura en doble hèlix antiparal·lela [79]. En el parell de bases del tipus Hoogsteen l'adenina gira 180° al voltant de l'enllaç glicosídic (conformació *syn*), d'aquesta manera els ponts d'hidrogen que es formen entre l'adenina i la timina són diferents de l'aparellament Watson-Crick.

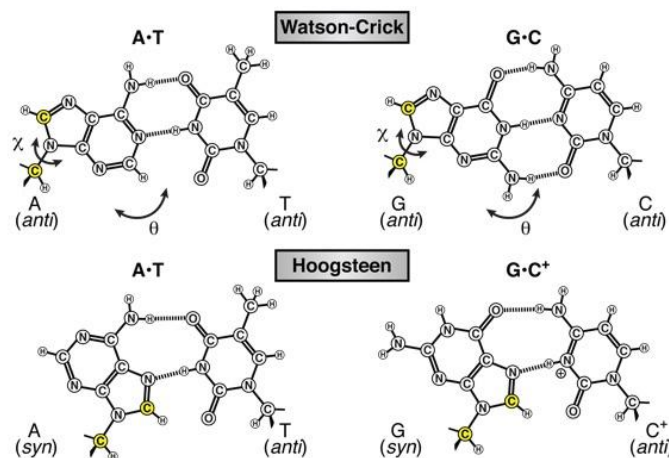


Fig. 1.29. Aparellament Hoogsteen [80].

2. Plantejament i objectius

L'objectiu final d'aquest treball és l'estudi de la interacció de les proteïnes HMG amb DNA mitjançant l'obtenció de l'estructura atòmica que formen els seus complexos per cristal·lografia de raigs X. Per a assolir aquest objectiu s'ha treballat amb oligonucleòtids rics en ATs i diferents membres de dues famílies de les proteïnes HMG (les HMGA i les HMGB) així com amb pèptids corresponents als seus motius d'unió al DNA.

L'obtenció de l'estructura tridimensional dels complexos de les proteïnes HMGA i HMGB amb oligonucleòtids comporta diferents etapes. En primer lloc, l'obtenció de proteïna prou pura i en quantitat suficient, a continuació la realització d'assajos cristal·logràfics per a aconseguir cristalls aptes dels complexos amb DNA i finalment la resolució de l'estructura a partir dels corresponents diagrames de difracció.

D'altra banda també s'ha volgut estudiar l'efecte en la topologia del DNA que produeixen les proteïnes HMGB i la transició estructural a estructures del tipus PX-DNA en DNAs circulars contenint fragments de seqüències d'ATs alternants.

Així doncs, aquesta tesi comprèn els següents objectius parcials:

- Obtenció d'HMGA: expressió i purificació de proteïnes HMGA1 humanes i construcció, expressió i purificació del mutant HMGA1b(1-79).
- Obtenció d'HMGBs: expressió i purificació dels fragments Box AS, Box BS i Box ABS de la proteïna HMGB1 de mamífer i de la proteïna NHP6A (HMGB de llevat).
- Cristal·lització de les proteïnes HMGA i pèptids sintètics corresponents als seus dominis *AT-hook* amb oligonucleòtids rics en ATs.
- Cristal·lització dels fragments Box AS, Box BS i Box ABS de la proteïna HMGB1 i la proteïna NHP6A amb oligonucleòtids rics en ATs.
- Difracció de raigs X dels cristalls obtinguts d'aquests complexos, anàlisi dels diagrames de difracció i resolució de l'estructura per reemplaçament molecular de les millors difraccions.
- Estudi de les interaccions i dels mecanismes de reconeixement entre les proteïnes HMG i el DNA a partir de l'estructura tridimensional dels seus complexos.

- Estudi de l'efecte en la superhelicitat de DNAs circulars dels fragments Box AS i Box ABS de la proteïna HMGB1 i l'NHP6A mitjançant assajos de superenrotllament en presència de topoisomerasa I.
- Estudi de la transició a una estructura del tipus PX-DNA de plasmidis amb inserts ATs mitjançant electroforesis bidimensionals en gels d'agarosa i anàlisi per microscòpia de força atòmica.

3. Materials i mètodes

Per a abordar la determinació de l'estructura tridimensional dels complexos de les proteïnes HMGA i HMGB amb el DNA, en primer lloc, cal obtenir co-cristalls de la proteïna (o pèptids de la mateixa) amb oligonucleòtids sintètics. Els pèptids i els diferents oligonucleòtids rics en ATs utilitzats en aquest treball han estat sintetitzats a la Universitat Pompeu Fabra i a l'Institut Pasteur de París, respectivament. En canvi, les proteïnes HMG han estat obtingudes a partir de cultius bacterians mitjançant la tècnica del DNA recombinant com s'explicarà a continuació.

Posteriorment els cristalls obtinguts es difractaran amb raigs X. En aquest treball s'han fet servir les instal·lacions del sincrotró Grenoble ESRF (*European Synchrotron Radiation Facilities*) i del sincrotró ALBA (Cerdanyola del Vallès), així com les instal·lacions de raigs X de la *Colorado University Cancer Center* i de la Plataforma Automatitzada de Cristal·lografia del Parc Científic de Barcelona. Finalment, amb l'estudi de les intensitats de les reflexions s'aborda la resolució de l'estructura pel mètode del reemplaçament molecular i l'afinament de l'estructura trobada.

Per tant l'esquema simplificat del procediment de treball seria el següent:

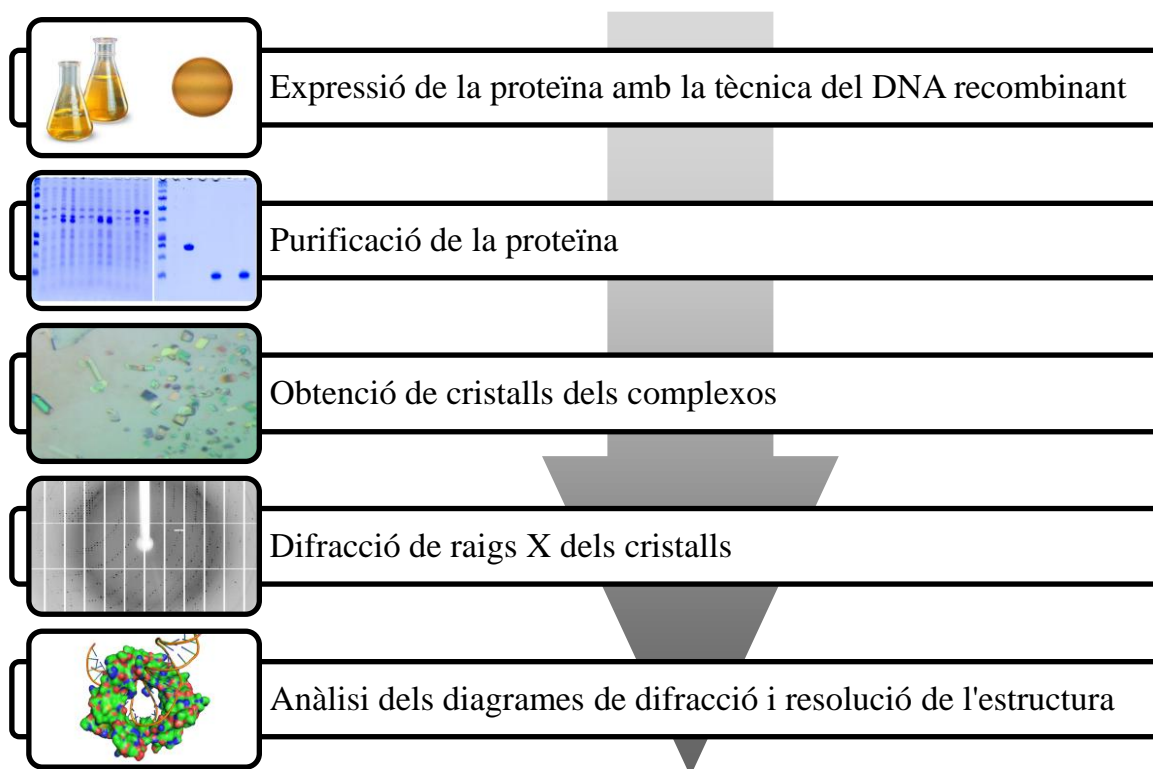


Fig. 3.1. Esquema del procés d'obtenció de l'estructura dels complexos proteïna-DNA.

3.1. Proteïnes utilitzades

Al llarg d'aquest treball s'han estudiat diferents membres de la família de les HMG. En el grup de les HMGA humanes s'han estudiat les proteïnes HMGA1a i HMGA1b i el fragment HMGA1b(1-79). Del grup de les HMGB s'ha treballat amb fragments de l'HMGB1 (Box AS, Box BS i Box ABS) (v. fig. 3.6) i amb la proteïna HMG-box de llevat, l'NHP6A. Les proteïnes HMGB han estat expressades i purificades durant l'estada al *Department of Pharmacology* de la *University of Colorado* (USA) sota la supervisió de la Dra. Mair Churchill. A l'annex A es recullen les seqüències, les masses moleculars i d'altres característiques d'aquestes proteïnes.

L'obtenció de totes aquestes proteïnes s'ha realitzat amb la tècnica del DNA recombinant però s'han fet servir diferents soques i vectors d'expressió (o plasmidis) en cada cas. A més, els passos per a aconseguir una mostra de proteïna pura i en abundància a partir d'un cultiu bacterià han estat diferents pels requeriments de cada proteïna o família de proteïnes. En la taula 3.1 es recull la soca i el tipus de plasmidi utilitzat per a l'expressió de la proteïna i els passos seguits durant la purificació dividits en les tres fases principals (captura, purificació intermèdia i poliment). De forma general es pot dir que en la captura s'aïlla i es concentra la proteïna d'interès, en la purificació intermèdia s'eliminen gran part de les impureses i finalment en el poliment s'arriba al màxim grau de puresa.

Taula 3.1. Resum dels principals passos per a l'obtenció de les diferents proteïnes.

Nom de la proteïna	Expressió de la proteïna		Purificació de la proteïna				
	Soca <i>E. coli</i>	Plasmidi	Captura		Purificació intermèdia		Poliment
HMGA1a HMGA1b	Rosetta(DE3) pLysS	pET-3a	Extracció àcida	-	Cromatografia de bescanvi catiònic	-	-
HMGA1b(1-79)*	Rosetta(DE3) pLysS	pET-3a	Extracció àcida	-	Cromatografia de bescanvi catiònic	-	Cromatografia d'exclusió molecular
HMGB1 (Box AS, BS, ABS)	Rosetta(DE3) pLysS	pGEX-2T	Sonicació	-	Cromatografia d'afinitat (Purificació GST)	Cromatografia de bescanvi catiònic	Cromatografia d'exclusió molecular
NHP6A	BL21(DE3)	pET-11a	Sonicació	Precipitació amb sulfat d'amoni	Cromatografia de bescanvi aniònic	Cromatografia de bescanvi catiònic	Cromatografia d'exclusió molecular

*Obtinguda per mutagènesi dirigida a partir d'HMGA1b (v. 3.2.1).

A continuació es detalla la metodologia seguida per a l'obtenció d'aquestes proteïnes. Degut als seus passos de purificació diferenciats, l'obtenció de les proteïnes HMGA, HMGB i NHP6A s'explica per separat en els apartats següents.

3.2. Obtenció de les proteïnes HMGA

3.2.1. Mutagènesi dirigida

Per a l'obtenció del mutant HMGA1b(1-79), que conté els tres *AT-hooks* però no la cua C-terminal carregada negativament (v. fig. 4.1), s'ha fet servir la tècnica de la mutagènesi dirigida. Aquesta és una tècnica molt potent que ens permet canviar la seqüència d'un gen conegut en un punt específic. Per tal de produir aquest mutant ha calgut modificar el DNA plasmídic que codifica la proteïna original HMGA1b introduint-hi un codó stop en la posició corresponent a l'aminoàcid 80. Hem realitzat la mutació amb el kit de mutagènesi dirigida *QuickChange™ Site-Directed Mutagenesis Kit*, de la casa comercial Stratagene.

Per a realitzar la tècnica de la mutagènesi dirigida (v. fig. 3.2), es parteix d'un vector d'expressió o plasmidi motlle de doble cadena i de dos *primers* sintètics i complementaris que contenen la mutació. Cadascun dels *primers* s'uneixen complementàriament a cada una de les cadenes del plasmidi. A continuació els *primers* s'estenen durant el cicle tèrmic per l'acció de la DNA polimerasa *PfuTurbo*, generant d'aquesta manera el plasmidi mutat. Aquest plasmidi mutat contindrà un *nick* (falta d'un enllaç fosfodièster) en cada cadena. A continuació la mostra es tractarà amb *Dpn I*. La *Dpn I* és una endonucleasa capaç de digerir el DNA metilat i en aquest cas s'utilitza per a digerir el plasmidi motlle o parental. Finalment el plasmidi que conté la mutació desitjada es transformarà en una soca de cèl·lules supercompetents XL1-Blue (en aquest moment es repararan els *nicks*).

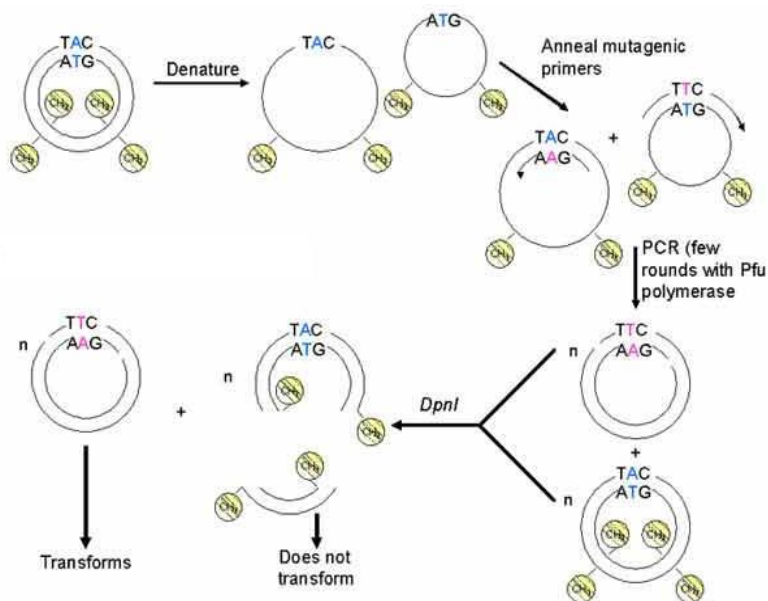


Fig. 3.2. Esquema de la mutagènesi dirigida.

En el nostre cas el plasmidi motlle o parental és el corresponent a la proteïna sencera HMGA1b, obtingut prèviament al nostre laboratori, i els *primers* sintètics van ser adquirits de la casa comercial *Thermo Fisher Scientific*. Els *primers* van ser dissenyats per tal que fossin complementaris al fragment del gen d'interès i que continguessin la mutació desitjada, a més a més aquests han de complir uns certs requisits que es recullen a la taula 3.2.

Taula 3.2. Disseny dels primers.

Disseny dels primers
<ul style="list-style-type: none"> - Contenir la mutació i hibridar-se en la mateixa seqüència però en cadenes oposades del plasmidi. - Han de ser d'entre 25-45 parells de bases de longitud i tenir una $T_m \geq 78^\circ\text{C}$: $T_m = 81.5 + 0.40 (\%GC) - 675/N - \%Mismatch$ <ul style="list-style-type: none"> %GC: tant per cent de bases G i C N: nombre de bases de l'oligonucleòtid %Mismatch: tant per cent de desaparellament amb la seqüència original - La mutació ha d'estar al mig del <i>primer</i>, amb 10-15 bases per cada banda. - Es recomana que continguin un mínim d'un 40% en %GC i acabar en una o més C o G.

Els *primers* han de complir les condicions que es mostren a la taula.

Després d'analitzar diferents possibilitats es van triar les dues cadenes complementàries que es mostren a continuació en ordre 5' → 3' (el triplet stop està indicat amb un requadre):

Primer 5' HMGA1b(1-79) GGCAGACCCAAAAACTG TAG AAGGAGGAAGAGG

Primer 3' HMGA1b(1-79) CCTCTTCCTCCTT CTA CAGTTTTTTGGGTCTGCC

Breument, el protocol seguit a continuació és el següent: en primer lloc preparar la barreja de reacció, a continuació cal desenvolupar els cicles tèrmics, tot seguit digerir les cadenes parentals amb l'endonucleasa *Dpn I* i finalment realitzar la transformació en cèl·lules supercompetents XL1 Blue. A partir d'un cultiu d'aquestes cèl·lules purificarem el nou DNA plasmídic ("midiprep") i comprovarem per seqüenciació que s'ha incorporat la mutació desitjada.

Barreja de reacció

La barreja de reacció s'ha preparat d'acord amb la taula 3.3. Cal comentar que en aquest cas hem treballat amb una quantitat de plasmidi motlle de 20 ng i que també hem afegit DMSO (dimetil sulfòxid) fins al 3% per tal d'evitar la formació d'estructures secundàries.

Taula 3.3. Components de la barreja de reacció.

Barreja de reacció
<ul style="list-style-type: none"> • 5 µl tampó reacció 10x • 2 µl plasmidi (10 ng/µl) • 1,25 µl primer 5' (100 ng/µl) • 1,25 µl primer 3' (100 ng/µl) • 1 µl dNTP mix (barreja de nucleòtids) • 1,5 µl DMSO (3% final) • Volum adient d'aigua per a arribar a 50 µl (en aquest cas, 38 µl) <p style="text-align: center;">A continuació afegir 1 µl de DNA polimerasa <i>PfuTurbo</i> (2,5 U/µl)</p>

Cicles tèrmics

Els cicles tèrmics els hem realitzat amb un miniciclador que hem programat segons la taula següent que ens facilita el fabricant. A la barreja anterior cal afegir-li 30 µl d'oli mineral abans de començar els cicles tèrmics per a evitar evaporacions durant el procés.

Taula 3.4. Detalls dels cicles tèrmics.

Cicles tèrmics			
Segment	Cicles	Temperatura	Temps
1	1	95 °C	30 segons
2	12-18*	95 °C	30 segons
		55 °C	1 minut
		68 °C	2 minuts/kb de plasmidi

*El nombre de cicles del segon segment depèn de si desitgem aconseguir una única mutació en una base, en més d'una base d'un mateix aminoàcid o delecions (o insercions) de més d'un aminoàcid, en aquests casos es treballaria amb 12, 16 o 18 cicles respectivament.

En aquesta taula tenim dos paràmetres que depenen del tipus de mutació i de la grandària de la mostra amb la que treballem, el nombre de cicles del segment dos i el temps del període a 68°C, respectivament. En el nostre cas, per a un plasmidi de 5kb (v. fig. 3.3) necessitaríem un temps de 10 minuts a 68°C per cicle i el nombre de cicles ha estat de 12.

Per a acabar el tractament tèrmic hem portat les mostres a 4°C.

Digestió amb Dpn I

Com hem comentat, en aquesta etapa l'endonucleasa *Dpn I* digereix les cadenes parentals. Hem procedit de la següent manera:

- Treure amb compte l'oli mineral de la mostra.
- Afegir 1 µl de l'enzim de restricció *Dpn I* (10 U/µl).
- Incubar 1 hora a 37°C.

Transformació

Hem fet servir les cèl·lules XL1 Blue supercompetents del *kit* per a realitzar la transformació. Hem fet un control negatiu amb cèl·lules no transformades per a veure que el procediment era correcte. Concretament:

- Aliquotar 50 µl d'XL1 Blue en dos eppendorfs.
- En un d'ells afegim 1 µl del DNA tractat amb *Dpn I* (l'altre serà el control).
- Barrejar i posar en gel 30 minuts.
- Xoc tèrmic: posar a 42°C durant 45 segons.
- Posar en gel 2 min.
- Afegir 0.5 ml de medi NZY+.
- Incubar 1 hora a 37°C a 250 rpm.
- Plaquejar en LB + Amp (50 µg/ml) la mostra i el control (la mostra repartida en dues plaques, una amb 150 µl i l'altra amb 350 µl).
- Incubar a 37°C O.N.

Al dia següent només s'observa creixement en les plaques amb cèl·lules transformades.

Cultiu en medi líquid

- Picar una única colònia en un tub amb 1 ml d'LB amb ampil·lina (50 µg/ml).
- Incubar un mínim de 6 hores a 37°C i 250 rpm.
- Passar a erlenmeyers amb 100 ml de LB + Amp (50 µg/ml).
- Incubar a 37°C i agitació (250 rpm) O.N.

Els medis de cultiu que hem fet servir en aquest protocol són els que es recullen a l'annex B.

Obtenció del DNA plasmídic

Per tal d'obtenir preparacions pures de DNA plasmídic hem realitzat Midi-preps (o “midis”) utilitzant el kit *Plasmid purification kit* de Qiagen (v. apartat 3.11).

Seqüenciació del DNA

Per tal de verificar que la mutació s'hagi introduït correctament es seqüencia el nou DNA plasmídic. Les mostres obtingudes es van portar a seqüenciar al Servei de Seqüenciació del Parc Científic de Barcelona. Normalment portem una mostra d'aproximadament 800 ng. El resultat de la seqüenciació és un cromatograma amb la seqüència del DNA analitzat. Cal destacar que no només cal comprovar la incorporació de la mutació desitjada sinó que també cal comprovar que no s'hagi introduït cap variació en la resta de la seqüència.

3.2.2. Expressió en *E. coli*

La introducció en una cèl·lula hoste d'un vector d'expressió (plasmidi) contenint la seqüència de DNA que codifica una determinada proteïna d'interès farà que les cèl·lules transformades siguin capaces de sintetitzar la proteïna que desitgem. El plasmidi pot presentar diferents marcadors de selecció de les cèl·lules transformades, com per exemple resistència a antibiòtics.

En el nostre cas hem fet servir el plasmidi pET-3a que conté el gen de la proteïna HMGA1a humana. Aquest plasmidi va ser cedit gentilment pel Dr. Reeves (*Washington State University*) i conté també un gen que confereix resistència a l'antibiòtic ampil·lina (gen Amp^r). D'altra banda per a l'obtenció de les proteïnes HMGA1b i el fragment HMGA1b(1-79) hem fet servir el plasmidi pET-3a però contenint aquest cop el DNA corresponent a aquestes formes (fig. 3.3).

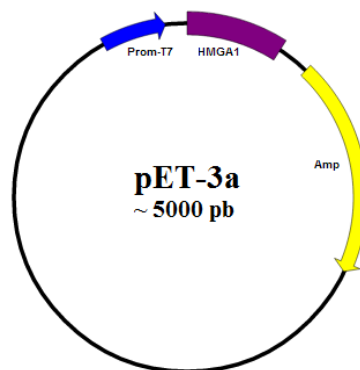


Fig. 3.3. Esquema del plasmidi pET-3a.

A més del gen corresponent a la proteïna HMGA1a humana o a l'HMGA1b(1-79) el plasmidi conté el gen Amp^r que conferirà al bacteri resistència a l'antibiòtic ampil·lina. Figura preparada amb el programa *Savvy* [<http://www.bioinformatics.org/savvy/>].

Les cèl·lules hoste més utilitzades són el bacteri *Escherichia coli*, algunes espècies de *Bacillus* i el llevat *Saccharomyces cerevisiae*. Per a l'expressió de l'HMGA1 hem fet servir la soca Rosetta(DE3)pLysS d'*Escherichia coli*.

El procés de transformació de les cèl·lules hoste no és efectiu al 100%. És a dir, no totes les cèl·lules incorporaran el nou material genètic. Així, per tal d'evitar la multiplicació de les cèl·lules no transformades i seleccionar només les transformades, s'afegeix ampicil·lina al medi de cultiu. En presència d'aquest antibiòtic només les cèl·lules transformades seran capaces de créixer ja que a l'incorporar el plasmidi amb el gen de l'HMGA1 hauran incorporat també el gen Amp^r.

L'inici de l'expressió de la proteïna per part del bacteri es fa per inducció amb IPTG (isopropil- β -D-tiogalactòsid). La presència d'aquest inductor fa que es posi en marxa el procés de síntesi de la proteïna recombinant (v. annex C).

El protocol que hem seguit per a l'expressió de les proteïnes HMGA1 és el que es mostra a continuació. Tots el processos s'han de fer a la flama amb material i medi estèril.

Transformació

- Sobre 100 μ l de cèl·lules competents, o sigui amb la paret cel·lular permeabilitzada, de la soca Rosetta(DE3)pLysS d'*Escherichia coli* afegir 1 μ l de la dissolució de DNA plasmídic (aproximadament 0,1 μ g/ μ l). Fer també un control només amb cèl·lules competents.
- Deixar 10-20 min en gel.
- Xoc tèrmic: 5 min a 37°C al bany.
- Deixar 2 min en gel.

Amb aquests tractaments de refredament i escalfament facilitem que el DNA plasmídic entri dins de les cèl·lules competents.

- Afegir 1 ml de medi LB sense antibiòtics.
- Incubar 1 hora a 37°C al bany.
- Centrifugar durant 2 min a 4000 rpm (micròfuga). Els bacteris queden al fons.
- Eliminar sobrenadant fins deixar aproximadament 100 μ l.
- Resuspendre els bacteris i sembrar-los sobre una placa d'agar contenint LB + Amp (50 μ g/ml) (cultiu en medi sòlid).
- Deixar *overnight* (O.N.) a 37°C a l'estufa.

- L'endemà veurem com a la placa de cèl·lules transformades han crescut algunes colònies de bacteris, mentre que al control no, ja que les cèl·lules no transformades no són resistents a l'antibiòtic.

Cultiu en medi líquid

- Picar una colònia de bacteris transformats sobre un tub amb 3 ml LB + 50 µg/ml Amp.
- Incubar a 37°C i agitació (250 rpm) durant unes 6 hores (o fins observar terbolesa al tub).
- Abocar cultiu sobre 50 ml LB + 50 µg/ml Amp.
- Incubar a 37°C i agitació (250 rpm) O.N.

Inducció de l'expressió

- Inocular suficient volum del cultiu anterior en 1 litre de LB + 50 µg/ml Amp per a assolir una $DO_{600} = 0,1-0,2$.
- Incubar a 37°C i agitació (250 rpm) fins $DO_{600} = 0,6-0,7$.
- Inducció amb IPTG 1 mM.
- Incubar a 37°C i agitació (250 rpm) durant 3 hores.
- Centrifugar a 4000 rpm durant 10 min a 4°C en un rotor GSA.
- Eliminar el sobrenadant.
- Guardar el sediment (bacteris) a -80°C fins al seu processat.

Un cop començada la inducció s'agafaran alíquotes a temps 0 h, 1:30 h i 3 h per a veure l'absorbància i fer controls electroforètics del procés en gels d'SDS com veurem en l'apartat 3.6.1. Per tant, farem:

- Separar 1 ml per a mesurar DO_{600} .
- Agafar una alíquota de 125 µl.
- Centrifugar a 4000 rpm durant 5 min (micròfuga).
- Eliminar sobrenadant.
- Resuspendre el sediment en 30 µl de tampó de mostres per a electroforesi.
- Guardar a -20°C fins al desenvolupament de l'electroforesi.

Els medis de cultiu que hem fet servir en aquest protocol són els que es recullen a l'annex B.

3.2.3. Tractament del sediment bacterià

Per tal d'obtenir les proteïnes HMGA1 que s'han produït a l'interior dels bacteris hem de seguir el següent procediment [81]. Per a 1 litre de cultiu:

- Resuspendre el sediment en 100 ml de 5% HClO_4 /0.5% Tritó X-100 fred. Ho fem pipetejant contínuament fins a obtenir una bona dispersió.
- Deixar 30 min en gel com a mínim, invertint de tant en tant.
- Centrifugar a 10000 rpm durant 10 min a 4°C (rotor GSA, Sorvall).
- Descartar el sediment.
- Afegir àcid tricloroacètic (TCA) al sobrenadant fins a una concentració final del 25% (33.3 ml). D'aquesta manera fem precipitar la proteïna.
- Deixar en gel d'una hora a O.N. invertint de tant en tant.
- Centrifugar a 10000 rpm durant 15 min a 4°C (rotor GSA, Sorvall).
- Descartar sobrenadant. El sediment de proteïna precipitada pot ser fluid, per tant hem d'anar amb compte al decantar.
- Resuspendre en 30 ml de H_2O i passar a tubs Corex de vidre.
- Afegir TCA fins a concentració final del 25%.
- Deixar 30 min en gel (invertir de tant en tant).
- Centrifugar durant 10 min a 4°C (rotor SS34, Sorvall).
- Descartar el sobrenadant.
- Rentar el sediment amb 60 ml d'acetona acidificada (acetona contenint 10 mM HCl) freda.
- Centrifugar a 10000 rpm durant 10 min a 4°C (rotor SS34, Sorvall).
- Descartar el sobrenadant amb compte. Millor fer-ho just acabar de centrifugar i amb l'ajut d'una pipeta Pasteur.
- Rentar el sediment amb 60 ml d'acetona freda (-20°C).
- Centrifugar a 10000 rpm durant 10 min (rotor SS34, Sorvall).
- Descartar el sobrenadant amb compte.
- Rentar el sediment amb 60 ml d'acetona freda (-20°C).
- Centrifugar a 10000 rpm durant 10 min (rotor SS34, Sorvall).

- Descartar el sobrenadant amb compte.
- Assecar al buit el sediment.
- Guardar el sediment sec o bé redissoldre'l en el volum i tampó adient per a la seva purificació. Conservar a -20°C.

3.2.4. Purificació per cromatografia

Un cop realitzat el tractament del sediment bacterià descrit a l'apartat anterior, les proteïnes HMGA no es troben encara pures sinó barrejades amb d'altres proteïnes d'origen bacterià. A més a més, en el cas de la proteïna HMGA1a i HMGA1b, com s'explicarà a l'apartat de resultats, apareixen també formes truncades per l'extrem C-terminal degut a l'activitat proteolítica residual de la cèl·lula hoste. Caldrà també separar aquestes formes truncades de la proteïna recombinant ja que per a poder cristal·litzar la proteïna serà necessari que sigui el més pura possible. Per a poder assolir el grau de puresa requerit es va recórrer a la cromatografia de bescanvi iònic. En el cas de les proteïnes HMGA1b(1-79) es va intentar arribar a un grau major de puresa afegint un pas més de purificació amb una cromatografia d'exclusió molecular.

Cromatografia de bescanvi iònic

La cromatografia és un procés en el qual els components d'una barreja, dissolts en una fase mòbil, es desplacen amb diferent velocitat a través d'una fase estacionària. En el cas de la cromatografia d'intercanvi iònic la fase estacionària conté grups carregats que interaccionen electrostàticament amb ions de signe contrari a la fase mòbil. Els suports o matrius que constitueixen la fase estacionària carregada poden ser polímers sintètics (resines) o polímers naturals. Si la matriu té grups funcionals d'unió carregats positivament, es tracta d'un sistema de bescanvi aniònic i els contraions de la fase mòbil són anions. De la mateixa manera, el bescanvi catiònic s'aconsegueix amb una matriu amb grups funcionals d'unió amb càrrega negativa i els contraions són cations.

En el cas de la proteïna HMGA1a, al tractar-se d'una proteïna amb càrrega neta positiva, s'ha utilitzat la cromatografia de bescanvi catiònic. En aquest treball s'ha utilitzat la matriu MACRO-PREP 25S (Bio-Rad) seguint el protocol descrit per Reeves i Nissen [81]. També s'ha fet ús d'una columna d'intercanvi catiònic dèbil amb matriu de carboximetil cel·lulosa, la CM52 (Whatman), per a purificar les proteïnes HMGA1b i HMGA1b(1-79).

La matriu s'empaqueta dins d'una columna (de vidre generalment en el nostre cas) que disposa en el seu extrem inferior d'un filtre permeable capaç de retenir la matriu però permetre el flux de la fase mòbil. L'empaquetament de la matriu a la columna s'ha fet seguint les instruccions del fabricant.

La fase mòbil s'impulsa a través de la fase estacionària mitjançant una bomba peristàltica que, a més, ens assegura un flux constant.

La cromatografia de bescanvi iònic es desenvolupa en tres etapes ben diferenciades:

- 1) *Equilibrat de la fase estacionària.* En aquesta etapa els grups carregats de la columna estan units als contraions (cations en el nostre cas).
- 2) *Aplicació de la mostra i rentat.* La mostra ha d'estar en la mateixa dissolució emprada en l'equilibrat de la columna. Convé centrifugar la mostra abans de carregar-la a la columna per a eliminar possibles restes de material insoluble que podrien obturar la columna. Les proteïnes amb càrrega neta positiva podran desplaçar els contraions i s'uniran a la matriu de la columna. Les proteïnes amb la mateixa càrrega que la matriu no s'uneixen i són eluïdes durant el rentat.
- 3) *Elució de les proteïnes unides a la matriu.* Es pot fer per canvis de pH o bé de força iònica (més habitualment). Així, si fem passar per la columna dissolucions amb concentracions creixents del contraió, eluiran primer les proteïnes menys carregades positivament i finalment les de major càrrega positiva, unides més fortament a la matriu.

Per a poder tornar a utilitzar la columna cal reequilibrar-la passant almenys 10 volums de la dissolució d'equilibrat per volum de matriu empaquetada a la columna.

Les dissolucions utilitzades rutinàriament per a purificar les proteïnes HMGA1 han estat les que es recullen a la figura següent.

Bescanvi iònic (Catiònic)	
Columna	Macro-Prep 25S (Biorad)
	CM52 (Biorad)
Tampons	Equilibrat: 50 mM KCl, 25 mM KH ₂ PO ₄ , pH 7 Gradient: concentracions creixents de KCl
	Equilibrat: 50 mM NaCl, 25 mM Tris, pH 7 Gradient: concentracions creixents de NaCl

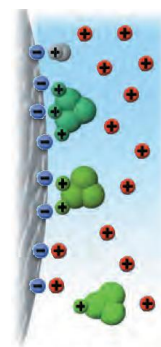


Fig. 3.4. Detalls de la cromatografia de bescanvi catiònic utilitzada i dibuix esquemàtic del principi de separació d'aquesta columna.

L'aplicació de concentracions creixents a la columna es pot fer de manera gradual (gradient continu) o bé de forma esglaonada (gradient isocràtic). El gradient continu és eficaç per a separar soluts que tendeixen a sortir en un interval de concentracions d'eluent properes, mentre que el discontinu ens permet separar fàcilment altres contaminants de naturaleses més diferents. Durant el present treball s'han emprat les dues estratègies (v. apartat 4.1).

El volum eluït de la columna es fracciona mitjançant un col·lector de fraccions automàtic programat per a recollir un volum determinat en cada tub.

Mesurant l'absorbància a 220 nm, regió on absorbeix l'enllaç peptídic entre els aminoàcids de les proteïnes, s'obté el corresponent cromatograma. També es pot mesurar l'absorbància a 276 nm on absorbeixen aquelles proteïnes que presentin aminoàcids amb grups aromàtics. Per a comprovar que no existeixi contaminació per part d'àcids nucleics es pot també controlar l'absorbància a 260 nm.

Les fraccions que presenten pics d'absorbància a 220 nm s'analitzen electroforèticament (veure apartat 3.6.1). Aquelles fraccions corresponents a la proteïna d'interès es reuneixen, es dialitzen i es concentren immediatament (veure apartat 3.5) o bé es conserven a -20°C fins al seu ús.

Cromatografia d'exclusió molecular

La cromatografia d'exclusió molecular o també anomenada filtració per gel s'ha dut a terme per fer el pas de poliment de la purificació de les mostres d'HMGA1 obtingudes de la cromatografia de bescanvi catiònic.

Aquesta cromatografia permet separar diferents molècules segons la seva massa molecular. La matriu de la columna està formada per micropartícules esfèriques poroses de polímers altament hidrofílics com polisacàrids o poliacrilamides. Aquest entramat d'esferes amb diferents mides de porus permetrà l'entrada a les molècules petites retenint la seva elució de la columna. En canvi, les molècules grans seran incapaces d'entrar dins dels porus i per tant passaran a través de la columna més fàcilment. És important destacar que, en aquest cas, a diferència de les columnes de bescanvi iònic, la proteïna no s'uneix a la matriu de la columna i per tant la composició del tampó no afecta directament a la resolució (grau de separació entre pics del cromatograma). D'aquesta manera les mostres s'elueixen isocràticament, és a dir, amb el mateix tampó per a l'equilibrat i l'elució. A la fig. 3.5 es descriuen les condicions d'aquesta cromatografia.

Exclusió molecular (Filtració per gel)	
Columna	Superdex 75 10/300 GL (GE Healthcare)
Tampons	Equilibrat i elució: 100 mM NaCl, 10 mM Tris-HCl, pH 7,5

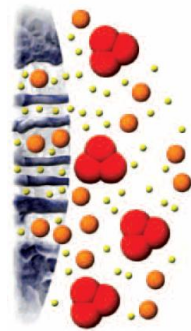


Fig. 3.5. Detalls de la cromatografia d'exclusió molecular utilitzada i dibuix esquemàtic del principi de separació d'aquesta columna.

Per a dur a terme una cromatografia d'exclusió molecular, un cop carregada la mostra que es vol purificar és suficient passar una mica més d'un volum del tampó per volum de matriu empaquetada a la columna (cv), normalment passem 1,5 cv.

La resolució de la cromatografia i per tant la bona separació entre la nostra proteïna i altres impureses depèn de molts factors com el volum de la mostra, el quocient entre el volum de la mostra i la columna, les dimensions de la columna, la mida de les micropartícules i la seva distribució i densitat, la mida dels porus, el flux i la viscositat de la mostra. En el nostre cas per a la cromatografia d'exclusió molecular s'ha triat la columna Superdex 75 (GE Healthcare) (amb 25 ml de matriu ja empaquetada), que és una columna d'alta selectivitat amb un rang de separació de 3000-70000 Mw adient a la mida de les proteïnes d'interès. De forma general s'ha carregat un volum de no més d'un 0.5% del volum empaquetat per tal d'optimitzar la resolució. Per això ha estat necessari concentrar les mostres eluïdes de la cromatografia de bescanvi catiónic abans de realitzar la cromatografia d'exclusió molecular.

S'ha treballat amb un equip de cromatografia de mitja pressió FPLC (*Fast Protein Liquid Chromatography*) AKTA Purifier 10 (GE Healthcare). Aquest equip permet el carregat automàtic de la mostra a la columna, la dosificació dels tampons a la columna, la recollida de les fraccions i la lectura de les absorbàncies a 220, 260, 280 nm durant el desenvolupament de la columna.

De la mateixa manera que anteriorment, es realitzarà l'anàlisi electroforètica de les mostres eluïdes de la columna i s'ajuntaran i concentraran les fraccions més pures (veure apartat 3.6.1 i 3.5.1 respectivament).

3.3. Obtenció de les proteïnes HMGB

3.3.1. Expressió en *E. coli*

Les proteïnes HMGB van ser expressades i purificades durant l'estada de doctorat al laboratori de la Dra. M. E. A. Churchill (*Department of Pharmacology, University of Colorado*). Les bases per a l'obtenció d'aquestes proteïnes a partir d'un cultiu bacterià són les mateixes que en el cas de les HMGA (veure apartat 3.2.2) i es va fer servir la mateixa soca de bacteris *E. coli*, Rosetta(DE3)pLysS. La diferència més significativa en l'obtenció de les proteïnes HMGB és la utilització del sistema de fusió del gen de la Glutatió S-transferasa (*Glutathione S-transferase Gene Fusion System*). Aquesta tècnica consisteix en l'expressió de la nostra proteïna d'interès, l'HMGB, fusionada a una proteïna coneguda, en el nostre cas la glutatió S-transferasa (GST), per tal de facilitar la seva purificació per cromatografia d'afinitat. Aquest tipus de constructe constituït per la proteïna d'interès amb una cua proteica s'anomena proteïna de fusió. La introducció de cues per a facilitar la purificació és una tècnica molt utilitzada per a l'obtenció de proteïnes recombinants. Un cop purificada per afinitat la proteïna de fusió, es talla aquesta cua GST (v. 3.3.3) obtenint d'aquesta manera una mostra de la proteïna d'interès força pura.

Així doncs, per tal d'expressar la nostra proteïna de fusió fem servir un plasmidi que codifica la proteïna de fusió HMGB amb una cua N-terminal de glutatió S-transferasa (o cua GST). A la fig. 3.6 podem veure l'esquema del plasmidi que hem fet servir per a expressar els tres fragments de la proteïna HMGB1 (Box AS, Box BS i Box ABS). Els tres plasmidis ens van ser facilitats al mateix laboratori de la Dra. M.E.A. Churchill [82].

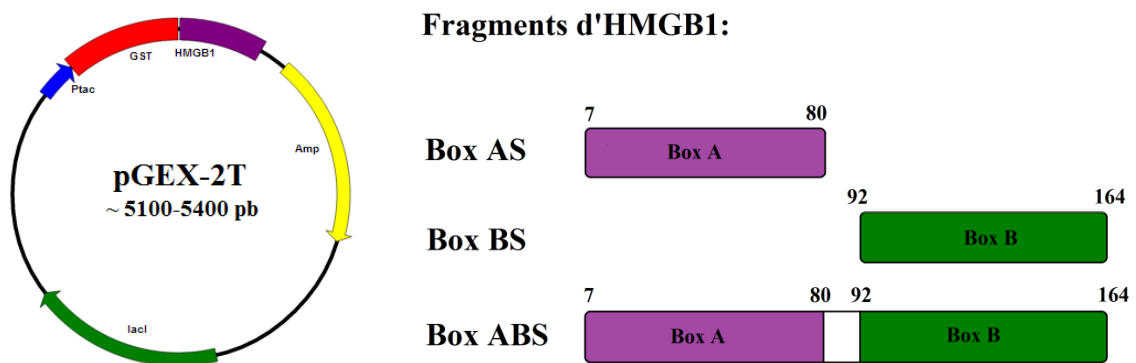


Fig. 3.6. Esquema del plasmidi pGEX-2T que codifica la proteïna de fusió GST-HMGB1 (esquerra) i dels fragments de la proteïna HMGB1 amb els que hem treballat (dreta).

Els tres plasmidis així construïts permeten expressar els tres fragments de l'HMGB1 amb una cua GST a l'extrem N-terminal. Box AS: HMGB1(Δ 7-80); Box BS: HMGB1(Δ 92-164); Box ABS: HMGB1(Δ 7-164). La figura del plasmidi ha estat preparada amb el programa Savvy [<http://www.bioinformatics.org/savvy/>].

El protocol per a l'expressió d'1 litre de cultiu es descriu a continuació i els medis de cultiu que hem fet servir en aquest protocol són els que es recullen a l'annex B.

Transformació

- Descongelar en gel les cèl·lules competents de la soca Rosetta(DE3)pLysS del *kit* comercial de Novagen i barrejar suament per tal d'assegurar que les cèl·lules estiguin resuspeses homogèniament.
- Aliquotar 20 µl de les cèl·lules competents en tubs eppendorf prèviament refredats. Agafem tants tubs com plasmidis tinguem més un per fer el control.
- Afegir 1 µl del DNA plasmídic (normalment 1-20 ng/µl) a tots els tubs amb cèl·lules excepte el control, i barrejar suament.
- Deixar els tubs en gel 5 min.
- Escalfar els tubs 30 s a 42°C al bany d'aigua (xoc tèrmic).
- Deixar en gel 2 min.
- Afegir 80 µl del medi de cultiu SOC a temperatura ambient a cada tub (InvitrogenTM).
- Incubar 1 hora a 37°C i 250 rpm.
- Preescalfar a 37°C les plaques contenint LB + 34 µg/ml Cam i 50 µg/ml Amp.
- Sembrar 50 µl de cèl·lules en la placa i deixar absorbir durant uns minuts.
- Invertir la placa i deixar *overnight* (O.N.) a 37°C.

Cultiu en medi líquid

- Picar una única colònia de bacteris transformats sobre un tub amb 3 ml LB + 34 µg/ml Cam i 50 µg/ml Amp.
- Incubar a 37°C i agitació (250 rpm) durant unes 6 hores (o fins observar terbolesa al tub).
- Abocar el cultiu sobre 150 ml LB + 34 µg/ml Cam i 50 µg/ml Amp.
- Incubar a 37°C i agitació (250 rpm) O.N.

Inducció de l'expressió

- Inocular suficient volum del cultiu anterior en 1 litre de LB + 34 µg/ml Cam i 50 µg/ml Amp per a assolir una $DO_{600} = 0,1$. Normalment uns 60 ml.
- Incubar a 37°C i agitació (250 rpm) fins $DO_{600} = 0,8$.
- Inducció amb IPTG 0,5 mM.

- Incubar a 37°C i agitació (250 rpm) durant 3 hores.
- Centrifugar a 5000 g durant 10 min a 4°C en un rotor F9-4x1000Y.
- Eliminar el sobrenadant.
- Congelar súbitament el sediment de bacteris amb nitrogen líquid i guardar a -80°C fins el seu processat.

De la mateixa manera que amb l'expressió de les proteïnes HMGA, agafarem alíquotes a temps 0 h, 1:30 h i 3 h post-inducció per a mesurar-ne l'absorbància i fer controls electroforètics del procés en gels d' SDS (v. 3.6.1). Així, a cada temps, farem:

- Separar 1 ml per a mesurar DO_{600} .
- Agafar una alíquota de 50 μ l.
- Centrifugar a 4000 rpm durant 5 min (micròfuga).
- Eliminar sobrenadant.
- Resuspendre el sediment en 15 μ l de tampó de mostres per a electroforesi.
- Guardar a -20°C fins al desenvolupament de l'electroforesi.

Estoc congelat de cèl·lules transformades

Quan expressem la mateixa proteïna diverses vegades ens pot interessar guardar uns estoc de cèl·lules ja transformades per a començar el procés en el moment del pas a medi líquid.

Per a obtenir aquest estoc fem:

- Picar una única colònia de bacteris transformats sobre un tub amb 3 ml LB + 34 μ g/ml Cam i 50 μ g/ml Amp.
- Incubar a 37°C i agitació (250 rpm) durant unes 6 hores (o fins observar terbolesa al tub).
- Congelar el cultiu bacterià amb glicerol (0,9 ml de cultiu + 0,3 ml glicerol 60% autoclavat).
- Guardar a -80°C.

Iniciariem l'expressió de la proteïna a partir d'un cultiu de bacteris transformats i congelats amb glicerol així:

- Rascar o bé fondre uns pocs microlitres de la superfície de l'estoc congelat amb la punta blanca d'una pipeta.
- Inocular sobre 3 ml LB medi contenint 34 μ g/ml Cam i 50 μ g/ml Amp.

- Incubar a 37°C i agitació (250 rpm) durant unes 6 hores (o fins observar terbolesa al tub).
- Abocar el cultiu sobre 150 ml LB + 34 µg/ml Cam i 50 µg/ml Amp.
- Incubar a 37°C i agitació (250 rpm) O.N.
- Continuar amb la inducció de l'expressió.

3.3.2. Tractament del sediment bacterià

L'extracció de la proteïna de fusió de dins de les cèl·lules bacterianes es realitza en dues etapes. En primer lloc es porta a terme la lisi dels bacteris i així l'extracció de la proteïna recombinant al medi. Finalment es realitza una centrifugació que permet separar la fracció soluble (on es troba la proteïna d'interès) de les restes insolubles.

En aquest treball s'han lisat les cèl·lules per sonicació. La ruptura cel·lular, en aquest cas, és produïda pel fenomen de la cavitació gasosa degut a l'acció d'ones ultrasonores amb una freqüència de vibració de l'ordre de 50 kHz.

El protocol que hem realitzat és el següent:

- Resuspendre el sediment bacterià de 1000 ml del cultiu en 25 ml del tampó d'equilibrat de la columna d'afinitat (v. fig. 3.8) amb inhibidors de proteasa (1/2 comprimit de cOmplete Protease Inhibitor Cocktail Tablets, Roche) i una punta d'espàtula de DNasa en tubs falcon siliconats.
- Sonicar el sediment resuspès dos cops durant 2 minuts amb 50% d'output (Fisher Scientific™ Model 500 Sonic Dismembrator). És molt important mantenir les mostres en gel durant el tractament d'ultrasons per tal d'evitar l'escalfament de la mostra.
- Centrifugar a 15000 rpm durant 30 minuts a 4°C (rotor SS34, Sorvall).
- Descartar el sediment i carregar el sobrenadant a la columna d'afinitat.

3.3.3. Purificació de les proteïnes de fusió GST-HMGB

A l'arribar a aquest punt, tenim una mostra amb la nostra proteïna d'interès barrejada amb d'altres proteïnes d'origen bacterià. Per tal d'obtenir-la pura, en primer lloc es durà a terme una cromatografia d'afinitat aprofitant l'especificitat d'unió de la cues GST de la proteïna de fusió. Tot seguit com es veurà al següent apartat es realitzaran una cromatografia de bescanvi catiònic i una d'exclusió molecular per arribar a l'alt grau de puresa requerit per tal de cristal·litzar la proteïna.

Les proteïnes amb cues GST es poden purificar fàcilment amb una columna d'afinitat fent servir una matriu de sefarosa amb la molècula de glutatió immobilitzada. Hem fet servir la matriu Glutatió Sefarosa de GE Healthcare. D'aquesta manera, al carregar la mostra a la columna, la cua GST de la proteïna s'uneix a aquest lligand. Llavors es fa un rentat per a eliminar les impureses o d'altres proteïnes de la mostra de partida no unides a la columna. A continuació per tal de tallar la cua GST i així obtenir la nostra proteïna d'interès es fa una digestió amb una proteasa que talla específicament al lloc desitjat. En el nostre cas es farà servir l'enzim trombina que talla específicament la seqüència aminoacídica LVPR↓GS, situada entre la cua de la glutatió S-transferasa i el fragment d'HMGB1. Al fer l'elució de la columna obtindrem una mostra contenint la proteïna HMGB1 sense la cua i l'enzim trombina que separarem en el proper pas de la purificació. Finalment caldrà regenerar la columna de Glutatió Sefarosa, és a dir, portar-la a les condicions inicials per a poder utilitzar-la un altre cop. Ho farem desplaçant les cues GST unides a la matriu de la columna amb clorur de guanidina. A la fig. 3.7 es poden observar de forma esquemàtica els passos seguits per a la cromatografia i a la fig. 3.8 es detallen les dissolucions utilitzades en la purificació.

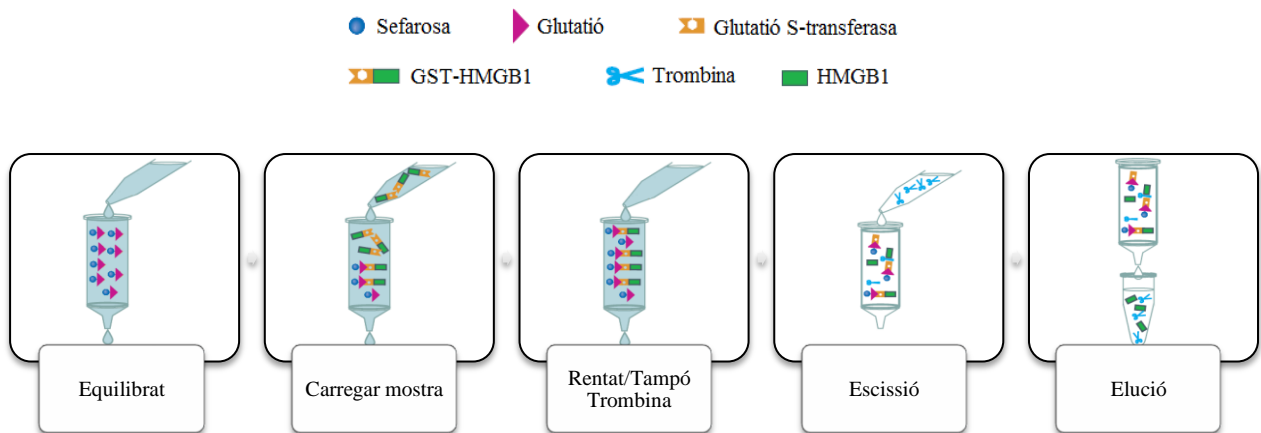


Fig. 3.7. Esquema de la purificació de la proteïna HMGB amb una columna de glutatió sefarosa.

Adaptat de GST Gene Fusion Handbook (GE Healthcare).

Afinitat	
Columna	Glutatió Sefarosa 4B (GE Healthcare)
Tampons	Equilibrat: 500 mM NaCl, 10 mM EDTA, 1 mM DTT, 20 mM Tris pH=7,9 Rentat: 1 M NaCl, 10 mM EDTA, 1 mM DTT, 20 mM Tris pH=7,9 T. trombina: 100 mM NaCl, 2,5 mM CaCl ₂ , 1 mM DTT, 20 mM Tris pH=7,9 Elució: 100 mM NaCl, 10 mM EDTA, 1 mM DTT, 20 mM Tris pH=7,9 Regeneració: 6 M Clorur de guanidina

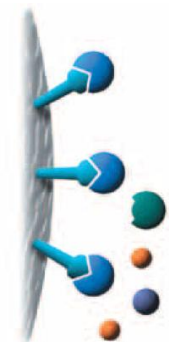


Fig. 3.8. Detalls de la cromatografia d'afinitat Glutatió Sefarosa i dibuix esquemàtic del principi de separació d'aquesta columna.

La matriu que hem fet servir ha estat empaquetada en una columna de vidre Bio-Rad però a diferència de les cromatografies explicades anteriorment no hem fet servir un flux continu per a l'elució de les mostres de la columna. En aquest cas el mètode de cromatografia és *semi-batch*, on la mostra es barreja durant un cert temps amb la matriu de la columna per a facilitar la unió i l'elució es fa deixant sortir l'eluent per gravetat.

El protocol fet servir és el següent (per un litre de cultiu):

- Equilibrar la matriu (microesferes o resina) de Glutatió Sefarosa fent passar tres vegades 25 ml del tampó d'equilibrat. Hem fet servir uns 2,5 ml de resina per litre de cultiu.
- Incubar el sobrenadant de la sonicació sobre la matriu durant 2 hores a 4°C amb rotació.
- Eluir el *flow through* (FT).
- Rentar amb 5 volums del tampó de rentat (passar els volums de columna d'un en un). Guardar l'eluent (W).
- Rentar amb 3 volums del tampó trombina. Guardar eluent (ThW).
- Tallar la cua GST amb 200 U de trombina (Fisher Bioreagents) en 1 ml del tampó trombina. Deixar incubant amb rotació *overnight* a 4°C.
- Afegir EDTA 0,5 M fins arribar a una concentració final de 10 mM per tal d'acabar la reacció.
- Eluir el FT de l'escissió (FTc).
- Eluir la proteïna amb 3 volums del tampó d'elució (EIW1, EIW2, EIW3).
- Regenerar la matriu de la columna passant 3 volums d'aigua, 3 volums de 6 M clorur de guanidina i 3 volums d'aigua. A continuació, si la columna no s'ha de fer servir en un temps, fer passar 3 volums d'etanol i conservar així.

Totes les mostres eluïdes s'analitzaran electroforèticament (v. 3.6.1) per a assegurar que el procés de purificació s'ha realitzat correctament. Les fraccions eluïdes després de l'escissió (FTc, EIW1, EIW2, EIW3) s'ajuntaran i es carregaran en una columna de bescanvi catiònic com s'explica a continuació (v. 3.3.4).

3.3.4. Cromatografia de bescanvi catiònic i d'exclusió molecular

Gràcies a la columna d'afinitat les mostres de proteïna HMGB1 s'obtenen força pures. Tot i així cal separar-les de l'enzim trombina que queda a les fraccions eluïdes de la columna Glutatió Sefarosa i d'altres impureses. Per això seguidament realitzem una columna de bescanvi catiònic. En aquest cas fem servir una columna HiTrap SP FF (GE Healthcare) de 5 ml de volum de resina empaquetat. Aquesta columna, adquirida ja empaquetada, té com a matriu sefarosa contenint el grup sulfopropil (SP) com bescanviador catiònic fort. Les cromatografies s'han realitzat en un sistema de purificació *AKTA Purifier* fent servir els tampons que es recullen a la fig. 3.9.

Bescanvi iònic (Catiònic)	
Columna	HiTrap SP FF (GE Healthcare)
Tampons	Equilibrat: 100 mM NaCl, 10 mM Tris pH=7,9, 1 mM EDTA, 1 mM DTT Gradient: concentracions creixents de NaCl (fins a 1M)

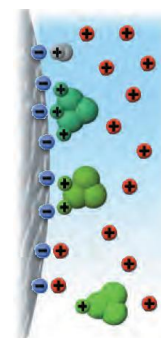


Fig. 3.9. Detalls de la cromatografia de bescanvi catiònic per a les proteïnes HMGB i dibuix esquemàtic del principi de separació d'aquesta columna.

Durant la cromatografia es llegiran les absorbàncies a 220, 260 i 280 nm i s'analitzaran electroforèticament els pics que presentin major absorbància a 280 nm (a diferència de l'HMGA, la proteïna HMGB sí que presenta grups aromàtics). Les fraccions amb major contingut de proteïna pura s'ajuntaran i concentraran (veure apartat 3.5.1). Tot seguit es realitzarà una columna d'exclusió molecular (descrita en més detall a l'apartat 3.2.4) per a acabar de purificar la mostra. En aquest cas s'han fet servir columnes HiLoad Superdex 75 i 200, específiques per a purificar proteïnes amb masses moleculars d'entre 3000 i 70000 Da i d'entre 10000 i 600000 Da respectivament, sent la primera la que hem triat preferentment. Els tampons fets servir per aquestes cromatografies es detallen a la fig. 3.10. Finalment, després de verificar la puresa de les fraccions eluïdes per electroforesi, s'ajuntaran i es concentraran les fraccions del pic d'aquesta cromatografia. Per a assegurar una mostra homogènia i pura es trien les fraccions de la part central del pic. De manera rutinària hem disminuït la concentració de sal per a conservar les mostres concentrades. Es podria fer per diàlisi però amb les proteïnes HMGB1 hem optat per fer un canvi de tampó en el procés de concentració (com s'explicarà a l'apartat 3.5.1). Així, pels assajos cristal·logràfics, concentrarem les mostres com a mínim a 10 mg/ml en el tampó 50 mM NaCl, 25 mM HEPES pH 7,4, 1 mM DTT i les conservarem aliquidades a -80°C.

Exclusió molecular (Filtració per gel)	
Columnna	HiLoad Superdex 75 16/600 PG (GE Healthcare)
	HiLoad Superdex 200 16/600 PG (GE Healthcare)
Tampons	Equilibrat i elució: 150 mM NaCl, 50 mM HEPES pH 7,4, 1 mM DTT

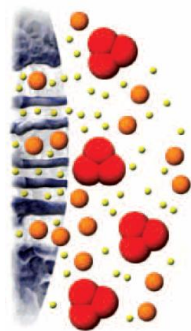


Fig. 3.10. Detalls de la cromatografia d'exclusió molecular per a les proteïnes HMGB i dibuix esquemàtic del principi de separació d'aquesta columna.

3.4. Obtenció de la proteïna NHP6A

3.4.1. Expressió en *E. coli*

El plasmidi per a l'obtenció de la proteïna de llevat NHP6A va ser cedit gentilment pel professor Reid Johnson del *Department of Biological Chemistry* de la *University of California*. La proteïna NHP6A s'ha expressat amb un sistema pET (en el plasmidi pET-11a) en la soca BL21(DE3) del bacteri *E. coli* [83] (veure fig. 3.11).

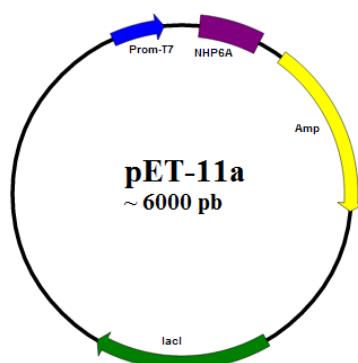


Fig. 3.11. Esquema del plasmidi pET-11a que codifica el gen de la proteïna de llevat NHP6A.

Figura preparada amb el programa Savvy [<http://www.bioinformatics.org/savvy/>].

El protocol per a l'expressió i la purificació d'aquesta proteïna ha estat adaptat del treball del grup d'investigació del Dr. Johnson [84]. Els medis de cultiu que hem utilitzat es recullen a l'annex B.

Transformació

- Descongelar en gel les cèl·lules competents de la soca BL21(DE3) del *kit* comercial d'Invitrogen i barrejar suaument per tal d'assegurar que les cèl·lules estiguin resuspeses homogèniament.
- Aliquotar 20 µl de les cèl·lules competents en dos tubs eppendorf prèviament refredats (un serà el control).
- Afegir 1 µl del DNA plasmídic (normalment 1-20 ng/µl) sobre les cèl·lules i barrejar suaument.
- Deixar els tubs en gel 5 min.
- Escalfar els tubs 30 segons a 42°C al bany (xoc tèrmic).
- Deixar en gel 2 min.

- Afegir 80 μ l del medi de cultiu SOC InvitrogenTM o bé 250 μ l de LB (Fisher) a temperatura ambient a cada tub.
- Incubar 1 hora a 37°C i 250 rpm.
- Preescalfar a 37°C les plaques contenint LB + 50 μ g/ml Amp.
- Sembrar 50 μ l de cèl·lules en la placa i deixar absorbir durant uns minuts.
- Invertir la placa i deixar *overnight* (O.N.) a 37°C.

Cultiu en medi líquid

- Picar una única colònia de bacteris transformats sobre un tub amb 3 ml LB + 50 μ g/ml Amp.
- Incubar a 37°C i agitació (250 rpm) durant unes 6 hores (o fins observar terbolesa al tub).
- Abocar el cultiu sobre 150 ml LB + 50 μ g/ml Amp.
- Incubar a 37°C i agitació (250 rpm) O.N.

Inducció de l'expressió

- Inocular suficient volum del cultiu anterior en 1 litre de LB + 50 μ g/ml Amp per a assolir una $DO_{600} = 0,1$. Normalment uns 20 ml.
- Incubar a 37°C i agitació (250 rpm) fins $DO_{600} = 0,6$.
- Inducció amb IPTG 1 mM.
- Incubar a 37°C i agitació (250 rpm) durant 3 hores.
- Centrifugar a 5000 g durant 10 min a 4°C en un rotor F9-4x1000Y.
- Eliminar el sobrenadant.
- Congelar súbitament el sediment de bacteris amb nitrogen líquid i guardar a -80°C fins al seu processat.

Agafarem alíquotes a temps 0 h, 1:30 h i 3 h per a veure l'absorbància i fer controls electroforètics.

3.4.2. Tractament del sediment bacterià

Com en el cas de les HMGB1, per a extreure la proteïna NHP6A de dins dels bacteris, hem sotmès la mostra a ultrasons i després hem fet una centrifugació a alta velocitat. Concretament, hem fet:

- Resuspendre el sediment bacterià d'1 litre de cultiu en 25 ml del tampó 500 mM NaCl, 50 mM HEPES pH=8, 1 mM EDTA, 1m M DTT amb inhibidors de proteases (1/2 comprimit de cOmplete Protease Inhibitor Cocktail Tablets, Roche) i una punta d'espàtula de DNasa en tubs falcon siliconats.
- Sonicar el sediment resuspès dos cops durant 2 minuts amb 50% d'output (Fisher Scientific™ Model 500 Sonic Dismembrator). És molt important mantenir les mostres en gel durant el tractament d'ultrasons per tal d'evitar l'escalfament de la mostra.
- Centrifugar a 15000 rpm durant 30 minuts a 4°C (rotor SS34, Sorvall).
- Descartar el sediment i realitzar una precipitació fraccionada del sobrenadant amb sulfat d'amoni (veure apartat següent).

Per a la purificació d'aquesta proteïna hem seguit els passos descrits a [85] amb lleugeres modificacions. Concretament, hem realitzat una precipitació de la proteïna, una cromatografia de bescanvi aniònic seguida d'una de catiònic i finalment una cromatografia d'exclusió molecular.

3.4.3. Precipitació amb sulfat d'amoni

Després de la sonicació i la centrifugació hem obtingut una fracció soluble amb la nostra proteïna d'interès (NHP6A), d'altres proteïnes d'origen bacterià, lípids i àcids nucleics, principalment. Per obtenir un extracte proteic enriquit en la nostra proteïna s'ha portat a terme una precipitació seqüencial amb sulfat d'amoni, $(\text{NH}_4)_2\text{SO}_4$.

Aquesta tècnica es basa en la separació de macromolècules mitjançant la precipitació de forma selectiva d'aquestes mitjançant l'addició de sulfat d'amoni. La solubilitat d'una proteïna a una baixa concentració d'ions augmenta a mesura que afegim sal, fenomen que es coneix com a *salting in*. No obstant això, si s'afegeix molta més quantitat de sal, la solubilitat de la proteïna disminueix novament. Aquest efecte de *salting out* és resultat de la competència per les molècules d'aigua dels ions afegits i les proteïnes en solució. És a dir, a l'afegir major quantitat de sal les molècules d'aigua solvataran els ions deixant els grups carregats de les proteïnes exposats, així podran interaccionar entre elles, donant lloc a la seva agregació i precipitació.

Així afegirem el sulfat d'amoni al sobrenadant de la sonicació en una concentració just per sota del punt de precipitació de la proteïna d'interès. Després de la centrifugació d'aquesta mostra les proteïnes precipitades no desitjades es descarten. A continuació s'afegeix més sal al sobrenadant per a precipitar la proteïna desitjada. Després d'aquesta centrifugació es recupera la proteïna en el sediment i es descarta el sobrenadant contenint els àcids nucleics. Com no coneixem el punt exacte de saturació en el que precipita la proteïna NHP6A, precipitarem a tres punts estàndards (30%, 60% i 100%).

Per conèixer els grams de sal de sulfat d'amoni que s'ha d'afegir per arribar a un tant per cent de saturació determinat es fa ús de taules o fórmules. En el nostre cas hem fet servir el programa de la web d'*EnCor Biotechnology Inc* que fa servir la següent equació, a la vegada extreta de "*Current Protocols in Protein Science*" (2001) [86].

$$m = V_i \frac{C_{sat}(M_f - M_i)}{M_{sat} - \left(\frac{v}{1000} \cdot 132,14 \left(\frac{g}{mol} \right) \cdot M_{sat} M_f \right)}$$

Eq. 3.1. Equació per al càlcul de la massa de sulfat d'amoni necessària per a arribar a un determinat percentatge de saturació.

Sent:

m : massa de sulfat amoni que cal afegir (g)

V_i : volum inicial (l)

C_{sat} : concentració en el punt de saturació (519,1 g/l a 4°C)

M_f : molaritat de la solució final (M)

M_i : molaritat de la solució inicial (M)

M_{sat} : molaritat en el punt de saturació (3,93 M a 4°C)

v : volum específic (0,54 ml/g)

S'ha de tenir en compte que el volum de la dissolució augmenta a l'afegir la sal d'amoni i que la seva solubilitat depèn de la temperatura, és a dir que els valors de la concentració, la molaritat en el punt de saturació i el volum específic varien amb la temperatura.

El protocol seguit és (treballant sempre a 4°C):

- Afegir la quantitat de sulfat d'amoni sòlid necessària per arribar a un 30% de saturació tenint en compte el volum actual. Barrejar durant una hora a 4°C.
- Centrifugar a 10000 g 15 minuts a 4°C (Rotor SS34, Sorvall). Redissoldre el sediment en 15 ml de 50 mM NaCl, 50 mM HEPES pH=8, 1 mM EDTA, 1m M DTT i guardar a 4°C.
- Sobre el sobrenadant afegir la quantitat necessària de sulfat d'amoni sòlid per arribar a una saturació del 60%. Barrejar durant una hora a 4°C.

- Centrifugar i redissoldre el sediment de la mateixa manera que en el pas anterior.
- Sobre aquest últim sobrenadant afegir la quantitat necessària de sulfat d'amoni sòlid per a arribar a una saturació del 100%. Barrejar durant una hora a 4°C.
- Centrifugar i redissoldre el sediment de la mateixa manera que en el pas anterior.
- Guardar aquest últim sobrenadant a 4°C.

Serà necessari analitzar electroforèticament les mostres obtingudes durant la precipitació per a verificar el contingut en NHP6A dels precipitats redissolts i de l'últim sobrenadant. A continuació es dialitzaran les mostres que continguin la proteïna d'interès per tal d'eliminar l'excés de sals i deixar en el tampó d'equilibrat de la primera columna de bescanvi, que és 50 mM NaCl, 50 mM HEPES pH=8, 1 mM EDTA, 1m M DTT. En el nostre cas hem dialitzat els precipitats a 60% i 100% de saturació i hem descartat el precipitat al 30% i el sobrenadant final.

3.4.4. Purificació per cromatografia

Les fraccions al 60% i al 100% de saturació procedents de la precipitació contenen encara un alt grau d'impureses d'altres proteïnes de naturaleses diferents. Per això realitzarem una columna de bescanvi aniònic i tot seguit una de catiònic. La proteïna bàsica NHP6A serà eluïda al carregar la mostra a la columna de bescanvi aniònic en el FT (*flow through*), que presenta una matriu carregada positivament. D'aquesta manera quedaran unides a la columna les proteïnes àcides i així podrem separar-les de la mostra inicial. Hem fet servir la columna Toyopearl DEAE 5PW (Tosoh Bioscience). Aquesta columna de bescanvi dèbil presenta una resina de TSK-gel amb grups funcionals de dietilaminoetil (DEAE) immobilitzats.

Bescanvi iònic (Aniònic)	
Columna	TSK-Gel Toyopearl MD-G DEAE 5PW (Tosoh Bioscience)
Tampons	Equilibrat: 50 mM NaCl, 50 mM HEPES pH=8, 1 mM EDTA, 1 mM DTT Gradient: concentracions creixents de NaCl (fins a 1M)

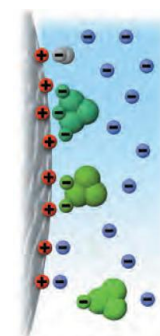


Fig. 3.12. Detalls de la cromatografia de bescanvi aniònic utilitzada per a la proteïna NHP6A i dibuix esquemàtic del principi de separació d'aquesta columna.

El FT d'aquesta cromatografia presentarà altres proteïnes de naturalesa bàsica que caldrà separar en una segona cromatografia, en aquest cas de bescanvi catiònic. Fem servir una columna HiTrap SP FF (GE Healthcare) com en el cas de les HMGB1, aplicant un gradient de condicions creixents de sal per tal d'eluir de manera fraccionada les proteïnes bàsiques.

Finalment es realitzarà una cromatografia d'exclusió molecular per acabar d'obtenir una mostra prou pura pels assajos cristal·logràfics.

Les tres cromatografies (bescanvi aniònic, catiònic i exclusió molecular) s'han realitzat en un sistema de purificació *AKTA Purifier* monitoritzant les absorbàncies a 220, 260 i 280 nm i fent servir els tampons que es recullen a la fig. 3.12, fig. 3.13 i fig. 3.14, respectivament. Les mostres s'han seleccionat i ajuntat visualitzant les fraccions de cada columna en gels SDS-PAGE i s'han concentrat per a disminuir-ne el volum (v. 3.5.1) prèviament a carregar-se a cadascuna.

Bescanvi iònic (Catiònic)	
Columna	HiTrap SP FF (GE Healthcare)
Tampons	Equilibrat: 50 mM NaCl, 50 mM HEPES pH=8, 1 mM EDTA, 1 mM DTT Gradient: concentracions creixents de NaCl (fins 1M)

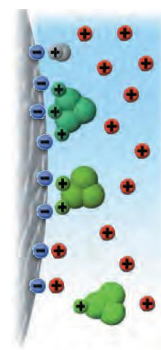


Fig. 3.13. Detalls de la cromatografia de bescanvi catiònic utilitzada per a la proteïna NHP6A i dibuix esquemàtic del principi de separació d'aquesta columna.

Exclusió molecular (Filtració per gel)	
Columna	HiLoad Superdex 75 16/600 PG (GE Healthcare) Superdex 75 10/300 GL (GE Healthcare)
Tampons	Equilibrat i elució: 150 mM NaCl, 50 mM HEPES pH 7,4, 1 mM DTT

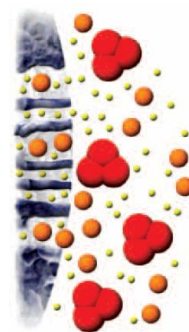


Fig. 3.14. Detalls de la cromatografia d'exclusió molecular utilitzada per a la proteïna NHP6A i dibuix esquemàtic del principi de separació d'aquesta columna.

3.5. Altres tècniques per a l'anàlisi de proteïnes

3.5.1. Concentració per ultrafiltració

Els dispositius de concentració permeten la concentració i purificació de fluids biològics i solucions macromoleculares a través d'una membrana anisotròpica mitjançant la centrifugació. Al llarg d'aquest treball ha estat necessària la concentració per ultracentrifugació de les proteïnes HMG amb dues finalitats. En primer lloc, per a disminuir el volum de les mostres per carregar en columnes d'exclusió molecular (on el volum de la mostra afecta a la resolució d'aquesta) i en segon lloc per a obtenir mostres un cop pures suficientment concentrades per a cristal·lització. En alguns dels casos, a més a més de concentrar la mostra inicial, s'han fet servir els concentradors per a dessalar les mostres de proteïna, és a dir, canviar el tampó de la mostra per un altre diferent o amb menys sals.

El procés comença introduint la mostra al dispositiu concentrador i a continuació es centrifuga. La força centrífuga impulsa a través de la membrana els solvents i els soluts de baixa massa molecular fins al vial de filtrats, mentre que els soluts d'alta massa molecular (com la nostra proteïna) queden retinguts per la membrana. La mida dels porus de la membrana (o MWCO, *Molecular Weight Cut Off*) és una característica específica del concentrador i s'ha de triar segons la mida de la proteïna HMG que es vol concentrar. Per a finalitzar, depenent del tipus de concentrador que es fa servir, la mostra concentrada es transfereix a un tub recollidor de concentrat girant el dispositiu o bé es pipeteja la mostra directament del recipient de reserva. Tot aquest procés es realitza a 4°C.

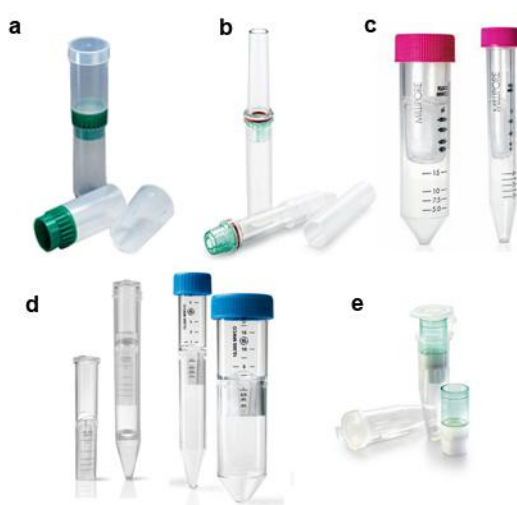


Fig. 3.15. Els diferents concentradors fets servir al llarg d'aquest treball.

a) Centriplus. b) Centricon. c) Amicons-ultra. d) Vivaspins. e) Microcons.

En funció del volum de mostra, de la massa molecular de la proteïna que es vol concentrar i del volum al que es vol arribar hem necessitat un tipus de dispositiu o altre. A la fig. 3.15 s'observen els diferents concentradors utilitzats i a la taula 3.5 les principals especificacions de cadascun i amb quina proteïna o fragment de les HMG s'han utilitzat.

El procediment és el següent:

- Rentar els concentradors passant mig volum de la capacitat del concentrador de la solució en la que tenim la mostra (centrifugant a la velocitat especificada pel proveïdor). Buidar.
- Centrifugar les mostres abans de concentrar-les, a 7500 g 10 min a 4°C.
- Omplir els concentradors amb les mostres fins el volum màxim especificat.
- Centrifugar a la velocitat adient fins que la major part del volum de mostra hagi passat al recipient col·lector del filtrat. Extreure el filtrat, afegir més mostra i tornar a centrifugar. Seguir aquest procediment fins assolir el grau de concentració desitjat.
- Per a recuperar la mostra concentrada:

Acoblar el Centriplus/Centricon/Microcon el tub recollidor de concentrat, girar i centrifugar a baixa velocitat.

En el cas dels Vivaspins i Amicons pipetejar la mostra directament.

Taula 3.5. Especificacions dels concentradors utilitzats i proteïna concentrada.

Nom comercial	Tipus	MWCO	Capacitat ¹	Max. força centrífuga ¹	Volum final límit	Proteïna concentrada
Centriplus (Millipore)	YM3 yellow	3000	10 ml (15 ml)	3000 g	500 µl	HMGA1b(1-79)
Centricon (Millipore)	YM10 green	10000	2 ml	5000 g	25 µl	HMGA1a
Vivaspin (GE Healthcare o Sartorius stedim biotech)	20	3000	16 ml (20 ml)	8000 g (5000 g)	50 µl	HMGB1 Box AS i BS, NHP6A
	20	5000	16 ml (20 ml)	8000 g (5000 g)	50 µl	HMGB1Box ABS
	6	3000	6 ml (6 ml)	10000 g (4000 g)	30 µl	HMGA1b(1-79)
	4	5000	4 ml (4 ml)	10000 g (4000 g)	20 µl	HMGA1b(1-79) HMGA1b
Amicon ultra (Millipore)	4	3000	3,5 ml (4 ml)	7500 g (4000 g)	50 µl	NHP6A
	4	5000	3,5 ml (4 ml)	7500 g (4000 g)	50 µl	HMGB1Box ABS
Microcon (Millipore)	YM3 yellow	3000	0,5 ml	14000 g	5 µl	HMGB1 Box AS, BS i ABS, NHP6A

¹Valors per a rotor d'angle fix (o rotor basculant); els centricons i microcons només es fan servir amb rotor d'angle fix.

3.5.2. Determinació de la concentració

La determinació de la concentració de les proteïnes HMGA1 i HMGB1 s'ha fet mesurant l'absorbància a 220 nm i a 280 nm, respectivament, i utilitzant l'equació de Lambert-Beer. En el cas de les HMGA1 es realitza la mesura a 220 nm (enllaç peptídic) ja que aquestes proteïnes no presenten grups aromàtics que absorbeixin a 280 nm.

$$A = \varepsilon \cdot C \cdot l$$

Eq. 3.2. Equació de Lambert-Beer.

Sent:

A: absorbància.

ε : factor d'extinció a la longitud d'ona considerada (veure annex A).

l: pas de llum (1 cm pels espectrofotòmetres estàndards i 1 mm o 0,2 mm pel Nanodrop).

C: concentració de la mostra.

Per tant, coneixent el coeficient d'extinció podem determinar la concentració a partir del valor de l'absorbància de la mostra. Per a llegir les absorbàncies es fa servir un espectrofotòmetre. Posem en una cubeta de quars un volum de 800 μl del tampó en el que es troba la mostra (fem el blanc amb aquest mateix tampó) i afegim 1 μl de la mostra. Llegim l'absorbància a la longitud d'ona desitjada. Podem afegir un altre microlitre i tornar a llegir l'absorbància per tal de tenir dos valors de la concentració i treballar amb el promig. Durant l'estada a la *University of Colorado* l'espectrofotòmetre que es va utilitzar va ser un Nanodrop. En aquest cas només és necessari aplicar 1,8 μl de la mostra de la que es vol saber la concentració en el suport per a mostres de l'aparell.

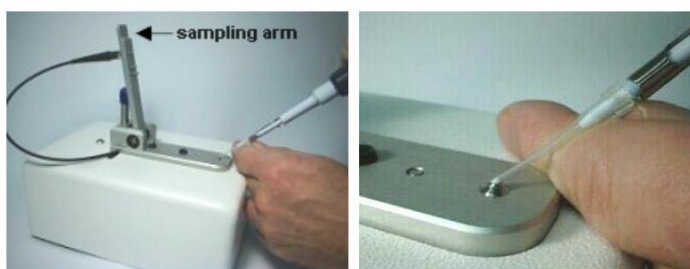


Fig. 3.16. Imatge d'un Nanodrop.

Podem veure un detall del suport on es diposita la mostra per a llegir-ne l'absorbància.

3.5.3. Diàlisi

La tècnica de la diàlisi permet canviar la dissolució en la qual està la mostra proteica per una altra de diferent composició. És molt habitual usar aquesta tècnica per a dessalar les mostres.

Es necessita una bossa de membrana semipermeable d'un diàmetre de porus concret que ens determina la massa molecular mínima (MWCO) que serà retinguda per la membrana. S'han utilitzat les membranes de la casa Spectra/Por[®] (Spectrum[®]) amb un límit de massa retinguda de 6-8000 MWCO per a les proteïnes HMGA1a i HMGA1b i de 3500 MWCO per al fragment HMGA1b(1-79). Per a la proteïna NHP6A s'ha utilitzat una membrana de la casa comercial Fisherbrand amb 3500 MWCO.

Un cop tallat el tros de membrana necessari cal deixar-la en aigua destil·lada durant 20 minuts per tal d'hidratar la membrana i extreure'n les restes de glicerol. Passats els 20 minuts, les fraccions proteiques s'introdueixen en la membrana i es tanquen els extrems amb pinces especials de la casa Spectrum.

La mostra, dins la bossa de diàlisi adequada, es submergeix en un recipient que contingui la nova dissolució. Cal realitzar suficients canvis de la solució del tampó de diàlisi i utilitzar agitació ja que així s'arriba a l'equilibri més ràpidament. Les diàlisis es realitzen a 4°C per a minimitzar el perill de degradació de la proteïna.

Per tal d'obtenir una mostra per a purificar en una cromatografia hem fet servir com a tampons de diàlisi els tampons corresponents a la primera solució tampó de la columna. En canvi, per a la preparació de mostres per a cristal·lografia hem fet servir un tampó amb un contingut baix en sals.

3.6. Tècniques electroforètiques per a proteïnes

3.6.1. Electroforesi en gels de poliacrilamida-SDS (SDS-PAGE)

L'electroforesi és una tècnica que es basa en la migració diferencial de molècules carregades sota l'acció d'un camp elèctric que ens permet separar i visualitzar proteïnes i àcids nucleics. Els elements necessaris per a portar a terme una separació electroforètica són en primer lloc una font de tensió que proporciona un camp elèctric mitjançant dos elèctrodes (càtode i ànode) entre els que es produeix una diferència de potencial. També cal una cubeta on es situen els elèctrodes en els seus extrems, un suport electroforètic (un gel de poliacrilamida en el nostre cas) i un tampó d'electroforesi. Durant el procés la mostra que inicialment ocupa la part superior del gel es desplaça segons la seva càrrega cap a la part inferior impulsada pel camp elèctric.

Els gels de poliacrilamida s'obtenen gràcies a l'entrecruament de llargues cadenes de monòmers d'acrilamida amb N,N'-metilè-bisacrilamida. En funció de la concentració d'acrilamida s'obtenen entramats més o menys densos. La reacció de polimerització s'inicia amb la formació de radicals lliures del monòmer, que es produeixen per radicals lliures d'oxigen, degut a l'acció d'ions persulfat. Les amines terciàries com el N,N,N',N'-tetrametilen-diamina (TEMED) s'utilitzen com a catalitzadors d'aquesta reacció, ja que causen la formació de radicals lliures del persulfat. Aquesta reacció està fortament inhibida pels alts nivells d'oxigen, per tant la solució ha de ser desgasificada per a aconseguir una formació de gel reproducible. Els monòmers (en pols o en dissolució) són neurotòxics per absorció cutània o inhalació. Una vegada s'ha donat la polimerització, la toxicitat és molt reduïda però encara hi pot haver presència d'alguns monòmers lliures.

Normalment, hem fet servir en aquest treball gels de poliacrilamida-SDS utilitzant el sistema discontinu de tampons introduït per Laemmli el 1970 [87]. Aquest tipus de gels consta d'una petita part de gel empiladora o *stacking* de baix tant per cent d'acrilamida seguida de la part separadora o *resolving* de major tant per cent d'acrilamida. L'SDS (dodecil sulfat de sodi) és un detergent aniònic que s'utilitza per a desnaturalitzar proteïnes. L'SDS revesteix les proteïnes i els hi confereix càrrega neta negativa proporcionalment a la seva massa. S'uneix una molècula d'SDS cada 2 aminoàcids (1,4 g d'SDS per gram de proteïna aproximadament). Així, les proteïnes migraran cap al pol positiu de manera inversament proporcional a la seva mida.

Una unió uniforme de l'SDS és crítica per a una bona migració electroforètica; per això s'escalfen les proteïnes a 100°C en presència d'un excés d'SDS. L'SDS s'afegeix tant als gels com al tampó de mostra i de cubeta.

Muntatge i preparació dels gels

L'electroforesi en placa és la forma més habitual i se sol desenvolupar en posició vertical. Per a preparar el motlle del gel, s'utilitzen dos vidres entre els quals es col·locaran uns separadors de gruix variable (normalment de 0,75 mm o 1,5 mm) que determinaran el gruix del gel. Per a evitar que el sistema vessi, s'han de segellar els dos extrems laterals i l'inferior amb l'ús d'una goma adequada. A la part superior del gel hi posarem la pinta de tefló del mateix gruix que els separadors i amb un nombre variable de dents. Una vegada acabada la polimerització, retirarem la pinta i obtindrem una sèrie de pous on es dipositaran les mostres que s'han d'analitzar.

Després de fer el muntatge, es preparen les barreges d'acrilamida (taula 3.6). Es comença per la part inferior o gel separador. Una vegada preparada la barreja, s'aboca entre els vidres fins 0,5 cm per sota del nivell de la pinta. S'afegeix una fina capa d'isopropanol per sobre del gel separador, d'una banda per a evitar el contacte amb l'oxigen, que alenteix la polimerització, i d'altra banda perquè quedi una superfície ben plana. Una vegada polimeritzada la part inferior del gel, s'aboca l'isopropanol, es neteja amb aigua i s'asseca amb paper de filtre. Ara ja es pot preparar la part superior o gel empilador. Una vegada abocat entre els vidres, s'hi col·loca la pinta.

El gel polimeritzat es col·loca en la cubeta d'electroforesi que contindrà el corresponent tampó (taula 3.7).

Preparació de la mostra

La mostra es prepara barrejant-la amb el tampó de mostres (taula 3.8) el qual conté un colorant per a poder visualitzar la migració. Se li afegeix un agent reductor per a desfer complexos que estiguin units per ponts disulfur i també glicerol per a donar densitat a la barreja. Per a permetre una bona unió de l'SDS convé bullir la mostra durant 3-6 minuts abans de carregar-la al gel. El volum de mostra que podem carregar dependrà de la mida dels pous. La quantitat de proteïna aplicada hauria de ser de l'ordre d'1-20 µg. L'aplicació de la mostra es fa amb una xeringa Hamilton. Quan la banda de colorant (front de migració) es troba a 1 cm del final, l'electroforesi es pot donar per finalitzada. Habitualment s'han desenvolupat les electroforesis a un voltatge constant d'entre 100-130 V.

Detecció de les proteïnes

Per tal de visualitzar les proteïnes separades en el gel hem de tenyir-les. Fem servir la tinció amb blau de Coomassie, isopropanol i àcid acètic, que permetrà la fixació de les proteïnes al gel i la tinció d'aquestes.

Al finalitzar l'electroforesi, separem el gel de les plaques de vidre i el submergim en la solució de colorant (taula 3.9) i amb agitació. La tinció sol durar 60 minuts. Després de la tinció, tot el gel queda tenyit de color blau intens i cal destenyir-lo per a eliminar l'excés de colorant. Se submergeix el gel en la solució destenyidora (taula 3.10) i es deixa destenyint tota la nit. D'aquesta manera obtindrem un gel transparent amb bandes blaves de proteïna.

Conservació dels gels

Per últim podem conservar el gel assecant-lo. Aquesta tècnica consisteix en posar una capa de cel·lofana mullada sobre un vidre, posar-hi el gel a sobre, folrar-lo amb una altra capa de cel·lofana mullada i subjectar-ho amb unes pinces. A continuació es deixa assecar un parell de dies.

Taula 3.6. Composició dels gels de poliacrilamida-SDS.

Gel poliacrilamida-SDS		
Reactius	Separador (10 ml) (15% Acrilamida, 0,4% Bisacrilamida, 0,375 M Tris, 0,1% SDS)	Empilador (3 ml) (6% Acrilamida, 0,16% Bisacrilamida, 0,125 M Tris, 0,1% SDS)
Solució A	5 ml	0,6 ml
Solució B	2,5 ml	-
Solució C	100 µl	30 µl
Solució D	-	0,75 ml
H₂O	2,3 ml	1,59 ml
TEMED	4,5 µl	3 µl
PSA 10%	75 µl	30 µl
<p><i>Solució A: acrilamida 30%, bisacrilamida 0,8%</i> <i>Solució B: 1,5 M Tris-HCl pH=8,8</i> <i>Solució C: 10% SDS</i> <i>Solució D: 0,5 M Tris-HCl pH=6,8</i></p>		

Reactius necessaris per a preparar un gel de poliacrilamida d'1,5 mm de gruix.

Taula 3.7. Composició del tampó de cubeta per a electroforesi en gels de poliacrilamida-SDS.

Tampó de cubeta per a gels de poliacrilamida-SDS		
Reactius	5x (1 litre)*	Concentració 1x
Tris	30 g	0,01 M
Glicina	144 g	0,075 M
SDS	5 g	0,02 %

*Normalment es prepara el tampó 5x i es dilueix per a fer servir el tampó 1x durant l'electroforesi.

Taula 3.8. Composició del tampó de mostres per a electroforesi en gels de poliacrilamida-SDS.

Tampó de mostres per a gels de poliacrilamida-SDS		
Reactius	4x (10 ml)	Concentració 1x
Tris 2 M pH 6,8	1,25 ml	62 mM
Glicerol	4 ml	10 %
SDS	0,8 g	2%
2-mercaptoetanol	2 ml	5 %
Blau de bromofenol	Punta d'espàtula	Concentració corresponent

Taula 3.9. Composició de la solució de tinció per a gels de poliacrilamida-SDS.

Solució de tinció per a gels de poliacrilamida-SDS		
Reactius	Per a 2 litres	Concentració
Isopropanol	500 ml	25 %
Àcid Acètic	200 ml	10 %
Blau de Coomassie G-250	5 g	0,25 %

Taula 3.10. Composició de la solució destenyidora per a gels de poliacrilamida-SDS.

Solució destenyidora per a gels de poliacrilamida-SDS		
Reactius	Per a 2 litres	Concentració
Isopropanol	200 ml	10 %
Àcid Acètic	200 ml	10 %

Al laboratori de la Dra. Churchill a Colorado la forma de preparar els gels ha diferit una mica de l'explicada fins ara, tot i que els fonaments són els mateixos. En aquest cas les solucions per a preparar els gels de poliacrilamida-SDS es venen mig preparades, per a agilitzar-ne la realització. Així només cal afegir TEMED i PSA a la solució comercialitzada per la casa Amresco per obtenir un gel sense empilador al 15% d'acrilamida (veure taula 3.11). Rutinàriament es realitzaven gels petits de 0.75 mm de gruix. De la mateixa manera el tampó de cubeta s'adquiria en solució concentrada 20 cops (Amresco) i només era necessari diluir-la. El tampó de mostres i les solucions tenyidora i destenyidora es realitzaven seguint les mateixes receptes ja comentades.

Taula 3.11. Composició dels gels de poliacrilamida-SDS (Colorado).

Reactius	15 % Gel poliacrilamida - SDS
NEXT GEL® 15 % Acrylamide (Amresco)	10 ml
TEMED	6 µl
PSA 10%	60 µl

En aquest cas la solució amb acrilamida ja es comercialitza preparada i només cal afegir el TEMED i el PSA. Reactius necessaris per a preparar dos gel de poliacrilamida petits (10 ml) de 0,75 mm de gruix.

3.6.2. Electroforesi de cristalls de proteïnes i pèptids

A vegades l'observació dels diagrames de difracció, ja sigui perquè són típics de DNA o no gaire bons, no ens permet saber si el cristall obtingut presenta contingut proteic. Molts cops és una informació necessària per a saber si aquell assaig de cristallització està ben encaminat (és a dir, variant lleugerament les condicions obtindrem un bon cristall) o si no mereix la pena continuar en aquella línia de condicions. Per a saber si un cristall conté proteïna o pèptid hem realitzat anàlisis electroforètiques.

Pels cristalls de proteïna es pot realitzar una anàlisi electroforètica en gel d' SDS-PAGE com s'ha explicat a l'apartat 3.6.1. Però pels cristalls de pèptids, que són cadenes aminoacídiques molt curtes, cal fer ús de gels amb un entramat de poliacrilamida major. La composició d'aquest gel es detalla a la taula 3.12. És una variant dels gels de poliacrilamida-àcid acètic-urea (AU-PAGE) desenvolupats per Panyim i Chalhley (1969) [88], però amb major percentatge d'acrilamida i amb dues fases.

Taula 3.12. Composició dels gels AU-PAGE per a cristalls de pèptids.

Gels AU-PAGE per a cristalls de pèptids		
Reactius	Separador (6 ml) (30% Acrilamida, 1% Bisacrilamida)	Empilador (1,6 ml) (4% Acrilamida, 0,2% Bisacrilamida)
60% Acril./2% Bisacril.*	3 ml	-
20% Acril./1% Bisacril.	-	320 µl
Àcid Acètic	0,75 ml	200 µl
Urea 9 M	1,66 ml	444 µl
Tiourea	5,25 mg	1,8 mg
H₂O₂	33,75 µl	7,5 µl
PSA 10%	0,58 ml	635 µl

Reactius necessaris per a preparar un gel de poliacrilamida de 0,75 mm de gruix. *Conservar aquesta solució a temperatura ambient per a evitar la seva precipitació.

El tampó de cubeta en aquest cas és una solució al 5% d'àcid acètic i la composició del tampó de mostres es troba a la taula 3.13. No hem afegit cap colorant a la mostra ja que aquests, al migrar a una velocitat semblant a la dels pèptids, podrien interferir en la visualització de les bandes. Per a controlar el front de migració, però, s'han carregat als dos pous extrems del gel tampó de mostres 1x amb els colorants catiónics pironina G (rosa) i verd de metil respectivament.

Taula 3.13. Composició del tampó de mostres per a electroforesi en gels de poliacrilamida-àcid acètic-urea.

Tampó de mostres per a gels de poliacrilamida-àcid acètic-urea		
Reactius	2x (10 ml)	Concentració 1x
Urea	4,8 g	4 M
Àcid acètic	0,5 ml	2,5 %
β-mercaptoetanol	14 µl	10 mM

La mida dels pèptids (de 7 a 9 residus en aquest treball) és també un inconvenient a l'hora de tenyir els gels per a visualitzar les bandes. Els pèptids no quedarien ben fixats al tenyir-los de manera convencional. És per això que hem realitzat una fixació dels pèptids més forta, amb el reactiu àcid pícric, que va ser descrita per Stephano l'any 1986 [89]. En primer lloc es porta a terme la fixació submergint el gel en la solució fixadora uns 30 minuts (v. taula 3.14), s'esbandeix amb aigua milli-Q i a continuació el gel es tenyeix amb plata (v. 3.6.3). S'ha optat per la tinció amb plata perquè és una tècnica de tinció més sensible i per tant ens permetrà la visualització de les bandes amb menys quantitat de proteïna en comparació a la tinció amb blau

de Coomassie. Aquesta tècnica s'ha utilitzat també per a la tinció dels cristalls de proteïna en gels d'SDS. Si els pèptids es trobessin en suficient quantitat en el gel es podrien tenyir en una solució de tinció que tingués la mateixa composició que la solució fixadora amb àcid pícric però a més a més amb Coomassie blue.

Taula 3.14. Composició de la solució de fixació amb àcid pícric i metanol.

Fixació amb àcid pícric	
Reactius	Fixador pícric-metanol (250 ml)
0,1 M Àcid pícric	175 ml
Metanol	67,5 ml
H₂O	7,5 ml

3.6.3. Tinció amb plata

La tinció amb plata és una tècnica molt estesa per a la tinció de proteïnes en gels de poliacrilamida. En aquest treball s'ha fet servir aquesta tècnica per a la visualització de proteïnes i pèptids en les electroforesis de mostres de cristalls i puntualment en gels EMSA (v. 3.10.1). El seu principal avantatge és la seva alta sensibilitat en comparació a la tinció amb Coomassie blue. El mecanisme en què es basa és la reducció dels ions plata Ag^+ a plata metàl·lica Ag^0 a les bandes de proteïna [90].

La tinció amb plata convencional és molt laboriosa, llarga i les bandes de proteïna queden poc contrastades. És per això que hem fet servir el *SilverQuest Silver Staining kit* de la casa comercial Invitrogen. Aquest *kit* permet detectar fins a 0,3 ng de proteïna sent 30 cops més sensible que el Coomassie blue. El protocol i els principis de cada pas es recullen a la taula 3.15. És important treballar amb aigua ultrapura, preparar els reactius a mesura que es van necessitant, evitar tocar el gel amb les mans o objectes metàl·lics, esbandir els guants utilitzats amb aigua desionitzada si es necessita tocar el gel, i treballar sempre amb material ben net.

Taula 3.15. Protocol per a la tinció amb plata amb el *SilverQuest Silver Staining kit* (Invitrogen).

Tinció amb plata		
Pas	Reactius	Descripció
Fixació	Fixador (40 % etanol, 10 % àcid acètic)	Eliminació d'ions i detergent del gel que poden interferir, i evitar la difusió de la proteïna fora de la matriu del gel
Rentat	Aigua Milli-Q	Eliminar l'excés de fixador
Sensibilització	Sensibilitzador (30% etanol, 10% sensibilitzador ¹)	Incrementar la sensibilitat i el contrast en la tinció
Rentat	Aigua Milli-Q	Eliminar l'excés de sensibilitzador i rehidratació del gel
Tinció	Tenyidor (1% tenyidor ¹)	Unió dels ions de plata a la proteïna i formació d'una imatge latent
Rentat	Aigua Milli-Q	Eliminar l'excés de tenyidor
Revelat	Revelador (10% revelador ¹ + 1 gota de potenciador ¹)	Reducció dels ions de plata a plata metàl·lica a les bandes de proteïna resultant en la seva visualització
Detenció	Aturador¹ (afegir a la solució anterior)	Complexar qualsevol plata lliure per a evitar major reducció

¹ Solucions proporcionades pel *kit*.

3.7. Preparació de les mostres d'oligonucleòtids

Els oligonucleòtids es reben de l'Institut Pasteur congelats i en forma de pols blanc, amb un grau de puresa d'entre un 95 i un 98%. És necessari dissoldre l'oligonucleòtid i després calcular la concentració per espectrometria UV abans de fer-los servir. Els primers oligonucleòtids utilitzats es dissolien en aigua, però més endavant es va optar per dissoldre'ls en 25 mM del tampó cacodilat de sodi ja que evita la degradació de la mostra gràcies al seu efecte antimicrobià. Pel càlcul de la concentració, en primer lloc es realitza un espectre d'absorció des de 220 nm fins a 320 nm amb 500 µl d'una dilució de la mostra d'oligonucleòtid, on observarem que el màxim es troba a 260 nm ja que és la longitud d'ona on absorbeixen les bases nitrogenades del DNA. A continuació amb el valor d'absorbància a 260 nm es calcula la concentració a partir de l'equació de Lambert-Beer (veure Eq. 3.2).

En aquest cas caldrà fer servir el coeficient d'extinció de l'oligonucleòtid a la longitud d'ona de 260 nm. Aquest coeficient d'extinció es calcula per a cada oligonucleòtid amb l'equació de Cantor *et al.* (1970) [91].

$$\varepsilon_{260} = \varepsilon_{DpEpFpGp...KpL} = 2(\varepsilon_{DpE} + \varepsilon_{EpF} + \varepsilon_{FpG} + \dots + \varepsilon_{KpL}) - \varepsilon_E - \varepsilon_F - \varepsilon_G - \dots - \varepsilon_K$$

Eq. 3.3. Equació de Cantor *et al* (1970).

Segons aquesta equació, el valor de ε_{260} d'un oligonucleòtid qualsevol, $DpEpFpGp...KpL$, es pot trobar a partir dels coeficients d'extinció de cada tipus d'enllaç entre dues bases (ε_{DpE} , ε_{EpF} , ε_{FpG} , ..., ε_{KpL}) i el de les bases dels nucleòtids no terminals (ε_E , ε_F , ε_G , ..., ε_K). Els valors d'aquests coeficients els trobem a la taula 3.16.

Taula 3.16. Coeficients d'extinció a 260 nm, 25°C, pH neutre i cadena senzilla de DNA [91].

DNA	Coef. d'extinció ε (l/mmol cm)	DNA	Coef. d'extinció ε (l/mmol cm)
pdA	15,4	dCpdG	9,0
pdC	7,4	dCpdT	7,6
pdG	11,5	dGpdA	12,6
pdT	8,7	dGpdC	8,8
dApdA	13,7	dGpdG	10,8
dApdC	10,6	dGpdT	10,0
dApdG	12,5	dTpdA	11,7
dApdT	11,4	dTpdC	8,1
dCpdA	10,6	dTpdG	9,5
dCpdC	7,3	dTpdT	8,4

S'ha de tenir en compte que a l'hora de fer els càlculs per a la preparació de gotes caldrà fer servir la concentració de la cadena doble. Per tant, haurem de dividir el resultat que obtenim entre dos.

Els valors dels coeficients d'extinció i les concentracions calculades per als estocs mare dels oligonucleòtids preparats al llarg d'aquest treball es recullen a la taula que es troba a continuació:

Taula 3.17. Coeficients d'extinció i concentracions dels oligonucleòtids preparats.

Oligonucleòtid	Coef. d'extinció ϵ_{260} (l/mmol cm)	Volum (μ l)	Concentració estoc inicial (mM)
CGCATTAAATGCG*	114,5	50	8,20
AATAATTATT (1)	107,2	100	3,66
AATAATTATT (2)		50	3,46
GGGAAATTTCCC	114,9	40	3,30
GTTATTAATAAC	106,0	50	4,06
AATAAATTTATT (1)	127,3	90	3,73
AATAAATTTATT (2)		60	4,52
ATAAATTTATTA	130,0	30	4,47
AATAAATTTATTT	135,4	80	3,66
ATATCGATAT	106,2	40	7,47
AATAATCGCGATTATT	163,1	30	2,55
ATATATATATATAT*	155,4	10	2,40
CATATATATG*	105,2	10	3,98
CGCGTTTTTCGCGAAAAACGCGG (1)	215,3	20	2,54
CGCGTTTTTCGCGAAAAACGCGG (2)		30	3,13
GAATTAACGTACGTTAATTCG (1)	212,3	15	1,1
GAATTAACGTACGTTAATTCG (2)		20	1,87

*En aquests casos els oligonucleòtids ja es trobaven dissolts i només es va calcular la seva concentració.

3.8. Cristal·lització dels complexos

Per tal de determinar l'estructura a nivell atòmic dels complexos proteïna-DNA la tècnica més usada és la difracció de raigs X, que és la que hem fet servir en aquest treball.

Cal partir de proteïna abundant i pura, després aconseguir cristalls d'aquesta proteïna amb els oligonucleòtids adients i optimitzar aquests cristalls. A continuació, amb la difracció de raigs X, tindrem un espectre de difracció que ens permetrà obtenir un mapa de densitat electrònica. Finalment caldrà preparar un model molecular de l'estructura compatible amb el mapa de densitat electrònica. L'anàlisi d'aquest model ens hauria de permetre un enteniment de la funció de la proteïna de la qual hem obtingut l'estructura.

Per tant, el primer pas pels estudis de difracció de raigs X és aconseguir cristalls adients per a sotmetre'ls a difracció. En aquest apartat s'abordan els fonaments que governen la formació de cristalls i la metodologia seguida per a l'obtenció de bons cristalls. A l'apartat següent s'explicaran els conceptes bàsics de la cristal·lografia, les bases de la difracció de raigs X de cristalls i el procés d'obtenció de l'estructura tridimensional dels complexos a partir dels diagrames de difracció.

3.8.1. Fonaments de la cristal·lització

Els cristalls de qualsevol substància es produeixen a partir d'una solució sobresaturada mitjançant l'addició ordenada de molècules de solut a un nucli inicial de cristal·lització. La condició de sobresaturació suposa, en el nostre cas, que la quantitat de proteïna-DNA o pèptid-DNA dissolta sigui superior a la que accepta la solució en el punt de saturació. Per tant, l'objectiu serà l'obtenció d'una solució sobresaturada que generi uns pocs nuclis de cristal·lització sobre els que creixin uns pocs cristalls, grans i ben ordenats.

Si la concentració del complex sobrepassa excessivament el punt de saturació, passem d'una situació metastable a una situació inestable, i si a més a més es realitza de forma abrupta, es produeix la precipitació del complex com a agregats amorfs, no ordenats, que impedeixen l'estudi amb la difracció de raigs X (fig. 3.17). Així, la sobresaturació s'ha de realitzar poc a poc i superar lleugerament la frontera de la situació de metaestable a inestable per tal d'obtenir cristalls ordenats. D'aquesta manera, es pretén que es generin uns pocs nuclis de cristal·lització i que l'excés de solut sigui reclutat, de forma lenta i ordenada, per tal d'engrandir els escassos cristalls naixents (i que no es formin nous nuclis de cristal·lització ni precipiti desordenadament en formes amorfes). L'estratègia perquè això es produeixi és introduir a la solució del complex

substàncies que disminueixin la seva solubilitat, generalment aconseguint una reducció gradual de la quantitat de solvent. En altres paraules, per tal d'arribar a la sobresaturació de la mostra cal augmentar gradualment la concentració del complex eliminant aigua i introduint canvis que afectin a la seva solubilitat (taula 3.18).

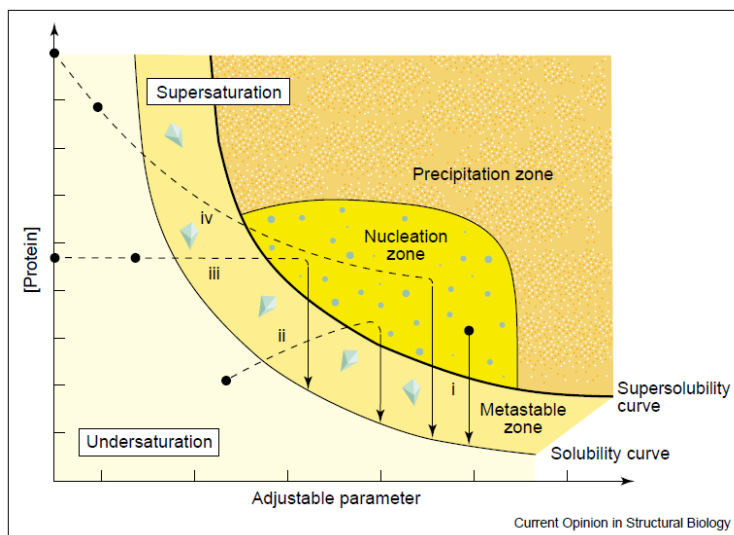


Fig. 3.17. Diagrama del procés de cristal·lització d'un cristall [92].

Podem observar en el diagrama la zona metastable, la zona de nucleació i la zona de precipitació. El paràmetre ajustable a l'eix d'abscisses pot ser la concentració de precipitant, el pH, la temperatura o d'altres. Cal sobresaturar la mostra per a passar de la zona metastable a la zona de nucleació lentament a fi que es produeixi la formació de nuclis de cristal·lització.

Taula 3.18. Factors que afecten la solubilitat del complex.

Factors que afecten la solubilitat
<ul style="list-style-type: none"> • Eliminació d'aigua per evaporació. • Canvis de pH. • Canvis de temperatura. • Baixa/alta força iònica. • Precipitants polimèrics, per exemple el polietilenglicol (PEG). • Solvents orgànics volàtils, per exemple l'etanol i l'isopropanol. • Solvents orgànics poc volàtils, per exemple el metilpentanodiol (MPD). • Lligands específics.

Per últim hem de tenir en compte que en la cristal·lització de complexos la relació de concentracions entre la proteïna i el DNA és un factor que pot determinar la seva unió i afectar la cristal·lització.

La cristal·lització de macromolècules és un procés molt complex en el que, com ja hem vist, intervenen molts factors. Normalment s'opta per abordar la cristal·lització d'una proteïna, complex o DNA realitzant assajos de prova i error variant les condicions de cristal·lització com s'explicarà posteriorment a la metodologia de cristal·lització (v. 3.8.2). En qualsevol cas hem de garantir un requisits mínims del complex abans d'iniciar els assajos de cristal·lització:

- Per tal d'obtenir una sobresaturació és molt important que la proteïna, el pèptid i el DNA de partida estiguin a una concentració elevada. Per a portar-ho a terme, en el cas de la proteïna, hem de concentrar la mostra amb ultrafiltració (v. 3.5.1). En el cas de l'oligonucleòtid i en el del pèptid, simplement els dissoldrem en un volum petit d'aigua. D'altra banda la quantitat de mostra de partida ha de ser elevada ja que per a la cristal·lització és necessari fer un gran nombre de proves.
- La puresa de la mostra ha de ser tan alta com sigui possible ja que les impureses poden impedir la cristal·lització o introduir irregularitats que disminueixin la qualitat del cristall obtingut. En el cas de la proteïna, aquesta puresa es determina amb els mètodes d'electroforesi (v. 3.6.1). Pel que fa al DNA i als pèptids, tots dos ens arriben amb un grau de puresa garantida d'un 98-99%, el primer de l'Institut Pasteur i el segon de la Universitat Pompeu Fabra.
- L'homogeneïtat de la mostra és també molt important. Les heterogeneïtats poden ser intrínseques de la mostra o degudes a la seva manipulació. En el nostre cas hem de separar perfectament les formes truncades de la proteïna de la forma sencera.
- Hem de garantir l'estabilitat de les mostres biològiques per a la seva conservació després de la purificació i en les condicions de cristal·lització. Per tant és necessari treballar sempre a temperatures baixes. És important també evitar el creixement de bacteris o fongs, per a fer-ho podem conservar el DNA en cacodilat.
- Verificar la formació del complex també és molt important. Mitjançant la realització d'assajos de canvi en la mobilitat electroforètica (EMSA) (v. 3.10.1) podem veure si és forma complex per aquella proteïna i oligonucleòtid i si aquesta unió és forta.

3.8.2. Metodologies de cristal·lització

La tècnica més comú d'obtenció de cristalls és la difusió de vapor. Aquesta tècnica consisteix en preparar una petita gota (normalment d'1 a 5 μ l) barrejant el complex proteïna/pèptid-oligonucleòtid amb el tampó de cristal·lització a parts iguals i mantenir-la dins d'una cambra hermètica. Aquesta cambra té un reservori que conté un volum de fins a un mil·lilitre del precipitant corresponent a una determinada concentració, sense que la gota i el reservori entrin en

contacte (v. fig. 3.18). La pressió de vapor d'aigua és major a la gota que al reservori. Per tant la gota perd poc a poc aigua fins que les concentracions de precipitant en la gota i en el reservori són iguals.

Existeixen principalment dues modalitats d'aquesta tècnica: la de gota penjada i la de gota assentada. En la modalitat de la gota penjada, la gota es disposa sobre un còbreeobjectes siliconat, que després s'inverteix i s'utilitza per a segellar el reservori fent servir gel de silicona, de manera que la gota queda dins del recinte segellat, separada de la solució de precipitant present al pou (veure fig. 3.18 a). En la segona modalitat, la gota assentada, la gota es disposa sobre una elevació còncava situada per sobre del nivell de la solució de precipitant (veure fig. 3.18 b).

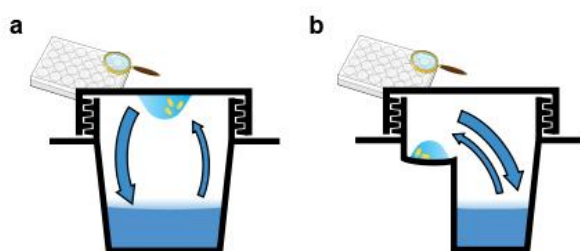


Fig. 3.18. Esquema de les dues modalitats de cristallització per difusió de vapor.

a) Gota penjada. b) Gota assentada.

Per a fer les gotes podem fer servir les clàssiques plaques Linbro de poliestirè de 24 pous (Hampton Research). Actualment aquest tipus de placa s'ha optimitzat en les anomenades plaques d'alt rendiment (*high throughput*) que permeten la realització de gotes de baix volum i es cobreixen amb làmines adhesives per a mantenir l'estanquitat. Podem trobar diferents tipus depenent del volum de gota (de 0,1 a 10 μl) i el nombre de pous (amb 24, 48 o 96). Les de menor volum estan pensades per a la realització automatitzada de gotes amb robots. Per últim podem fer ús de càpsules individuals (ETS; Caubere, França). A la fig. 3.19 es mostren aquests diferents sistemes. Al llarg d'aquest treball s'han fet servir plaques d'alt rendiment i càpsules.

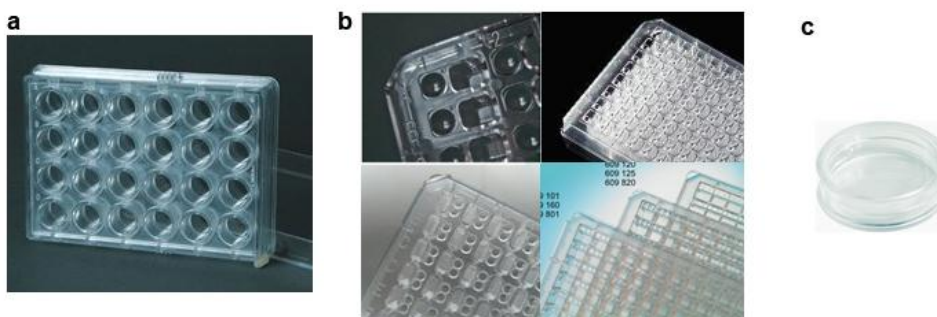


Fig. 3.19. Imatges d'una placa Linbro (a), de diferents plaques d'alt rendiment (b) i d'una càpsula (c).

Per tal de trobar les condicions òptimes de cristal·lització cal fer un gran nombre d'experiments en els que es variïn els diferents factors de cristal·lització (v. taula 3.18). Per aquest motiu s'han dissenyat estratègies de mostreig en el que s'explora tot l'espai de cristal·lització. Diferents empreses comercialitzen aquestes matrius de recerca incloent les condicions amb major èxit de cristal·lització descrites a la bibliografia. La primera d'aquestes matrius va ser la publicada per Jancarik i Kim [93]. En aquest treball, hem fet servir generalment els tampons Natrix HR2-116 de la casa comercial Hampton Research especialment dissenyada per a la cristal·lització d'àcids nucleics i complexos d'àcids nucleics i proteïnes [94]. Aquests tampons contenen diferents condicions de sals, precipitants polimèrics i solvents orgànics, a diferents pHs (v. annex D).

En els últims 10 anys s'han aconseguit molts avenços en el camp de la cristal·lització de proteïnes fent servir l'automatització i miniaturització dels assajos de cristal·lització. La utilització de robots per a la realització de gotes ha minimitzat els volums a nanolitres i ha disminuït considerablement el temps d'experimentació. Tot i així, la realització automatitzada de gotes en centres especialitzats és força cara i, si bé és cert que la realització d'assajos cristal·logràfics amb robots permet fer més quantitat de proves amb la mateixa quantitat de mostra i en menys temps, no assegura l'obtenció de cristalls òptims. En aquest treball s'han realitzat 4 plaques d'aquesta forma, dos plaques a la Universitat de Colorado amb el robot PHOENIX (Art Robbins Instruments) i la matriu comercial JCGS Core Suite I (Qiagen) [95] i dos plaques al Parc Científic de Barcelona amb el robot Microsys 4000 XL (Cartesian Technologies) i les matrius PAC9mod i PAC10 (preparades a la Plataforma Automatitzada de Cristal·lografia) (veure l'annex D). En el primer cas s'han realitzat gotes d'1 µl amb les proteïnes HMGB1 Box A i Box B i en el segon cas de 200 nl amb la proteïna NHP6A. A la fig. 3.20 hi ha les imatges dels dos robots.



Fig. 3.20. Fotografies del robots fets servir per a la realització d'assajos cristal·logràfics.
Cartesian (esquerra) i Phoenix (dreta).

Per tal de disminuir la solubilitat del complex proteïna-oligonucleòtid i intentar obtenir cristalls hem fet servir dues aproximacions. O bé anar augmentant gradualment la concentració de precipitant en el pou (en càpsules) o bé realitzar diferents assajos cadascun amb una concentració de precipitant més elevada al pou (especialment en les plaques). Normalment hem treballat a temperatura constant de 4°C.

La formació de cristalls es detecta mitjançant la microscòpia amb llum polaritzada. S'han de revisar les plaques o càpsules periòdicament i especialment després de dos dies d'haver fet canvis en les condicions de les gotes.

Quan es forma un cristall prou bo per a la difracció amb raigs X, és a dir un cristall amb alta birrefringència, costats rectes, sense esquerdes i uniforme, el congelem. Per a fer-ho agafem el cristall amb un *loop*, format per un llaç de fil de niló molt fi, i el submergim en nitrogen líquid. Aquest fil de niló està unit a la cara interior de la tapa d'un vial (fig. 3.21). El conservem dins del vial ple de nitrogen líquid, inserit en una canya que introduïm en un dipòsit amb nitrogen fins a la seva difracció.

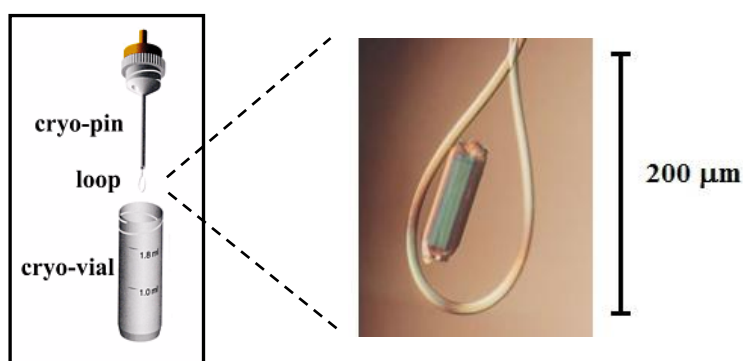


Fig. 3.21. Esquema d'un *loop* (esquerra) i detall amb un cristall (dreta).

És important una ràpida congelació dels cristalls per a aconseguir la vitrificació de l'aigua que envolta el cristall i la que es troba dins i així evitar la formació de gel i per tant l'aparició d'anells negres en els diagrames de difracció. En alguns casos s'ha fet servir la tècnica de la crioprotecció en la congelació de cristalls per tal d'evitar la formació de gel que dificulta l'anàlisi dels diagrames de difracció. Aquesta tècnica consisteix en submergir els cristalls en una solució crioprotectora de tal manera que el licor mare que l'envolta serà canviat per solució apropiada per a la vitrificació de l'aigua, a continuació els cristalls són congelats en nitrogen líquid com normalment. La composició d'aquesta solució crioprotectora (normalment precipitant, glicerol i la solució en la que es troba el cristall) influeix en la difracció del cristall i pot millorar la resolució de la difracció. No obstant això, també es pot córrer el perill que el cristall es fongui.

3.8.3. Optimització de cristalls

Algunes vegades el cristall obtingut és suficientment bo i de qualitat per a obtenir-ne l'estructura; altres cops és necessari optimitzar en nous assajos les condicions de cristallització dels cristalls obtinguts.

Després de realitzar un escombrat de condicions (amb una matriu comercial o una matriu preparada al laboratori) cal refinar les condicions que han produït els primers cristalls variant cada paràmetre com el pH, la concentració de precipitant, la concentració de complex, la temperatura o la concentració dels ions. A la fig. 3.22 s'observa un exemple d'estratègia d'optimització d'una condició positiva variant únicament la concentració de precipitant i el pH. Hi ha molts paràmetres i moltes consideracions a tenir en compte durant l'optimització per això aquest procés requereix paciència i observació.

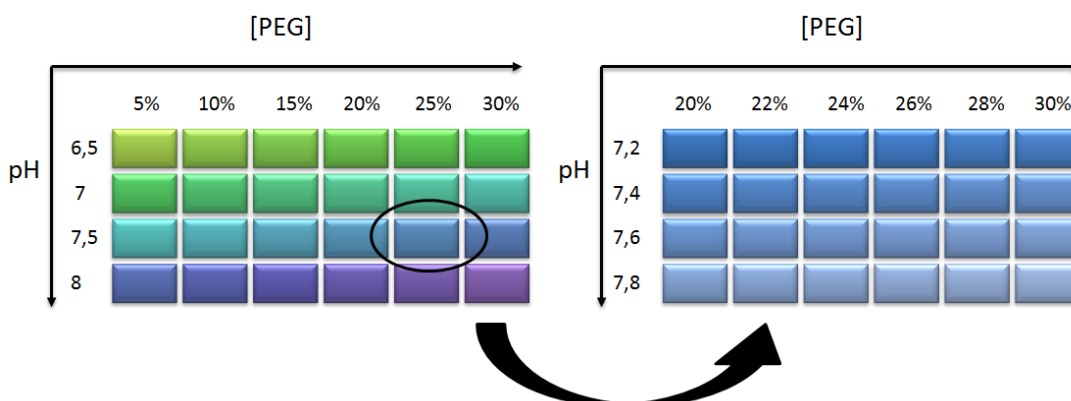


Fig. 3.22. Exemple d'una quadrícula de condicions per a l'optimització de cristalls.

En aquest cas es refinen les condicions de cristal·lització variant, per exemple, la concentració de precipitant (PEG) i el pH. Si la millor condició és 25%PEG i pH=7,5 s'acotaran els seus valors en la següent ronda d'assajos, i així successivament fins a obtenir el millor cristall.

S'ha d'arribar a un compromís entre els assajos d'optimització que es realitzen sobre una condició positiva en la que els cristalls obtinguts no són suficientment de qualitat com per a ser difractats i la realització de nous assajos explorant noves condicions amb, per exemple, altres sals, diferents seqüències d'oligonucleòtids o altres fragments de la proteïna d'interès.

Una manera d'aconseguir cristalls grans i ben ordenats, és el sembrat o *seeding*. Consisteix en introduir en la gota en fase metastable un cristall, un cristall molt o precipitat cristal·lí que actui com a nucli de cristal·lització. D'aquesta manera podem saltar-nos el pas de la nucleació, introduint aquests nuclis que creixen en la zona metastable fins a l'obtenció de cristalls aptes (veure fig. 3.17). Podem fer *seeding* de tres maneres: *macroseeding*, *microseeding* i *streak seeding* (fig. 3.23). En el *macroseeding* es transfereix un únic cristall normalment d'entre 5 i 50 μm prèviament rentat en una nova gota. En el *microseeding* s'obtenen solucions de microcristalls per dilució i aquests s'afegeixen a la gota. Per últim el *streak seeding* consisteix en agafar *seeds* tocant la superfície d'un cristall amb l'ajuda d'un *whisker* (bigoti de gat) i sembrar-los arrossegant-lo en diferents gotes noves de forma successiva.

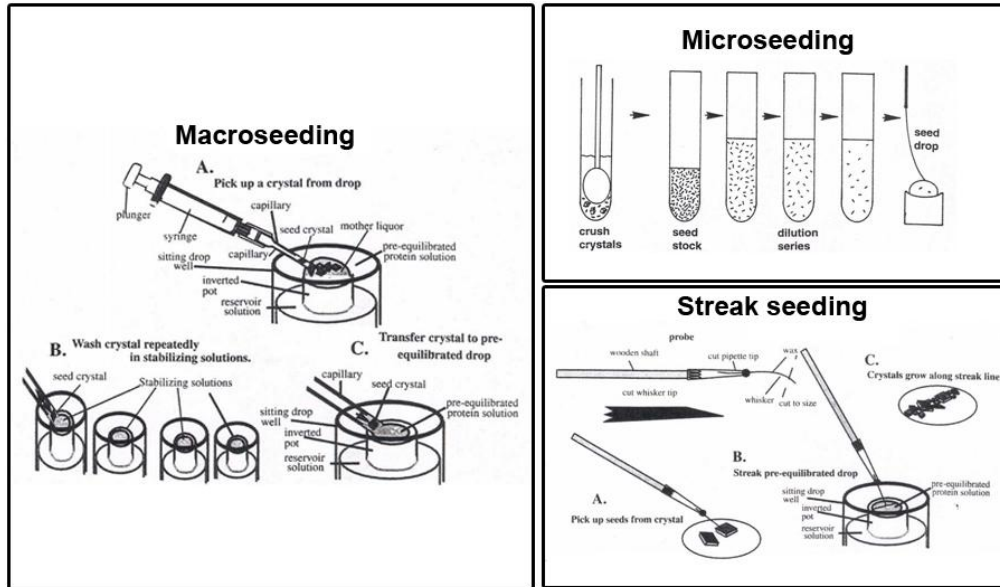


Fig. 3.23. Diferents tècniques de sembrat: *macro seeding*, *micro seeding* i *streak seeding*.

Imatge extreta de McPherson (1998) [97].

D'altra banda, si hem obtingut cristalls massa petits i en gran quantitat (degut a una ràpida cristal·lització, per exemple), podem fondre'ls i tornar-los a cristal·litzar, d'aquesta manera ja partim d'uns nuclis de cristal·lització i els nous cristalls que se'ns formen poden ser més grans i amb millor forma.

3.9. Difracció de raigs X i determinació de l'estructura

La cristal·lografia de raigs X és una de les tècniques físiques més potents per a la resolució de problemes estructurals de la matèria sòlida ordenada. L'estudi estructural correctament realitzat proporciona una fotografia tridimensional a escala atòmica del material cristal·litzat. A continuació farem un repàs dels fonaments dels cristalls i la seva simetria, explicarem el procés de difracció amb raigs X i per últim veurem com es tracten les dades de difracció per acabar obtenint l'estructura tridimensional.

3.9.1. Cristalls i simetria

Els materials cristal·lins tenen una estructura interna ordenada com una xarxa tridimensional en la que es repeteix una unitat geomètrica simple. Els tres elements que defineixen aquesta xarxa són la cel·la unitària, la unitat asimètrica i el grup espacial (fig. 3.24).

La *cel·la unitària* es defineix com el paral·lelepípede més petit que conté totes les operacions de simetria del grup espacial. Es treballa amb la longitud de cada vector d'aquest paral·lelepípede (a,b,c) i els angles que formen entre ells (α , β , γ) per a definir-la. Quan es treballa amb cel·les es considera que el contingut de la cel·la unitària està concentrat en un punt o més anomenats punts reticulars. Segons la localització d'aquests punts es diferencien 14 tipus de cel·les anomenades xarxes de Bravais (veure annex E).

S'entén per *unitat asimètrica (o motiu)* l'objecte que conforma la unitat fonamental de la cel·la unitària, de forma que qualsevol altra part del cristall pot aconseguir-se a partir d'ella mitjançant operacions de simetria. Pot estar formada per un o més àtoms, ions, molècules o parts de molècules.

El *grup espacial* és el conjunt d'operacions de simetria que s'han d'aplicar a la unitat asimètrica per a generar la cel·la unitària. A "International Tables for X-Ray Crystallography" [96] surten recollits els 230 grups espacials que podem obtenir. En el cas dels compostos orgànics es redueixen a 65 degut a la quiralitat que presenten aquestes molècules (veure annex E).

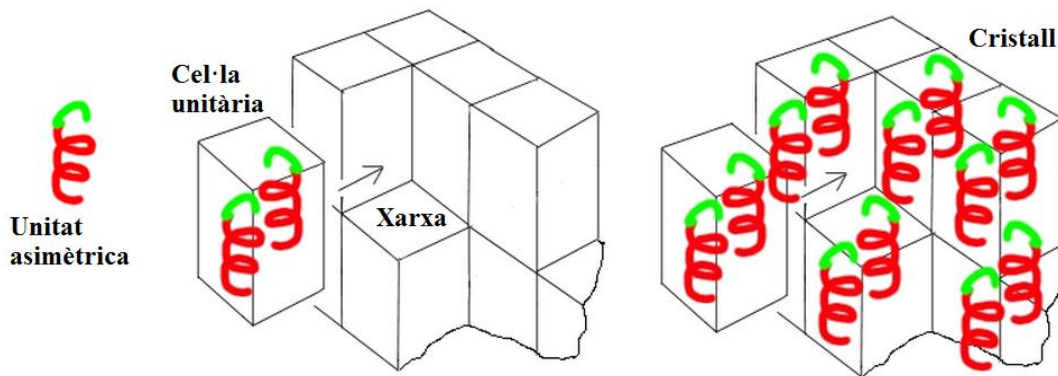


Fig. 3.24. Dibuix de la unitat asimètrica, la cel·la unitària, la xarxa i el cristall.

Adaptada de [<http://www.xtal.iqfr.csic.es/>].

Es fan servir els *índex de Miller* per a caracteritzar cada família de plans, són tres nombres enters (h,k,l) que designen els cops que els plans tallen cada costat de la cel·la. Concretament, es calculen considerant un pla qualsevol de la cel·la, es determina en quin punt talla cada eix cristal·logràfic, s'inverteix i s'escriu entre parèntesi.

La difracció de raigs X té lloc quan un raig incideix sobre la mostra en una determinada direcció en la qual els raigs difractats pels àtoms es combinen i reforcen. Per a descriure el fenomen de la difracció es pot fer ús de models físics i geomètrics basats, aquests últims, en el principi enunciat per Lawrence Bragg (1913), que estableix que la difracció de raigs X sobre la matèria es comporta com una reflexió òptica. Aquesta reflexió tindria lloc en uns plans virtuals de la xarxa cristal·lina, i així la difracció es produeix entre plans paral·lels distanciats entre si un valor d , per a un angle d'incidència θ , tal que la diferència del camí recorregut per les dues ones sigui igual a un múltiple de la longitud d'ona (veure fig. 3.25).

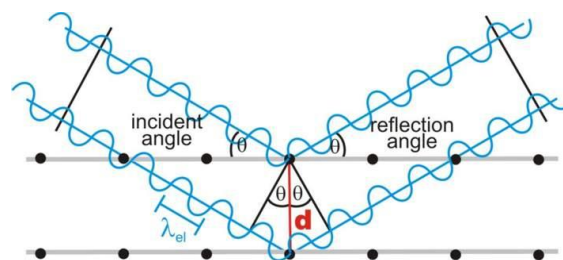


Fig. 3.25. Llei de Bragg ($\text{sen } \theta = n\lambda / 2d$).

Per tal que es produeixin tots els feixos difractats possibles, el cristall ha de ser girat de tal forma que tots els plans virtuals es col·loquin en la posició adequada per a complir la llei de Bragg. Aquest requeriment es pot comprendre millor si fem servir el concepte de *xarxa recíproca* i *el model d'Ewald*, que suposa una simplificació del model de Bragg. El model d'Ewald substitueix cada família de plans virtuals per un punt geomètric. Així totes les famílies de plans possibles

El cristall es col·loca en un goniòmetre que posseeix diferents graus de llibertat de gir, per tal de poder recollir les reflexions del major nombre de plans possibles. Normalment es fa girar el cristall en intervals petits, amb almenys un grau de llibertat, mentre incideixen en ell els raigs X.

El feix incident produeix calor i radicals lliures en la mostra analitzada. Per tant és important mantenir la temperatura del cristall congelat durant el procés de difracció fent servir un flux laminar de nitrogen fred (~100 K). Podem observar un difractòmetre a la fig. 3.27 on podem veure tots els elements que hem comentat.

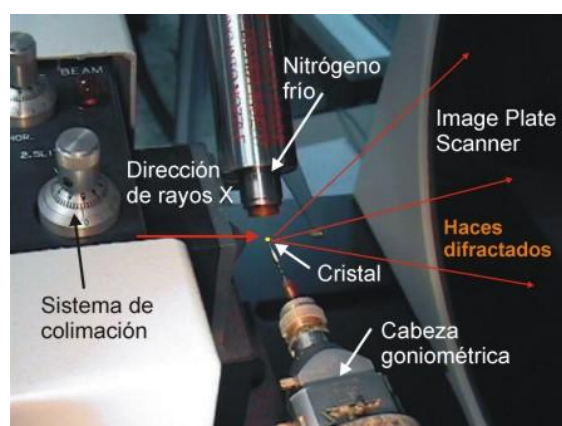


Fig. 3.27. Imatge d'un difractòmetre.

Imatge extreta de [<http://www.xtal.iqfr.csic.es/>].

En aquest treball vam portar els cristalls d'HMGA1a-DNA i pèptid-DNA a difractar al sincrotró ERSF (*European Synchrotron Radiaton Facility*) a la línia espanyola BM-16 a Grenoble, França. Els cristalls dels dominis de l'HMGB1 amb DNA van poder ser difractats al nou sincrotró ALBA a Cerdanyola del Vallès a la *beamline* XALOC (fig. 3.28).



Fig. 3.28. Imatge aèria de l'ERSF (*European Synchrotron Radiaton Facility*) a Grenoble (esquerra) i del sincrotró ALBA (dreta) a Cerdanyola del Vallès.

L'ALBA és el primer sincrotró construït a Espanya i va ser inaugurat el 2012. Imatges extretes de [<http://www.esrf.eu/>] i [<https://www.cells.es/>].

Als sincrotrons, els electrons són accelerats en primer lloc en un accelerador lineal (*linac*) i tot seguit en un accelerador circular (*booster*) fins a adquirir un alt nivell d'energia (de l'ordre de GeV). A continuació aquests electrons s'injecten a un gran anell circular d'emmagatzemament on circulen a una velocitat propera a la de la llum amb una energia constant. Els electrons són accelerats i desviats en aquest anell per diferents components magnètics, com imants dipol, ondulators i imants focalitzadors. Els imants dipol desvien els electrons i els ondulators ondulen la seva trajectòria contribuint ambdós components a l'emissió de raigs X. Són precisament aquests raigs X els que es fan servir en les línies d'investigació situades tangencialment al voltant de l'anell (fig. 3.29).

Els avantatges d'aquesta font de radiació en comparació a d'altres més convencionals (als tubs de raigs X o d'ànode rotatori) és la possibilitat de fer servir una àmplia gamma de longituds d'ona i l'alta intensitat de radiació, que permet disminuir el temps de realització dels experiments.

Tant l'ERSF com l'ALBA són sincrotrons de tercera generació que operen a una energia de 6 i 3 GeV, respectivament, i tenen uns anells d'emmagatzemament de 840 i de 270 m de perímetre en cada cas. A les *beamlines* aporten uns raigs X d'alta intensitat que abasten longituds d'ona d'entre 2,5 i 0,6 Å.

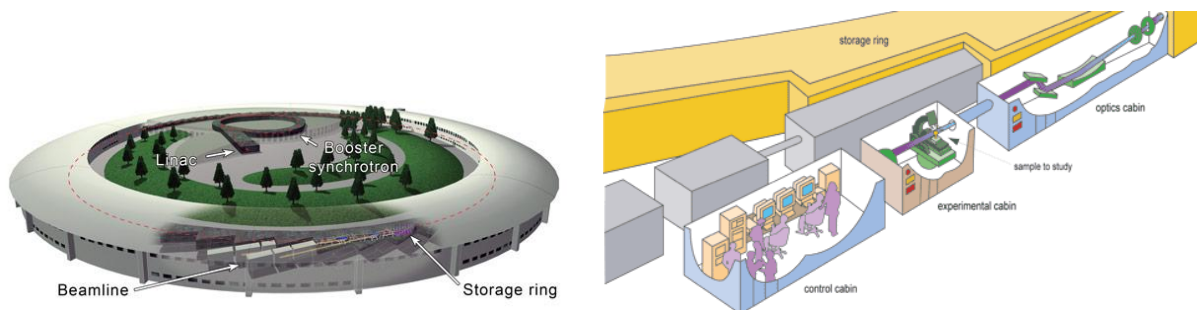


Fig. 3.29. Esquema d'un sincrotró i detall d'una *beamline* del sincrotró ERSF.

Podem observar la cambra òptica, la cambra experimental i la cambra de control. Imatges extretes de [<http://www.esrf.eu/>].

D'altra banda, també hem realitzat proves de difracció a la Plataforma Cristal·logràfica del Parc Científic de Barcelona i a les instal·lacions de raigs X del Colorado University Cancer Center. En aquest cas la font de radiació de raigs X es genera en un ànode rotatori.

Aquests generadors de raigs X consten d'un filament (cànode) i una placa d'un metall pur (normalment de coure o molibdè) (ànode). S'aplica un corrent de baix voltatge (uns 12 V) al filament, obtenint d'aquesta manera una diferència de potencial entre l'ànode i el cànode d'uns 50 kV. Gràcies a aquesta diferència de potencials es produeixen electrons que són projectats del cànode al ànode, provocant, en els àtoms d'aquest últim, una reorganització electrònica en els

seus nivells d'energia. En els generadors d'ànode rotatori, aquest té forma de cilindre i manté un gir continu que li permet dissipar la calor de forma més eficaç i, per tant, obtenir potencials majors de raigs X.

El difractòmetre de la Plataforma Cristal·logràfica del Parc Científic disposa d'un generador d'ànode de coure rotatori amb una potència de 40 kV (MicroMax-007, Rigaku), un detector Image Plate Mar345 SC (MarResearch) i un monocromador de miralls (Osmic Blue). Mentre que el generador de raigs X de la Universitat de Colorado era un Ru-H3R (Rigaku/MSU) i el detector un RAXIS IV ++ (Rigaku).

El generador ens proporciona uns raigs X d'una longitud d'ona constant d'1,5418 Å. Com hem comentat abans, els generadors de tubs o d'ànode rotatori tenen certs inconvenients respecte els sincrotrons però al ser més accessibles per a la realització d'experiments ens permeten fer proves amb alguns cristalls abans de portar-los a difractar al sincrotró.

Procediment experimental

En primer lloc, portem un a un els cristalls que hem conservat en nitrogen (veure apartat 3.8.2) fins al difractòmetre amb l'ajut d'un criotró (espècie de pinces que mantenen el cristall congelat) i l'acoblem al cap magnètic del goniòmetre. Dins del difractòmetre tenim un flux de nitrogen líquid que incideix sobre el cristall i assegura que aquest no es descongeli en cap moment. A continuació centrem el goniòmetre de tal manera que els raigs X incideixin en el cristall en els seus 360° de gir (l'angle de gir és el ϕ).

Amb el suport informàtic es controla la captació de dades i es controlen les variables del procés. Hem fet servir el mètode d'oscil·lació que permet la recollida d'imatges de petites oscil·lacions, obtenint així diversos nivells recíprocs en cada posició per temps petits de recollida. Les variables que podem controlar són:

- L'angle d'oscil·lació o angle que gira el goniòmetre durant la captura de cada imatge. És un angle petit.
- L'angle total girat pel goniòmetre durant la recollida.
- La distància focal o distància al detector.
- El temps d'exposició de cada imatge.
- La longitud d'ona (en el cas del sincrotró).

Finalment obtenim un conjunt de diagrames de difracció per a un cert gir del cristall on cada imatge correspon a una petita oscil·lació, per a uns valors determinats de longitud d'ona, distància focal i temps d'exposició. En la fig. 3.30 podem veure un exemple d'imatge de difracció d'un bon cristall.

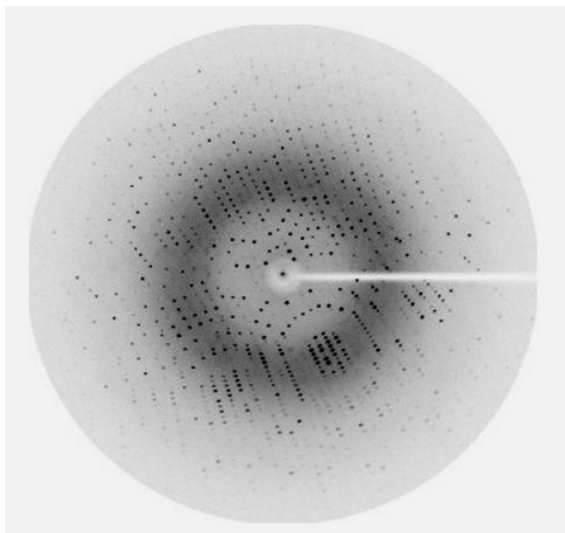


Fig. 3.30. Imatge de difracció d'un cristall.

Imatge extreta de [<http://www.xtal.iqfr.csic.es/>]

3.9.3. El problema de les fases

Per tal de determinar l'estructura d'una macromolècula cal determinar la densitat electrònica $\rho(xyz)$ per a cada punt de la cel·la unitària cristal·lina. Per a fer-ho partim de la informació de l'espai recíproc que hem obtingut de la difracció del cristall (espectre de difracció). La transformació de l'espai recíproc a l'espai cristal·lí implica la resolució d'una transformada de Fourier.

$$\rho(xyz) = \frac{1}{V} \sum_h \sum_k \sum_l |F(hkl)| \cos(2\pi(hx + ky + lz - \phi(hkl)))$$

Eq. 3.4. Funció de la densitat electrònica definida en cada punt de la cel·la elemental.

Aquesta equació representa la transformada de Fourier entre l'espai cristal·lí, definit per $\rho(xyz)$, i l'espai recíproc, definit pels factors d'estructura $F(hkl)$.

En aquesta equació apareixen unes magnituds anomenades factors d'estructura $F(hkl)$, que equivalen a la resultant de la dispersió de tots els àtoms de la cel·la, per a cada punt de l'espai recíproc.

Els factors d'estructura són ones que, per tant, es poden descriure pel seu mòdul $|F(hkl)|$ i la seva fase $\phi(hkl)$. El mòdul el podem trobar directament a partir de la intensitat $I(hkl)$ però la fase no es pot mesurar experimentalment.

En resum, a partir d'un conjunt de diagrames de difracció, i després d'alguns tractaments de dades, podrem trobar les intensitats i amb aquestes podrem calcular els mòduls dels factors d'estructura. Per a trobar les fases caldrà fer servir mètodes indirectes de càlcul per a poder resoldre l'anomenat problema de les fases. Un cop obtingudes les fases, podrem resoldre l'equació Eq. 3.4 que ens proporciona un mapa de densitat electrònica i per tant podrem determinar l'estructura.

L'obtenció de l'estructura a partir dels diagrames de difracció es pot resumir en quatre fases diferenciades (fig. 3.31):

- Anàlisi dels diagrames de difracció
- Resolució de l'estructura
- Afinament
- Validació

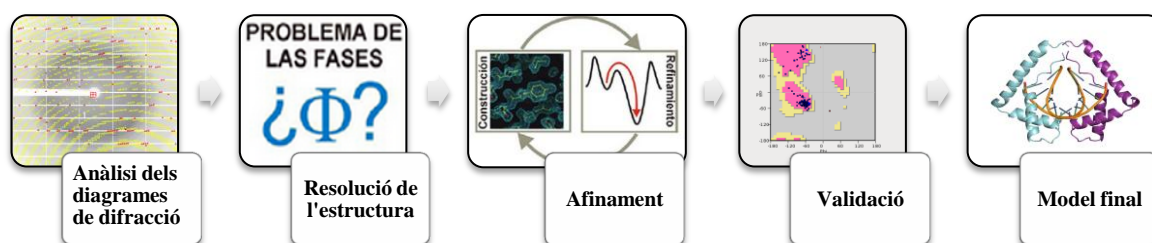


Fig. 3.31. Esquema dels diferents passos per a l'obtenció de l'estructura per difracció de raigs X.

Una revisió que dona una visió ampla, útil i pràctica sobre aquests passos a seguir per a la resolució d'estructures i enfocat a cristal·lògrafs amateurs és la feta per Wlodawer el 2013 [98].

3.9.4. Anàlisi dels diagrames de difracció

De l'anàlisi dels diagrames de difracció (posició i intensitat de les taques) es poden determinar els paràmetres de la cel·la unitària i el seu grup espacial, i obtenir la intensitat per a cada punt de l'espai recíproc. Aquesta anàlisi es realitza en tres etapes: l'indexat, la integració i l'escalat.

En l'indexat fem servir uns pocs diagrames de difracció per tal de determinar la xarxa de Bravais i les dimensions de la cel·la. Hem fet servir els programes *Denzo/HKL2000* [99] o *iMosflm* [100] que assignen a totes les reflexions observades els corresponents índexs de Miller. A partir d'aquests valors el programa pot trobar les direccions dels eixos del cristall i per tant la cel·la unitària. L'usuari a continuació defineix la cel·la i el grup espacial a partir del valor més baix de la matriu de distorsió que ens dóna el programa (compara les nostres dades amb els possibles grups espacials).

El pas següent és la integració, en el que per a calcular la millor posició del màxim d'intensitat s'assigna un perfil Gaussià a les taques, el radi del qual s'avalua amb les dades experimentals. Una de les opcions que es tenen en aquest pas és ajustar el perfil. Primerament es fa una predicció basada en el contorn d'altres reflexions en un radi determinat. A continuació es combina la informació del perfil predit amb l'actual i es minimitza per mínims quadrats la diferència d'aquests dos valors. La integració es realitza per a totes les taques dels diferents diagrames de difracció, obtenint-se d'aquesta manera un llistat de reflexions (h,k,l) amb les seves intensitats.

A continuació es realitza l'escalat, que té com a objectiu posar a la mateixa escala les intensitats de les diferents observacions de cada reflexió i extreure un fitxer que conté les intensitats promig de cada reflexió. L'escalat es realitza amb el programa *Scalepack/HKL2000* [98] o *Scala* [101]. A més ens proporciona valors estadístics per a avaluar els resultats obtinguts.

La qualitat de les nostres dades dependrà dels valors d'aquests paràmetres:

- Resolució màxima. La resolució màxima és el valor de resolució que agafem com a límit de la integració, com més petit millor.

Com a regla general es talla la resolució quan $I/\sigma > 2$ (intensitat de reflexió/desviació de la intensitat) per evitar la introducció d'error.

- $R_{\text{merge}} < 15\%$, sent $R_{\text{merge}} = \frac{\sum_{hkl} \sum_i (\bar{I}(hkl) - I(hkl)_i)}{\sum_{hkl} \sum_i (I(hkl)_i)}$
- χ^2 pròxim a 1, sent $\chi^2 = \sum k \frac{(I - \bar{I})^2}{N\epsilon^2}$ (només proporcionat per *Scalepack/HKL2000*).
- Completitud (% reflexions úniques), com major millor.
- Redundància (cops que es repeteixen les reflexions), com major millor.
- B-factor, (està relacionat amb el dany per radiació), B-factor > -10 .

La qualitat d'aquests paràmetres estadístics depenen directament de la qualitat dels diagrames de difracció però es poden intentar optimitzar en alguns casos. Es poden variar alguns paràmetres durant l'anàlisi de les imatges com ara la mosaïcitat, l'error estimat, el factor d'error d'escala, etc. (especialment en *Denzo-Scalepack/HKL2000*). A vegades cal extreure algunes imatges que empitjoren aquests valors, como per exemple les del final de la recollida quan el cristall comença a fer-se malbé per la incidència prolongada de raigs X.

Finalment es converteixen les intensitats en factors d'estructura amb el programa *Truncate* [102] del paquet de programes CCP4 [103].

$$I(hkl) = K \cdot A \cdot L \cdot p \cdot |F(hkl)|^2$$

Eq. 3.5. Relació entre el mòdul dels factors d'estructura i la intensitat.

K és un factor d'escala, A el factor d'absorció, L el factor de Lorentz i p el factor de polarització.

3.9.5. Resolució de l'estructura

L'objectiu final seria aconseguir un mapa de densitat electrònica. Si partim de la eq. 3.4 veiem que tenim els mòduls dels factors d'estructura però no la fase. En el cas de cristalls de macromolècules el problema de les fases es pot resoldre mitjançant tres mètodes: Reemplaçament Isomorf Múltiple (MIR, de l'anglès *Multiple Isomorphous Replacement*), Dispersió Anòmala Múltiple (MAD, de l'anglès *Multi-wavelength Anomalous Diffraction*) o Reemplaçament Molecular (MR, de l'anglès *Molecular Replacement*).

El mètode MIR consisteix en introduir a la cel·la àtoms que siguin grans dispersors de raigs X, com per exemple àtoms de nombre atòmic alt. Cal però que aquests àtoms no distorsionin l'estructura nativa de la proteïna i per tant siguin isomorfs amb els cristalls de proteïna nativa. La diferència entre l'espectre del cristall amb àtoms pesats i el natiu permet el càlcul del mòdul i la fase dels factors d'estructura.

El mètode MAD implica la difracció d'un cristall de proteïna que contingui un dispersor anòmal fort fent servir radiacions de diferents energies (és a dir, de diferents longituds d'ona). Analitzant les diferències entre les diferents dades de difracció és possible calcular una primera aproximació per a les fases dels factors d'estructura.

En aquest treball hem fet servir el mètode del reemplaçament molecular per tal de resoldre l'estructura. Consisteix en orientar un model estructural conegut amb similitud a la nostra estructura desconeguda. El problema consisteix en trobar com es troba empaquetat el model en la nostra estructura, és a dir, determinar la seva correcta orientació i posició precisa. Per això es fa

ús de les funcions de rotació i de translació per a dur a terme el reemplaçament molecular. És evident que posicionar les molècules en la cel·la és equivalent a posicionar-les a la unitat asimètrica, ja que el contingut en aquesta última ens generarà el contingut a tota la cel·la per les operacions de simetria del grup espacial.

Per exemple per a resoldre l'estructura d'un cristall d'HMGB i oligonucleòtid es pot fer servir com a model un B-DNA teòric de la mateixa seqüència i l'estructura resolta d'aquesta proteïna o d'altra que tingui un plegament similar. Caldrà realitzar diferents proves de reemplaçament molecular per trobar on estan posicionades exactament les molècules de DNA i proteïna a la unitat asimètrica. En els casos de complexos de macromolècules normalment es posiciona primer una (per exemple el DNA) i després l'altra (la proteïna).

El primer pas per a començar aquesta cerca és estimar el nombre de molècules de la unitat asimètrica. Per això es va fa servir el coeficient de *Matthews* (V_m) que equival al volum del cristall per unitat de massa molecular de proteïna $\text{\AA}^3/\text{Da}$. El rang de valors d'aquest coeficient i la seva freqüència ha estat analitzat basant-se en més de 15000 estructures del *Protein Data Bank* (PDB). El *Protein Data Bank* és una base de dades on estan emmagatzemades les estructures de RMN i raigs X de totes les estructures resoltes fins al moment. En l'estadística el coeficient de *Matthews* ha estat tractat separatament en el cas de complexos i àcids nucleics. El valor mitjà és de $2,69 \text{\AA}^3/\text{Da}$ per proteïnes, $2,59 \text{\AA}^3/\text{Da}$ per àcids nucleics i $3,08 \text{\AA}^3/\text{Da}$ per complexos [104]. Estimant V_m i aplicant la següent equació es pot aproximar fàcilment el nombre de molècules a la unitat asimètrica.

$$V_m = \frac{V}{M m n}$$

Eq. 3.6. Equació per al càlcul del coeficient de *Matthews*.

Sent:

V_m : coeficient de *Matthews*

V : volum de la cel·la unitària

M : massa molecular de la proteïna, l'àcid nucleic o el complex

m : multiplicitat del grup espacial (nombre de vegades que es repeteix la unitat asimètrica a la cel·la)

n : nombre de molècules de la proteïna, l'àcid nucleic o el complex

A la pàgina web del Dr. Rupp [<http://www.ruppweb.org/>], el software grafica la probabilitat d'ocurrència en funció del coeficient de *Matthews* (que és proporcional al nombre de molècules) per a uns valors determinants de cel·la, massa molecular i grup espacial. El valor estimat de molècules a la unitat asimètrica és el més probable però pot ser que no es compleixi i hi hagin més molècules o menys de les esperades depenent de si el cristall està més o menys solvatat.

D'altra banda també es poden fer servir el mapa de Patterson i la funció d'autorotació (*self-rotation function*) per a conèixer relacions addicionals de simetria dins del cristall (v. annex F).

A continuació cal preparar els models adients per a començar el reemplaçament molecular. En el cas dels oligonucleòtids hem fet servir principalment el programa *Turbo-Frodo* [105] que permet la creació de models de B-DNA i A-DNA amb seqüència específica i valors de *twist* i *rise* determinats per l'usuari (entre d'altres possibilitats). Per a les proteïnes es poden fer servir els models de proteïnes similars d'estructures ja resoltes per raigs X o RMN que es poden trobar al *Protein Data Bank* (PDB). Com ja veurem, els fragments o les molècules que no es localitzen per reemplaçament molecular poder ser construïdes durant l'afinament, però sempre és millor construir el menys possible. Els programes *Pymol* [106] i *Cerius²* [107] permeten la visualització de qualsevol model, la construcció de models i la realització de modificacions, entre d'altres.

Un cop obtingut el model adient, es realitzaria el mètode del reemplaçament molecular amb el programa *MolRep* [108] (del paquet CCP4), amb *AMoRe* [109] o amb *Phaser* [110] que a més a més ens proporcionen els paràmetres R i C. Aquests paràmetres ens permeten conèixer l'adequació de les solucions de posicionament del model respecte a les nostres dades.

$$R = \frac{\sum_{hkl} \left| |F(obs)| - k|F(calc)| \right|}{\sum_{hkl} |F(obs)|}$$

Eq. 3.7. Càlcul del paràmetre R.

F(obs): factor d'estructura observat en les difraccions.

F(calc): factor d'estructura calculat a partir del model teòric.

$$C = \frac{\sum_{hkl} \left[\left(|F(obs)|^2 - \overline{|F(obs)|^2} \right) \cdot \left(|F(calc)|^2 - \overline{|F(calc)|^2} \right) \right]}{\left[\sum_{hkl} \left(|F(obs)|^2 - \overline{|F(obs)|^2} \right) \cdot \sum_{hkl} \left(|F(calc)|^2 - \overline{|F(calc)|^2} \right) \right]^{1/2}}$$

Eq. 3.8. Càlcul del paràmetre C.

F(obs): factor d'estructura observat en les difraccions.

F(calc): factor d'estructura calculat a partir del model teòric.

R és un factor que representa l'error entre el model i la difracció. Durant cada cicle de refinament entre les dades i el model, és important que aquest factor es redueixi. S'accepta un valor entre un 15% i 25%. El factor C és un coeficient de correlació entre l'adequació del model i les dades observades, per tant ens interessa que sigui com més gran millor.

3.9.6. Afinament

Amb les fases dels factors d'estructura ja podem calcular la funció de densitat electrònica per a cada punt de la cel·la. La representació gràfica d'aquesta funció dóna lloc a l'anomenat mapa de densitat electrònica. A més del mapa de densitat també es calcula el mapa de diferències (veure fig. 3.32):

- Mapa de densitat $2F_oF_c \rightarrow |F(hkl)| = 2|F(obs)| - |F(calc)|$

Idealment els mapes $F(obs)$ i $F(calc)$ haurien de superposar-se.

- Mapa de diferències $F_oF_c \rightarrow |F(hkl)| = |F(obs)| - |F(calc)|$

Si el model fos correcte el mapa de diferències seria nul. Per tant, ens serveix per a observar el que ens falta i el que ens sobra.

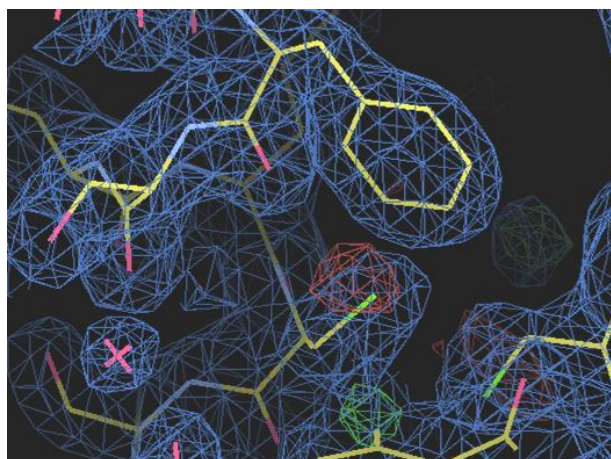


Fig. 3.32. Exemple d'un mapa de densitat electrònica (blau) amb el mapa de diferències (verd i vermell) sobre el traçat d'un model estructural proteïna-DNA.

Si el posicionament del model és correcte el mapa mostrarà zones amb densitat (blau) sobre les que es localitzaran les cadenes polipeptídiques o les cadenes de nucleòtids del model. Amb la informació que ens subministra el mapa de densitat podem corregir o ampliar el model estructural. Per exemple, es poden afegir residus que mostren densitat però no es trobaven al model inicial o moure cadenes laterals perquè segueixin el mapa de densitat. En aquest moment es fan servir programes informàtics per a agilitzar la convergència entre el model estructural i les observacions experimentals (mapa). El model estructural està definit per les seves coordenades de posicionament atòmiques (x, y, z) i un factor de temperatura (B) (mobilitat d'un àtom respecte a la seva posició) per a cada àtom. En aquesta etapa anomenada d'afinament es refinen aquests paràmetres (x, y, z, B) per tal d'optimitzar la concordança entre les observacions i les prediccions. En resum, en l'afinament s'ajusten millor les posicions dels àtoms i es poden afegir parts que faltessin (contraions, molècules d'aigua, residus, cadenes laterals, precipitants, etc).

L'afinament es pot portar a terme fent servir diferents mètodes com el mètode dels mínims quadrats, l'afinament de cos rígid (*rigid body refinement*), l'afinament per probabilitat màxima (*maximum likelihood refinement*) i l'afinament per recuit simulat (*simulated annealing refinement*). Al final de l'afinament s'apliquen uns pocs cicles d'afinament TLS (*Translational oscillations, Librational oscillations, Screw*) que permeten desplaçaments anisotròpics dels àtoms del model. Per a assegurar-nos que l'estructura millora realment amb cada cicle d'afinament i que el descens de R no es deu a motius matemàtics, es descarten un percentatge baix de reflexions que no s'afinen. En cada cicle es calcula el R_{free} corresponent a aquestes reflexions no afinades i el R normal; ambdós han de disminuir de manera similar. En aquest treball hem realitzat l'afinament amb el programa *Refmac5* [111].

Per a visualitzar els mapes de densitat i el de diferències superposats amb el model estructural hem fet servir el programa *Coot* [112]. Aquest programa ens permet la modificació manual del model estructural, mitjançant l'addició d'aigües, l'addició de contraions, la relocalització de les cadenes laterals de les proteïnes, l'addició de residus o cadenes laterals entre d'altres.

3.9.7. Validació

Per tal de validar l'estructura caldria comprovar que el mapa de densitat s'ajusta suficientment al model i que no hi ha zones de valor diferent a zero en el mapa de diferències. A més, cal verificar paràmetres estereoquímics establerts com xocs entre cadenes o distàncies i angles dels enllaços i paràmetres tèrmics raonables. Tota aquesta informació del model estructural queda reflectida en el fitxer pdb on consta la informació cristal·logràfica bàsica (cel·la i grup espacial) i les coordenades atòmiques per a cada àtom juntament amb el seu factor d'ocupació i el seu factor tèrmic (B).

En el cas de les proteïnes és important contrastar les conformacions que adopten els angles de torsió ϕ (phi) i ψ (psi) (v. annex G per a més informació). Aquests angles estan energèticament limitats a certs intervals definits al diagrama de *Ramachandran*, que mostra les regions permeses per a cada tipus d'estructura secundària (hèlix α , làmines β) (fig. 3.33).

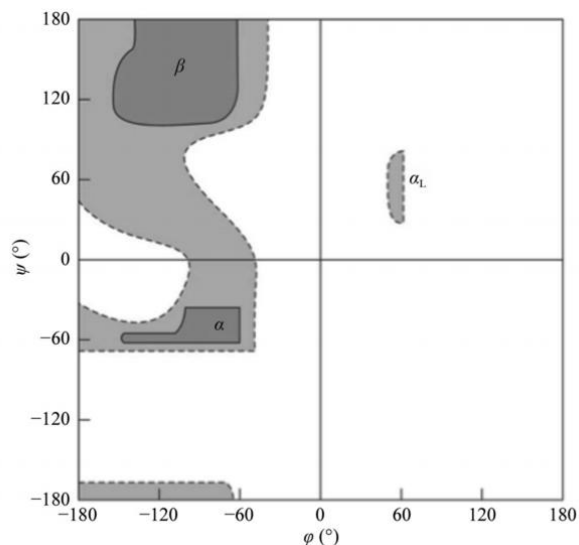


Fig. 3.33. Diagrama de Ramachandran per a les alanines i residus tipus alanina [113].

Es mostren les regions permeses (línies contínues) i les regions parcialment permeses (línies puntejades) per tots els residus menys les glicines i les prolines.

El programa *Coot* proporciona algunes opcions que ens permeten analitzar els xocs entre àtoms i les distàncies i angles d'enllaç. Però principalment hem fet servir el programa *Molprobit* [114] que permet analitzar tots aquests paràmetres i obtenir un fitxer de sortida amb tota aquesta informació destacant els *outliers* (paràmetres que surten de la norma), com es mostra a fig. 3.34.

El programa *Coot* permet realitzar moviments dels àtoms, cadenes laterals, etc. per tal d'optimitzar aquestes propietats estereoquímiques, posicionar les cadenes laterals en les posicions dels rotàmers més afavorits, moure els angles de torsió ϕ i ψ per caure dins de les regions permeses, etc. En alguns casos és necessari l'aplicació de restriccions de distàncies i angles d'enllaç durant l'afinament per a evitar que tinguin valors allunyats dels permesos. Els valors dels angles i de les distàncies d'enllaç per als àcids nucleics i per a les proteïnes han estat obtinguts de [115] i de [116] respectivament.

PROBITY

Viewing 1n78FH-multi.table

When finished, you should . Hint: Use File | Save As... to save a copy of this page.

All-Atom Contacts	Clashscore, all atoms:	25.46	24 th percentile* (N=576, 2.10Å ± 0.25Å)
	Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.		
Protein Geometry	Poor rotamers	6.36%	Goal: <1%
	Ramachandran outliers	0.21%	Goal: <0.2%
	Ramachandran favored	98.61%	Goal: >98%
	Cβ deviations >0.25Å	0	Goal: 0
	MolProbity score [^]	2.51	42 nd percentile* (N=11758, 2.10Å ± 0.25Å)
	Residues with bad bonds:	0.00%	Goal: 0%
Nucleic Acid Geometry	Residues with bad angles:	0.64%	Goal: <0.1%
	Probably wrong sugar puckers:	15	Goal: 0
	Bad backbone conformations [†] :	25	Goal: 0
	Residues with bad bonds:	0.00%	Goal: 0%
	Residues with bad angles:	8.67%	Goal: <0.1%

* 100th percentile is the best among structures of comparable resolution; 0th percentile is the worst.
[†] RNA backbone was recently shown to be rotameric. Outliers are RNA suites that don't fall into recognized rotamers.
[^] MolProbity score is defined as the following: 0.42574*log(1+clashscore) + 0.32296*log(1+max(0,pctRot(Cut-1))) + 0.24979*log(1+max(0,100-pctRama Favored-2)) + 0.5

#	Res	High B	Clash > 0.4Å	Ramachandran	Rotamer	Cβ deviation	Base-P perp. dist.	RNA suite conf.	Bond lengths.	Bond angles.
		Avg: 35.96	Clashscore: 25.46	Outliers: 2 of 932	Poor rotamers: 50 of 786	Outliers: 0 of 868	Outliers: 15 of 150	Outliers: 25 of 150	Outliers: 0 of 1086	Outliers: 19 of 1086
D 574	C	74.81	0.599Å HS with B 177 VAL CG2	-	-	-	suspect sugar pucker - δ & ε outlier (base-p distance indicates 2'-endo)	OUTLIER δ by none (triaged epsilon-1)	-	2 OUTLIER(S) worst is C2'-C3'-C3': 4.876 σ
D 575	C	76.57	0.576Å	-	-	-	-	OUTLIER Shrimp (traced)	-	-

Fig. 3.34. Sortida del programa Molprobity.

En la taula es mostra, entre d'altres, l'adequació dels angles i distàncies d'enllaç, la distribució dels angles de torsió dins del diagrama de *Ramachandran*, la concordança de les conformacions de les cadenes laterals amb els rotàmers preferents o l'absència de col·lisions entre àtoms no enllaçats.

El model final (un fitxer pdb) es diposita al *Protein Data Bank* (PDB) on es revisa i es valida. Un cop superat aquest pas, la macromolècula o complex rep un codi que l'identifica, quedant d'aquesta manera les seves coordenades accessibles a la comunitat científica.

3.10. Anàlisi de DNA mitjançant electroforesi

L'electroforesi d'àcids nucleics és un mètode habitual per a separar, identificar i purificar molècules o fragments de DNA (o RNA). Els àcids nucleics estan carregats negativament i per tant migraran cap al pol positiu (ànode) durant el desenvolupament de l'electroforesi. Ja que cada nucleòtid aporta una càrrega negativa procedent del grup fosfat, els àcids nucleics es desplaçaran a través del gel en funció de la seva mida, podent influir la seva conformació.

Cal tenir present que els àcids nucleics són un conjunt heterogeni de molècules en relació a la seva forma i mida. Així, a l'hora de realitzar una anàlisi electroforètica, cal diferenciar els DNA de doble cadena dels de cadena simple, i els lineals dels circulars (presentant aquests últims diferents graus de superenrotllament). Per tot això, l'electroforesi d'àcid nucleics es porta a terme en gels d'agarosa o de poliacrilamida a diferents concentracions depenent essencialment de la mida dels fragments a separar (taula 3.19).

Taula 3.19. Concentracions d'agarosa o poliacrilamida utilitzades per a la separació electroforètica de fragments lineals de DNA de diferents mides [117].

Concentracions d'agarosa i de poliacrilamida per a diferents mides de DNA	
Agarosa (%)	Interval de mides separables (kb)
0,3	5-60
0,6	1-20
0,7	0,8-10,0
0,9	0,5-7,0
1,2	0,4-6,0
1,5	0,2-3,0
2,0	0,1-2,0
Poliacrilamida (%)	Interval de mides separables (pb)
3,5	1000-2000
5,0	80-500
8,0	60-400
12,0	40-200
15,0	25-150
20,0	6-100

Els gels de poliacrilamida, realitzats en cubetes verticals, es fan servir per a fragments petits de DNA i oligonucleòtids (5-500 pb), així com per a la gran majoria dels RNA. En aquest treball hem fet servir aquest tipus de gel pels assajos del canvi de la mobilitat electroforètica per a visualitzar la formació de complexos entre oligonucleòtids i proteïnes (v. 3.10.1).

Els gels d'agarosa, realitzats en cubetes horitzontals, s'utilitzen per a fragments grans de DNA (des de 500 pb fins a 100 Mb). En aquest treball hem preparat gels d'agarosa per a visualitzar plasmidis d'entre 4 i 6 kb, tant per a electroforesis convencionals (3.10.2) com per a electroforesis bidimensionals (3.10.4).

3.10.1. Gels EMSA (assaig del canvi en la mobilitat electroforètica)

Per a visualitzar la formació del complex entre les proteïnes HMGA1 o HMGB1 i els oligonucleòtids utilitzats als assajos de cristal·lització es duen a terme els anomenats gels EMSA (de l'anglès *Electrophoretic Mobility Shift Assay*) o gels de retard.

L'assaig del canvi en la mobilitat electroforètica és un mètode sensible i ràpid per a detectar interaccions entre proteïnes i àcids nucleics. Aquesta tècnica es basa en l'observació de la disminució en la mobilitat electroforètica del complex proteïna-àcid nucleic en comparació a la mobilitat de l'àcid nucleic lliure (veure fig. 3.35 per a més detalls) [118]. Aquest assaig va ser originàriament descrit per Fried i Crothers (1981) [119] i per Garner i Revzin (1981) [120] i actualment la majoria dels assajos es basen en aquests. Molt sovint el desenvolupament dels gels EMSA es realitza amb un propòsit qualitatiu (verificar la formació de complex entre una proteïna i un oligonucleòtid concrets), encara que també sota unes condicions apropiades ens pot oferir resultats quantitius (estequiometria d'unió, afinitat i cinètica).

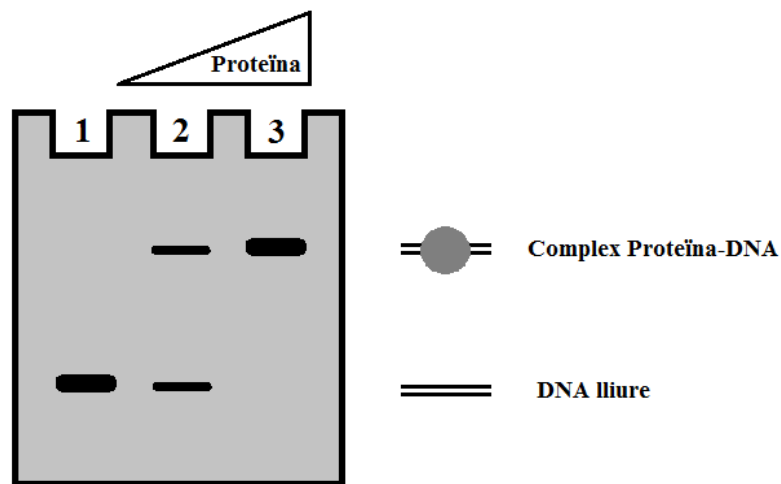


Fig. 3.35. Esquema d'un assaig de canvi en la mobilitat electroforètica.

Al gel EMSA es carreguen concentracions creixents de proteïna. A l'aplicar el camp elèctric es desplaçarà més fàcilment el DNA lliure mentre que el complex proteïna-DNA quedarà endarrerit degut a la seva major mida i menor càrrega. En aquest esquema, en el pou 1 s'ha carregat només DNA, en el 2 s'inicia la formació del complex i en el 3 tot el DNA està complexat.

Per a la realització dels gels EMSA s'han descrit moltes variants, en el nostre cas hem fet servir un protocol per a les proteïnes HMGA i un altre de millorat per a les HMGB. A la taula 3.20 i taula 3.21 es detallen les composicions dels gels, dels tampons de cubeta i dels tampons de mostra per a les HMGA i les HMGB, respectivament. De forma general, aquestes electroforesis es realitzen en gels d'entre 5 i 15% de poliacrilamida en condicions no desnaturalitzants i es desenvolupen en fred (4°C). Els tampons de mostra porten únicament solució tampó i glicerol (sense colorant de migració) per a no interferir en la formació dels complexos.

Taula 3.20. Composició dels gels EMSA per a les proteïnes HMGA.

Gel EMSA 7,5% Acrilamida – HMGA1	
Reactius	Per a 12 ml
30% Acrilamida/ 0,8% Bisacrilamida – Sol. A	3 ml
Glicerol	600 µl
1 M HEPES pH=8,3	600 µl
H₂O	7,8 ml
TEMED	12 µl
PSA 10%	60 µl
Tampó de cubeta	
Composició	Per a 1 litre
50 mM HEPES pH=8,3 0,1 M NaCl 3% glicerol	50 ml 1 M HEPES pH=8,3 25 ml 4 M NaCl 30 ml glicerol
Tampó de mostres	
Composició	Per a 10 ml
50 mM HEPES pH=8,3 0,1 M NaCl 30% glicerol	1 ml 1 M HEPES pH=8,3 0,5 ml 4 M NaCl 6 ml glicerol

Taula 3.21. Composició dels gels EMSA per a les proteïnes HMGB.

Gel EMSA 8 % Acrilamida – HMGB1	
Reactius	Per a 10 ml
30% Acrilamida/ 0,8% Bisacrilamida –Sol. A (40% Acril./ Bisacril. 37,5:1 – Biorad)	2,67 ml (2 ml)
5x TBE	660 µl
H₂O	6,61 ml (7,28 ml)
TEMED	10 µl
PSA 10%	50 µl
Tampó de cubeta	
Composició	Per a 1 litre
0,33x TBE	66 ml 5x TBE

Les mostres es preparen amb un tampó adient per a l'electroforesi (0,3x TBE contenint 3% glicerol) (v. taula 3.22) i per tant no els hi cal afegir cap tampó de mostra addicional.

Abans de realitzar l'electroforesi pròpiament s'ha de fer un *pre-run* (o pre-electroforesi) que consisteix en aplicar una diferència de voltatge al gel (més baixa que la de la pròpia electroforesi) durant mínim 30' abans de carregar les mostres. Això ens permet eliminar les possibles traces de persulfat d'amoni, equilibrar el gel amb el tampó de cubeta i assegurar una temperatura constant al llarg del gel.

A més a més una altra diferència important d'aquest tipus de gel és la de necessitar un temps d'incubació de les mostres preparades per a carregar. Aquest temps d'incubació permet la formació del complex (en el cas que es formi) i normalment és de mínim 30' a 4°C. Un exemple de les mostres preparades per a una incubació i que després es carregaran a un gel EMSA es troba a la taula 3.22 on podem observar la concentració creixent de la proteïna per a una quantitat fixada d'oligonucleòtid.

Taula 3.22. Reaccions d'incubació per a un gel EMSA en el cas de les proteïnes HMGB.

Ràtio molar (Proteïna: DNA)	0:1	0,5:1	1:1	1,5:1	2:1	2,5:1	3:1	3,5:1	4:1
5 µM DNA en 0,3x TBE (µl)	2	2	2	2	2	2	2	2	2
10 µM proteïna en 0,3x TBE (µl)	0	0,5	1	1,5	2	2,5	3	3,5	4
0,3x TBE (µl)	7,4	6,9	6,4	5,9	5,4	4,9	4,4	3,9	3,4
50 % Glicerol (µl)	0,6	0,6	0,6	0,6	0,6	0,6	0,6	0,6	0,6
Volum final	10 µl								

Un cop realitzada l'electroforesi, els gels es tenyeixen amb tenyidors específics d'àcids nucleics i es visualitzen amb llum ultraviolada a un transil·luminador-UV. Hem fet servir els tenyidors SYBR Gold i SYBR Green I (Invitrogen, Life Technologies) que són tenyidors més sensibles i més segurs que el bromur d'etidi (el tenyidor universal d'àcids nucleics però que és cancerigen).

3.10.2. Electroforesi en gels d'agarosa

L'electroforesi en gels d'agarosa és el mètode més comú per a l'anàlisi de DNA, però també es realitza aquesta tècnica per a l'obtenció i purificació d'aquest. Existeixen diferents tipus d'agaroses comercials depenent de la seva puresa, la seva temperatura de gelificació i fusió (més baixes si han estat modificades) i els seus nivells d'electroendosmosi (EEO). El flux electroendosmòtic va en contra de la direcció de migració del DNA, fet que perjudica la resolució, per tant és preferible triar una agarosa amb baixa EEO.

L'agarosa en pols es dissol en el tampó de cubeta en la quantitat necessària per a assolir el grau d'entrecruament adient i s'escalfa a 100°C. Un cop dissolta s'aboca sobre una petita safata de plàstic transparent (moltlle) tancada pels extrems amb cinta adhesiva. Es col·loca llavors la pinta de forma perpendicular a la safata, fent que els pouets quedin entre 1 i 2 mm del fons. L'agarosa al refredar-se gelifica. Un cop polimeritzat el gel, es retira la pinta i la cinta adhesiva per a permetre el pas de la corrent. El gel juntament amb la safata es col·loquen a la part central d'una cubeta horitzontal i aquest es cobreix totalment amb el tampó de cubeta (fig. 3.36). Amb aquesta disposició es dissipa millor la calor generada pel pas de la corrent. Finalment es carreguen les mostres al gel i es desenvolupa l'electroforesi.

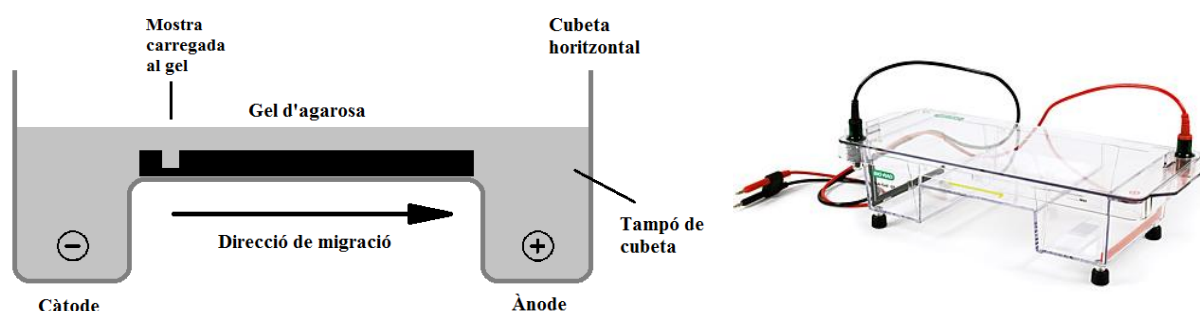


Fig. 3.36. Esquema de l'electroforesi horitzontal i imatge d'una cubeta per a aquest tipus d'electroforesi.

Existeixen diferents tamps de cubeta per a la realització d'aquest tipus de gels. Aquests contenen Tris-acetat i EDTA (TAE), Tris-borat (TBE) o Tris-fosfat (TPE) a concentracions de ~50 mM (taula 3.23). Tots ells funcionen bé i a vegades és una qüestió personal la tria d'un o d'un altre. De forma general es pot dir que el TAE és el que presenta menor capacitat de tamponar dels tres i és ineficient en electroforesis llargues. El poder de resolució per a DNAs d'alta massa molecular és lleugerament millor per al TAE que per al TBE o TPE, i pitjor per a DNAs de baixa massa molecular.

Abans de carregar el gel, la mostra de DNA (idealment en el tampó de cubeta corresponent) es barreja amb el tampó de mostres per a augmentar la seva densitat i afegir-li color. En el nostre cas, els colorants són el blau de bromofenol, que migra a través dels gels d'agarosa a una velocitat equivalent d'un DNA lineal de doble cadena de 300 pb, i el xilencianol que migra com un DNA lineal de doble cadena de 4 kb (en gels 0,5-1,4% d'agarosa i 0,5x TBE). Tots dos migren cap el pol positiu. A la taula 3.24 figura la composició del tampó de mostres fet servir.

En aquest treball s'han realitzat gels d'agarosa per a la visualització de les distribucions de topoisòmers de plasmidis (3.10.3) i dels plasmidis obtinguts en els assajos d'introducció de *supercoiling* (superenrotllament) (3.10.5). En aquests casos hem preparat gels petits de dimensions 6 x 8 cm i ~5 mm de gruix a l'1% d'agarosa en 0,5x TBE en el primer cas i en 1x TPE en el segon. Per a preparar-los s'ha fet servir 30 ml de tampó de cubeta i 0,3 g d'agarosa.

Les electroforesis en gels d'agarosa es solen desenvolupar entre 1-10 V/cm (els cm corresponen a la distància entre els elèctrodes) depenent de la resolució requerida i la mida dels fragments a separar.

Per a la detecció de les bandes de DNA es fa servir bromur d'etidi (BrEt). Aquest compost s'intercala entre les bases del DNA i s'observa en un aparell transil·luminador amb llum ultraviolada. Un cop finalitzada l'electroforesi, es submergeix el gel en 0,5-1,5 µg/ml BrEt durant mínim 30 minuts amb agitació. En alguns casos també s'ha fet servir el tenyidor SYBR Gold (Invitrogen, Life Technologies), que és més sensible i segur.

Taula 3.23. Tampons de cubeta per a l'electroforesi en gels d'agarosa [121].

Tampons de cubeta per a electroforesi		
Tampó	Solució de treball	Solució mare (1 litre)
TAE	1x 40 mM Tris-acetat 1 mM EDTA	50x 242 g Tris base 57,1 ml Àcid acètic glacial 100 ml 0,5 M EDTA pH=8
TPE	1x 90 mM Tris-fosfat 2 mM EDTA	10x 108 g Tris base 15,5 ml Àcid fosfòric (85%, 1,679 g/ml) 40 ml 0,5 M EDTA pH=8
TBE	0,5x 45 mM Tris-borat 1 mM EDTA	5x 54 g Tris base 27,5 g Àcid bòric 20 ml 0,5 M EDTA pH=8
TAE Mg*	1x 40 mM Tris-acetat 2 mM EDTA 12,5 mM MgAc	10x 48,4 g Tris base 11,4 ml Àcid acètic glacial 26,8 g MgAc 40 ml 0,5 M EDTA pH=8

*Específic per a l'assaig de la visualització de transicions estructurals de DNAs circulars en electroforesis bidimensionals (v. 3.10.4).

Taula 3.24. Tampó de mostres per a gels d'agarosa [121].

Tampó de mostres per a gels d'agarosa	
Reactius	6x Concentració
Blau de bromofenol	0,25%
Xilencianol	0,25%
Glicerol	30%

3.10.3. Obtenció d'una distribució de topoisòmers

Per tal de visualitzar la possible transició estructural que podria experimentar un determinat plasmidi (DNA circular de doble cadena) en un gel d'agarosa bidimensional (v. 3.10.4) és necessari prèviament l'obtenció d'una distribució de topoisòmers. Una distribució de topoisòmers és un conjunt de DNAs circulars amb diferent grau d'enrotllament (v. apartat 1.2.5). Els plasmidis es troben normalment de forma superenrotllada. Les topoisomerases són enzims capaços de modificar el grau de superenrotllament d'aquests DNA circulars mitjançant el trencament d'un o dos enllaços fosfodièsters, el desplaçament de la cadena de DNA i la formació de l'enllaç prèviament trencat. En el cas de la topoisomerasa I, aquesta talla una única cadena.

Topoisomerasa I



Fig. 3.37. Acció de la topoisomerasa I.

La incubació d'una mostra de DNA plasmídic amb topoisomerasa I, en les condicions adients, ens permetrà l'obtenció d'una població de molècules amb diferents graus de superenrotllament. Al ser analitzades electroforèticament, les mostres migraran al gel en funció del seu nombre d'enllaç (Lk) o grau de superenrotllament, desplaçant-se més ràpidament aquelles molècules que presentin un major enrotllament, i per tant estiguin més compactes (fig. 3.38).

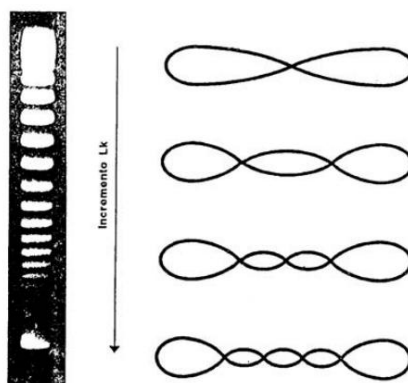


Fig. 3.38. Relació entre la mobilitat electroforètica i el grau de superenrotllament de DNAs circulars [117][115].

Si afegim topoisomerasa I a una mostra de molècules de DNA circular superenrotllat (negativament) aquestes passaran a estar completament relaxades fàcilment. Així, per tal d'obtenir una distribució de molècules amb diferents graus de superenrotllament, cal afegir en primer lloc un compost intercalant a la mostra. Aquest intercalant, que pot ser bromur d'etidi o cloroquina, desenrotllarà la doble hèlix reduint el gir (*twist*, Tw). Com a conseqüència, la molècula adquireix torsió o retorçament (*writhe*, Wr) positiu, formant voltes de superhèlixs. És a

dir, s'eliminaran voltes de superenrotllament negatiu. Així tindrem $\Delta Tw < 0$, $\Delta Wr > 0$ i $\Delta Lk = 0$. A l'afegir la topoisomerasa I, en presència de l'intercalant, aquesta només podrà desenrotllar les voltes de superhèlix que quedin (trencant enllaços i, així, canviant el valor d' Lk). Finalment al retirar l'intercalant l'efecte en forma de voltes que aquest havia produït s'anul·larà, formant-se el mateix nombre de voltes negatives que havia eliminat. En la fig. 3.39 observem un exemple de l'efecte de la topoisomerasa a diferents concentracions d'intercalant. Per exemple, en el cas de treballar amb baixa concentració d'intercalant (B, fig. 3.39), aquest s'uneix al DNA suficientment per a desenrotllar dues voltes (o eliminar dues voltes de superhèlix negatives). La topoisomerasa I elimina les tres voltes restants per a relaxar completament la molècula (pas *a*). Finalment després de retirar el lligand intercalant les dues voltes que havia desfet és manifesten amb un valor final de $Lk = -2$ (pas *b*).

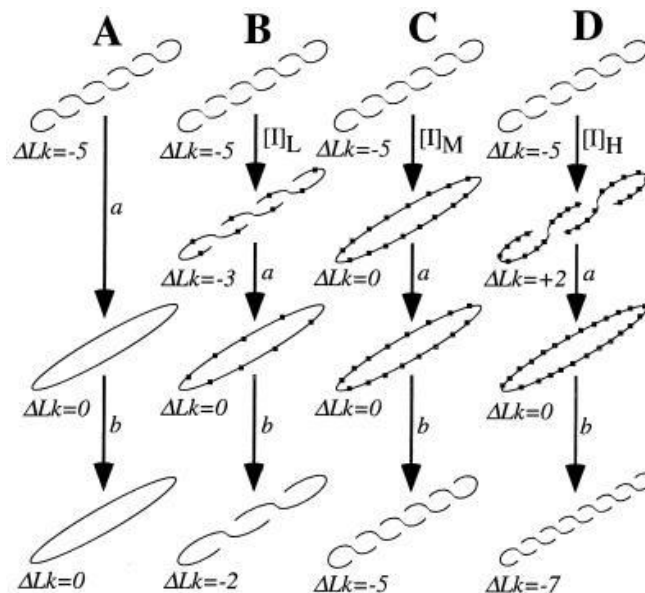


Fig. 3.39. Esquema de les reaccions de relaxació de la topoisomerasa I en presència de diferents concentracions d'intercalador [122].

El pas *a* representa la relaxació per Topoisomerasa I, el pas *b* la retirada de l'intercalant, els punts negres el lligand intercalant per a les quatre reaccions: (A) sense lligand intercalant, (B) amb baixa concentració d'intercalant, (C) amb concentració mitja d'intercalant i (D) amb alta concentració d'intercalant.

Així, canviant la concentració d'intercalant en les diferents reaccions amb topoisomerasa I obtindrem molècules amb diferents graus de superenrotllament. A més a més, la topoisomerasa no actua de forma homogènia així que per a cada reacció amb una concentració d'intercalant determinada s'obtindran molècules amb petites diferències en el grau de superenrotllament.

En aquest treball s'ha fet servir el bromur d'etidi com a intercalant per a obtenir els diferents graus de superenrotllament de la mostra de DNA circular. Hem realitzat les reaccions que es mostren a la taula 3.25 durant una hora. Després s'ha dut a terme una precipitació amb etanol per a purificar les mostres de DNA descartant així la resta de reactius i les sals.

Taula 3.25. Reaccions d'incubació per a l'obtenció d'una distribució de topoisòmers.

	Topo 1	Topo 2	Topo 3	Topo 4	Topo 5
0,266 µg/µl DNA (µl)	3	3	3	3	3
10 µg/ml BrEt (µl)	0	0,5	1	2	3
10x tampó reacció (µl)	4	4	4	4	4
Topoisomerasa I (µl)	0,5	0,5	0,5	0,5	0,5
Aigua (µl)	32,5	32	31,5	30,5	29,5
Volum final	40 µl				
Tampó de reacció 10x: 1 M NaCl, 50 mM MgCl ₂ , 350 mM Tris pH=7,5, 10 mM DTT					

Per a la precipitació amb etanol [121]:

- Afegir 4,4 µl de 3 M d'acetat de sodi.
- Barrejar bé i afegir exactament dos volums d'etanol absolut fred (88 µl) i un altre cop barrejar bé. Deixar 30 minuts en gel.
- Centrifugar 10 min a 2°C en micròfuga a la velocitat màxima de l'aparell (13200 rpm).
- Eliminar el sobrenadant amb molt de compte amb l'ajut d'una pipeta. Cal tenir cura de no pertorbar el sediment de DNA (el qual és transparent i per tant no veurem). Guardem aquest sobrenadant fins confirmar que hem realitzat el procés correctament.
- Emplenar l'ependorf amb el sediment amb 750 µl d'etanol al 70%.
- Centrifugar 2 min a 2°C en micròfuga al màxim de velocitat (13200 rpm).
- Eliminar el sobrenadant amb molt de compte amb l'ajut d'una pipeta. Igual que abans, reservem aquest sobrenadant.
- Deixar assecar la mostra de DNA a l'aire o al buit (Savant™ Speed-vac, Thermo Scientific™).
- Redissoldre el sediment de DNA en 10 µl del tampó utilitzat per a electroforesi.

Les diferents mostres de DNA redissoltes s'han analitzat electroforèticament en gels d'agarosa amb tampó 0,5x TBE (v. taula 3.23) per a visualitzar el grau de superenrotllament. Finalment s'han barrejat volums adients de les diferents mostres (és a dir, de les diferents reaccions) per a aconseguir una bona distribució de topoisòmers.

3.10.4. Gels bidimensionals

L'electroforesi en una dimensió presenta certes limitacions en la relació de la mobilitat electroforètica i el grau de superenrotllament. Així és difícil separar per electroforesi unidimensionals les molècules que presenten un grau de superenrotllament molt alt o molt baix. En aquestes condicions és útil l'aplicació d'una segona separació electroforètica, en direcció perpendicular i en unes condicions una mica diferents. Aquesta segona dimensió normalment es realitza en presència de l'intercalant cloroquina. A més a més, aquesta electroforesi bidimensional ens permet la visualització de transicions estructurals en les molècules analitzades.

Un cop obtinguda la distribució de topoisòmers (v. apartat anterior), aquesta barreja es carregarà en un gel i es desenvoluparà l'electroforesi en la primera dimensió (D1, fig. 3.40) en absència d'intercalant. Un cop finalitzada aquesta, el gel s'incubarà unes hores en el tampó de cubeta de la segona dimensió que conté l'intercalant cloroquina. L'intercalant produirà un subenrotllament local $\Delta T_w < 0$, que es compensa amb un superenrotllament positiu $\Delta W_r > 0$. És a dir, cada molècula s'enrotlla més (perd voltes negatives). Amb excés d'intercalant totes les molècules presentaran superenrotllament positiu. El gel D1 es gira 90° i llavors es desenvolupa la segona dimensió en presència d'intercalant (D2, fig. 3.40).

Els topoisòmers amb menor ΔL_k (menor mobilitat a D1) mostren major superenrotllament en la segona dimensió (major mobilitat). En canvi, els de major ΔL_k (major mobilitat a D1) tenen el mínim superenrotllament positiu a la segona dimensió (menor mobilitat). El punt r1 és la situació de les molècules relaxades (mínima mobilitat a D1). A la dreta de r1 queden les molècules que tenien un superenrotllament lleugerament positiu a la primera dimensió. Si fem servir menor concentració d'intercalant (concentració intermèdia), no hi ha suficient intercalant per tal que totes les molècules arribin a tenir superenrotllament positiu. El punt r2 indica la molècula que queda relaxada en la segona dimensió, les molècules amb major enrotllament negatiu que aquesta queden encara negatives i tenen major mobilitat.

Perquè es produeixi una transició estructural és necessari que s'arribi a un llindar de superenrotllament per al qual la tensió impulsi el canvi estructural. L'estructura alterada només apareix per a ΔL_k superiors al llindar (els valors més negatius), les molècules que tenen major mobilitat a D1. Un cop adoptada la nova estructura, com aquesta posseeix un menor grau de enrotllament, migrarà menys al gel. En la segona dimensió l'intercalant disminuirà el superenrotllament per sota del llindar, provocant una mobilitat normal. D'aquesta manera, el punt de transició estructural s'observa al gel com una discontinuïtat o salt en el perfil de taques (*j* en fig. 3.40).

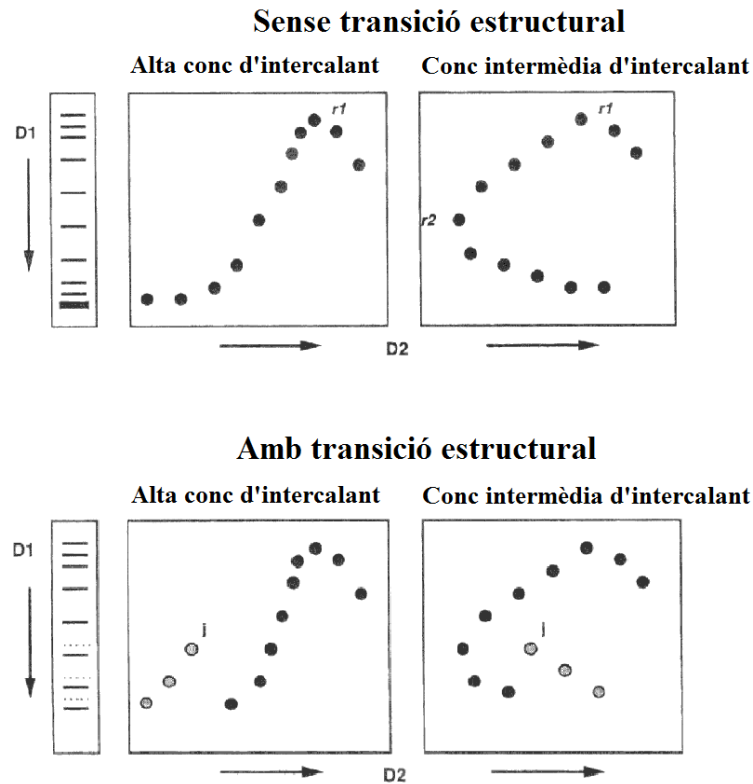


Fig. 3.40. Electroforesis bidimensionals amb i sense transició estructural amb alta i intermèdia concentració d'intercalant [123].

Un exemple de visualització d'una transició estructural d'un DNA circular és el que es mostra a continuació. En aquest cas s'analitza un plasmidi control al costat d'un plasmidi amb un insert d(pCpG)₁₆. Podem observar clarament el salt en el perfil de taques al produir-se la transició entre la forma B i la forma Z del DNA. Amb aquesta transició l'excés de superenrotllament negatiu es compensa en part per la formació de la forma Z que suposa una disminució en la torsió.

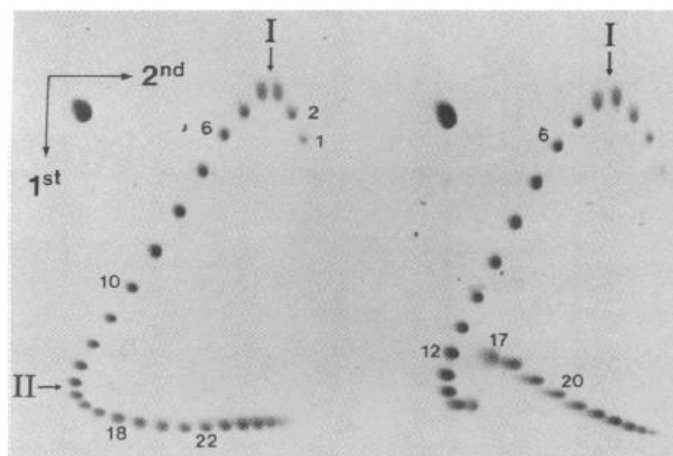


Fig. 3.41. Electroforesis bidimensionals de la transició de la forma B a la Z del DNA [124].

Per a veure la transició estructural en els nostres plasmidis amb inserts ATs hem realitzat electroforesis bidimensionals generalment en les condicions que es detallen a la taula 3.26. Hem fet servir unes cubetes horitzontals grans que permeten l'electroforesi de gels grans de 20 x 25 cm i ~0,5 cm de gruix. Per a preparar-los s'ha barrejat 300 ml d'1x TAEMg amb 6 g d'agarosa i s'ha utilitzat una pinta d'un únic pou quadrat. El contingut en sals d'aquests gels i els llargs temps d'electroforesis produeixen un alt escalfament del gel i una disminució del poder tamponant d'aquest. Per a evitar-ho les electroforesis s'han desenvolupat a 4°C, amb recirculació dels tampons i recanvi del tampó de cubeta de la primera dimensió (v. taula 3.26).

Taula 3.26. Condicions de les electroforesis bidimensionals.

Electroforesis bidimensionals (2% Agarosa)		
Etapa	Tampó	Condicions
1a Dimensió (1D)	1x TAEMg	42,5 hores a ~1,6V/cm (~75V) a 4°C Amb recirculació del tampó i canvi d'aquest cada 12 hores.
Incubació	1x TBE + 1,5 µg/ml cloroquina	6 hores
2a Dimensió (2D)	1x TBE + 1,5 µg/ml cloroquina	40 hores a ~1,5V/cm (~70V) a 4°C Amb recirculació del tampó.

La composició dels tampons TAEMg i TBE és l'especificada a la taula 3.23.

3.10.5. Assajos d'introducció de superenrotllament

Els octàmers d'histones així com la família de les HMGB són proteïnes capaces d'unir-se al DNA i doblegar-lo. Aquest canvi topològic del DNA induït per aquestes proteïnes pot ser visualitzat en un gel d'agarosa com un canvi en el grau de superenrotllament d'un DNA circular.

Per a veure aquesta introducció de superenrotllament d'aquestes proteïnes podem seguir el següent procediment. En primer lloc, la proteïna s'incuba amb DNA circular prèviament relaxat. La proteïna al forçar el doblegament del DNA canvia la seva torsió o retorçiment (*writhe*, Wr), formant voltes de superhelicitat. S'afegeix llavors l'enzim Topoisomerasa I, que mitjançant el canvi del nombre d'enllaç (*linking number*, Lk), dissiparà l'estrès topològic desfent aquestes voltes. Finalment la proteïna serà eliminada, així com la topoisomerasa I, d'aquesta manera s'eliminarà el retorçiment que provocava la proteïna però romandrà el canvi en el Lk provocat per la topoisomerasa. Finalment les mostres resultants s'analitzaran electroforèticament per a visualitzar el canvi en el grau de superenrotllament al variar la concentració de proteïna amb la que s'ha incubat el DNA circular (v. fig. 3.42).

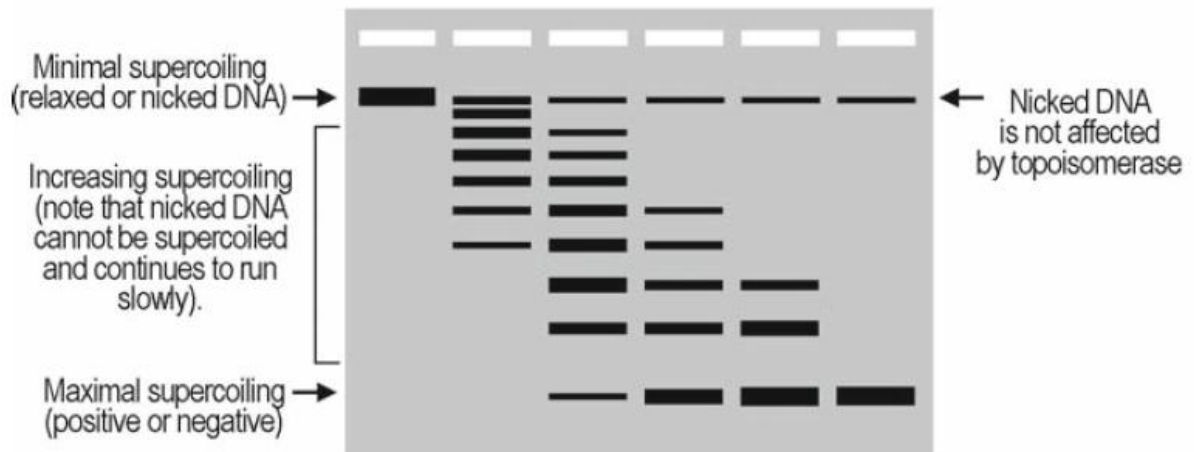


Fig. 3.42. Esquema d'un assaig d'introducció de *supercoiling* [125].

Hem seguit el següent procediment amb el volums indicats a la taula 3.27:

- Barrejar l'aigua, el tampó de reacció i el DNA circular *supercoiled* amb 0,5 μ l Topoisomerasa I.
- Incubació a 37°C durant 1 hora en un bany d'aigua, per a relaxar el plasmidi.
- Afegir el volum adient de proteïna en cada reacció i 0,5 μ l de Topoisomerasa I extres.
- Incubació a 37°C durant 1 hora en un bany d'aigua.
- Afegir i 1 μ l de 10% SDS i 0,5 μ l de 10 mg/ml proteïnasa K.
- Incubació a 37°C durant 30 minuts en un bany d'aigua.
- Afegir 4,3 μ l de tampó de mostres per a gels d'agarosa (4 μ l en el cas de la mostra SC).
- Desenvolupar l'electroforesi en un gel a l'1% d'agarosa petit amb 1x TPE de tampó de cubeta (normalment 2 hores a 90 V).
- Realitzar la tinció amb SYBR Gold i escanejar amb llum ultraviolada.

Taula 3.27. Reaccions d'incubació per a un gel *supercoiling*.

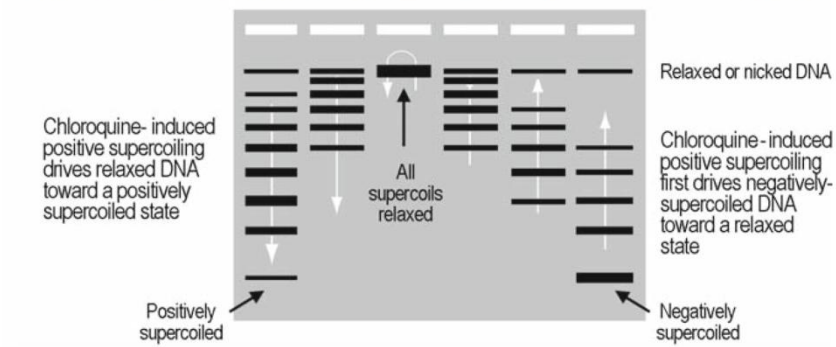
Ràtio molar (Proteïna: DNA)	SC*	0:1	300:1	1000:1	3000:1
0,08 µg/µl DNA (µl)	3,75	3,75	3,75	3,75	3,75
50 µM proteïna (µl)	-	-	0,6	2	6
10x tampó reacció (µl)	2	2	2	2	2
Aigua (µl)	14,25	13,25	12,65	11,25	7,25
Volum final	20 µl	19 µl			
Tampó de reacció 10x (alta força iònica): 1 M NaCl, 50 mM MgCl ₂ , 350 mM Tris pH=7,5, 10 mM DTT o bé					
Tampó de reacció 10x (baixa força iònica): 0,1 M NaCl, 50 mM MgCl ₂ , 350 mM Tris pH=7,5, 10 mM DTT					

*La mostra SC és un control del DNA circular superenrotllat, és a dir, la mostra de plasmidi de partida a la qual no li hem afegit ni topoisomerasa, ni proteïnasa K ni SDS.

Les proteïnes d'unió al DNA poden modificar la seva topologia, modificant els valors del gir (*twist*) i la torsió (*writhe*), per aquest motiu el superenrotllament pot passar en les dues direccions. Per a saber en quina de les dues situacions ens trobem es pot repetir l'electroforesi amb les mateixes mostres però aquest cop en presència d'intercalant. S'afegeix en aquest cas l'intercalant cloroquina al gel i al tampó de cubeta. Si l'efecte de superenrotllament es veu augmentat (les molècules es desplacen més ràpidament) vol dir que els topoisòmers presentaven superenrotllament positiu i la cloroquina ha fet que presentin major *supercoiling* positiu. Si pel contrari els plasmidis presentaven *supercoiling* negatiu, l'efecte de superenrotllament es veurà disminuït (menor mobilitat, degut al menor grau de *supercoiling* negatiu) podent-se arribar a formar topoisòmers positius a baixes concentracions de proteïna.

En el cas dels nucleosomes, per exemple, el DNA s'enrotlla sobre les histones format voltes levogires. Així un DNA relaxat a l'incorporar-se al nucleosoma tindrà superenrotllament negatiu associat a la proteïna i, per a compensar, haurà de desenvolupar superenrotllament positiu (*writhe*, *Wr*) i/o gir (*twist*, *Tw*) en algun lloc de la molècula. Un cop tractades aquestes mostres amb topoisomerasa i amb la subseqüent eliminació de la proteïna, produirà una distribució de topoisòmers amb superenrotllament negatiu [57].

Amb cloroquina al gel; efecte en DNA amb superenrotllament negatiu



Amb cloroquina al gel; efecte en DNA amb superenrotllament positiu

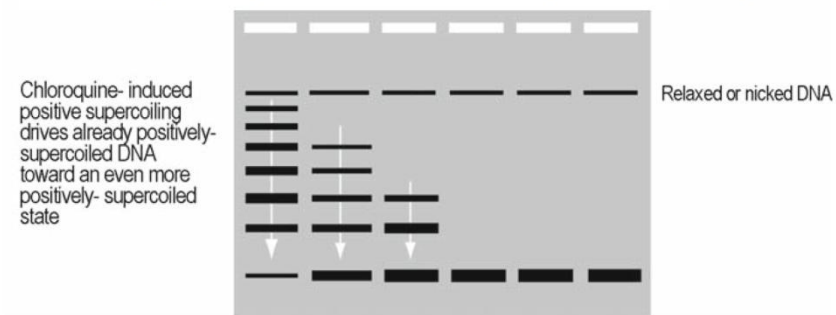


Fig. 3.43. Distributions de topoisòmers analitzada en un gel d'agarosa contenint cloroquina.

En la imatge superior els topoisòmers presentaven superenrotllament negatiu abans de l'electroforesi, en canvi en la inferior els topoisòmers mostraven superenrotllament positiu [125].

3.11. Obtenció i purificació de DNA plasmídic

Al llarg d'aquest treball ha estat necessari en algunes ocasions l'obtenció i purificació de DNA plasmídic. Per a realitzar aquest procés cal en primer lloc la transformació de cèl·lules competents específiques per al clonatge de DNA amb el plasmidi adient, la seva multiplicació en cultiu líquid i finalment la purificació d'aquest DNA plasmídic.

En el cas dels diferents plasmidis pSTAT-CEN 2 i pUC-19 (fets servir per a les anàlisis amb electroforesis bidimensionals) s'han transformat les cèl·lules XL1 Blue supercompetents i s'han purificat realitzant, en el primer cas, Maxi-preps amb el *GenElute HP Plasmid Maxi-prep kit* (Sigma) i en el segon cas, Midi-preps utilitzant el *Plasmid purification kit* (Qiagen). De la mateixa manera el plasmidi pET-3 contenint el gen de l'HMGAlb(1-79) també s'ha purificat realitzant aquestes Midi-preps. En el cas dels plasmidis codificant les proteïnes HMG-box s'han realitzat preparacions de DNA plasmídic per a verificar la seqüència d'aquest, introduint el plasmidi en cèl·lules DH5 α i purificant-los amb *QIAprep Spin Miniprep Kit* (Qiagen).

Obtenim de 5 a 15 μ g de plasmidi purificat a partir d'un cultiu d'1,5 ml amb el *kit* de les Mini-preps, mentre que el *kit* de Midi-preps ens proporciona rendiments d'entre 20-100 μ g per 100 ml de cultiu per a plasmidis de baix nombre de còpia (com és el nostre cas). En canvi, el *kit* de Maxi-preps ens pot arribar a purificar fins a 1,2 mg (per a 150 ml de cultiu).

La purificació de DNA amb aquests tipus de *kits* es basa en un procés de lisi alcalina seguit de la unió del DNA plasmídic a una columna en unes condicions adients de força iònica. A continuació s'eliminen les impureses de baixa massa molecular, proteïnes i RNA amb un rentat de la columna. Finalment s'elueix el DNA plasmídic.

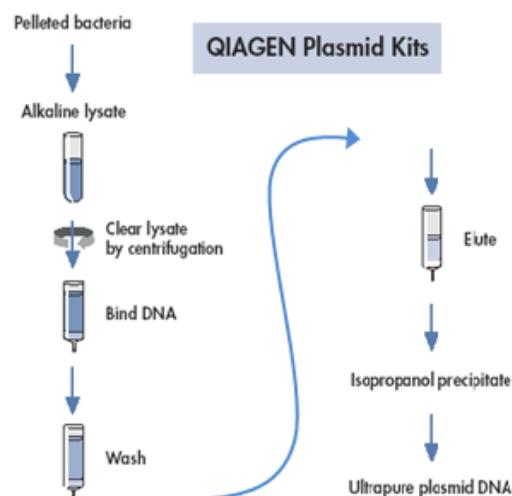


Fig. 3.44. Esquema del procés de la Midi-Prep.

3.12. Microscòpia de força atòmica (AFM)

La microscòpia de força atòmica (AFM, de l'anglès *Atomic Force Microscopy*) és un mètode que permet la visualització del perfil d'una superfície en tres dimensions amb un detall d'escala nanomètrica. L'AFM pot escanejar tot tipus de materials incloent estructures biològiques. Per a cada punt o píxel de la matriu bidimensional de la superfície de la mostra es realitza una mesura de l'alçada amb una sonda de força. D'aquesta manera s'obté finalment una imatge topogràfica de la superfície d'una mostra. Típicament, la mesura de l'alçada es mostra amb colors variant de més fosc per a alçades més baixes i més brillant per a les més altes, amb un gradient entre elles [126].

La sonda de força en la que es basa aquest mètode és una petita punta muntada a un *cantilever* que mesura la petita força d'interacció entre la punta i la superfície de la mostra, normalment de menys de 20 nN. El *cantilever* actua com un ressort que pot ser lleugerament desviat per aquestes petites forces. La desviació és mesurada amb un sistema òptic que produeix un senyal elèctric (fig. 3.45).

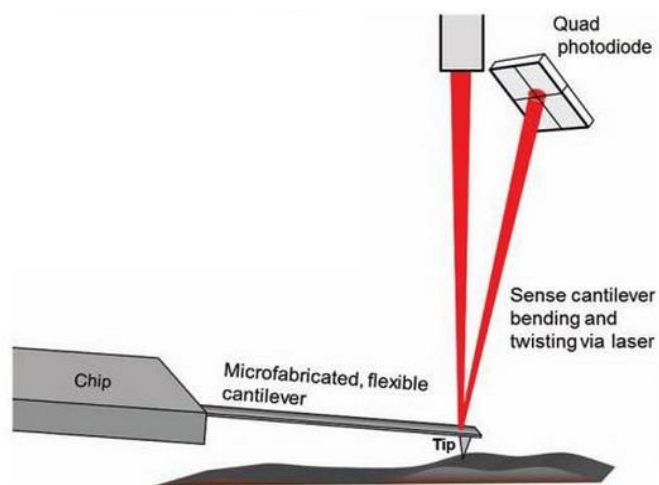


Fig. 3.45. Funcionament de l'AFM.

La interacció de la punta sobre la superfície de la mostra produeix unes forces (normal i lateral) que són mesurades pel gir i el doblegament del *cantilever* amb un sistema òptic de detecció. Imatge extreta de [126].

L'AFM pot ser realitzat de diferents maneres depenent del grau en què la punta està en contacte amb la superfície de la mostra: mode de contacte (la punta està en contacte tot el temps), mode intermitent o *tapping* (la punta oscil·la tocant intermitentment la mostra) o mode de no-contacte (la punta no toca la mostra). El mode intermitent o *tapping* és el mode més àmpliament usat i es fa servir per a materials tous com polímers i mostres biològiques. Aquest mode augmenta la resolució i evita tant el dany de la mostra per les forces laterals com la formació d'artefactes [127].

En aquest treball hem fet servir la microscòpia de força atòmica per a la visualització de les mostres de DNA plasmídic amb inserts ATs i poder observar la formació de transicions estructurals. Hem pogut realitzar aquests experiments a les instal·lacions del Grup de Biofísica de Biomembranes del Departament de Físicoquímica de la UB amb la col·laboració del Dr. Oscar Domènech i el responsable del grup Dr. Jordi Hernández Borrell.

Hem fet servir el mètode *tapping* per a escanejar les mostres fent servir una punta de nitrur de silici (Si_3N_4) per a la sonda en un aparell NanoScope V (Veeco). Hem preparat la mostra en un suport de mica de la següent forma:

- Exfoliar el suport de mica adherint i desadherint un tros de cinta adhesiva. D'aquesta manera la superfície on es diposita la mostra queda neta i carregada negativament.
- Dipositar 10 μl de la mostra de DNA plasmídic en el tampó 1x TAEMg (aproximadament a 1 $\text{ng}/\mu\text{l}$). Aquests ions Mg^{+2} fan de pont d'unió entre la mica carregada negativament i el DNA, fent que aquest quedi immobilitzat.
- Deixar absorbir 2 minuts.
- Netejar la superfície amb unes gotes d'aigua.
- Assecar les mostres amb aire comprimit.
- Escanejar el més aviat possible.

4. Resultats i discussió

4.1. Obtenció de proteïnes

4.1.1. Obtenció de les proteïnes HMGA

Les proteïnes HMGA1a i HMGA1b tenen una seqüència aminoacídica idèntica excepte pel fet que l'HMGA1b conté una deleció d'11 aminoàcids entre l'*AT-hook I* i l'*AT-hook II* (v. fig. 4.1 i fig. 1.6 de l'apartat 1.1.1). En el nostre laboratori hem treballat tant amb les proteïnes senceres com amb fragments de les mateixes. Concretament, en aquest treball s'han estudiat tant les HMGA1a i HMGA1b senceres com la forma proteica HMGA1b(1-79). L'HMGA1b(1-79) conté els tres *AT-hooks* i únicament difereix de la sencera en què no presenta la cua acídica carregada negativament. Amb l'eliminació d'aquesta zona acídica s'espera que aquesta forma cristal·litzi més fàcilment amb els oligonucleòtids.

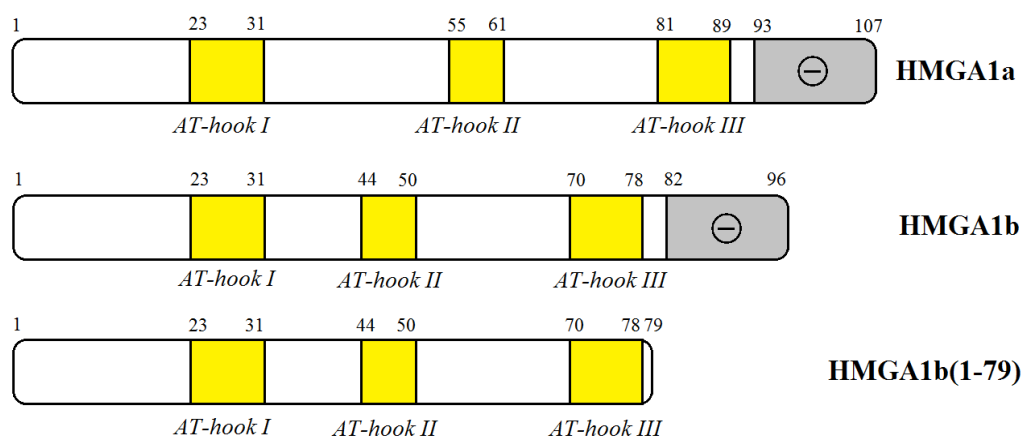


Fig. 4.1. Esquema de les proteïnes HMGA1a, HMGA1b i HMGA1b(1-79).

Ressaltat en groc els dominis *AT-hook* (carregats positivament) i en gris la regió acídica C-terminal.

L'obtenció de les proteïnes HMGA directament de cèl·lules humanes presenta moltes limitacions, és per aquest motiu que hem expressat les proteïnes d'interès en bacteris *E. coli* (v. apartat 3.2.2). En el cas de les proteïnes HMGA1a i HMGA1b humanes ja disposàvem al nostre laboratori del plasmidi que conté el gen que codifica aquestes proteïnes, en canvi el nou mutant HMGA1b(1-79) es va obtenir mitjançant la tècnica de la mutagènesi dirigida (v. apartat 3.2.1).

4.1.1.1. Expressió i purificació de la proteïna HMGA1a

Expressió en E. coli

Hem fet diverses expressions per a l'obtenció dels estocs necessaris de proteïna HMGA1a per tal de purificar-la i fer-la servir per a l'obtenció de cristalls amb oligonucleòtids. Generalment aquestes expressions s'han fet a partir d'un volum d'1 litre de cultiu bacterià. A continuació es detallen els resultats d'una d'aquestes expressions com a exemple.

Es va transformar la soca Rosetta(DE3)pLysS d'*E. coli* amb el plasmidi contenint el gen de l'HMGA1a (v. 3.2.2). Una de les colònies transformades es va utilitzar per a obtenir un cultiu líquid saturat a partir del qual es va inocular 1 litre de cultiu. L'expressió de la proteïna es va induir quan el cultiu presentava una $DO_{600} = 0,6-0,7$ per addició d'IPTG (v. annex C). El cultiu es va incubar durant tres hores més durant les quals els bacteris sintetitzen la nostra proteïna. Durant el període d'inducció es van agafar alíquotes de control a temps 0 h, 1,5 h i 3 h per a llegir-ne l'absorbància i fer controls electroforètics. A la fig. 4.2 es mostra l'electroforesi en un gel de poliacrilamida-SDS corresponent a aquesta inducció. S'observa com a l'augmentar el temps d'inducció augmenta la quantitat de proteïna expressada. Així mateix s'aprecia com l'HMGA1a s'obté també com a forma truncada durant l'expressió.

Passades les tres hores d'inducció es van recuperar els bacteris per centrifugació. Aquest sediment bacterià es va lisar en presència de 5% $HClO_4$ /0,5% Tritó X-100 i es va tractar tal i com es descriu a l'apartat 3.2.3 per tal de recuperar-ne la proteïna HMGA1a. Degut a la propietat de les HMGAs de poder ser solubilitzades amb 5% d'àcid perclòric, aquest tractament ens permet obtenir un extracte força enriquit en la proteïna HMGA1a.

Al carril EE (extracte enriquit) de la fig. 4.2 podem apreciar clarament com al finalitzar el procés de tractament del sediment bacterià obtenim una mostra on s'han eliminat la gran majoria de proteïnes d'origen bacterià mentre que obtenim majoritàriament la proteïna HMGA1a (comparar els carrils 3 h i EE de la fig. 4.2). De fet, les dues bandes intenses que observem a l'extracte enriquit corresponen concretament a la proteïna HMGA1a sencera (banda superior) i a una forma truncada (banda inferior) que es forma ja dins del bacteri. Aquesta forma truncada també es va considerar d'interès pel nostre treball i per tant també hem perseguit la seva purificació.

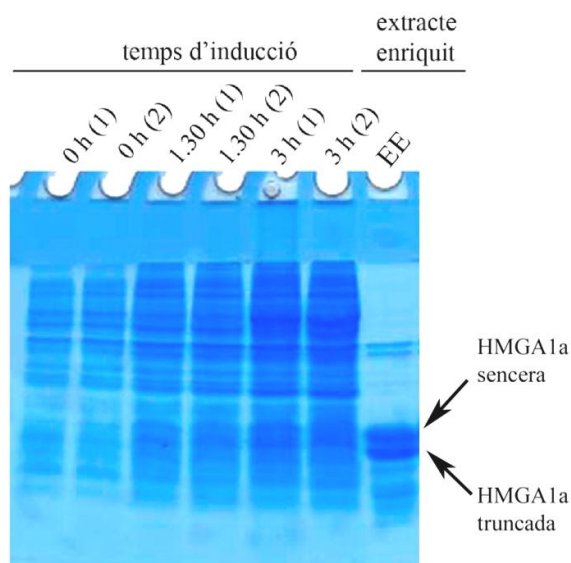


Fig. 4.2. Gel de poliacrilamida-SDS de la inducció i de l'extracte enriquit de l'HMGA1a.

El cultiu es va repartir en dos erlenmeyers (1 i 2) que es van controlar individualment. L'EE (extracte enriquit) correspon a la mostra enriquida en proteïna HMGA1a després del tractament del sediment bacterià.

Per tal de separar la forma sencera de la forma troncada de l'HMGA1a i alhora eliminar la resta de proteïnes bacterianes que ens han quedat a la mostra el procés de purificació es va continuar utilitzant la tècnica de la cromatografia de bescanvi iònic.

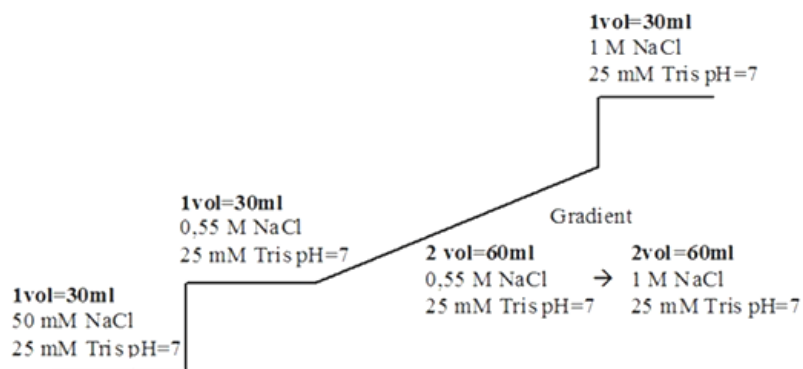
Purificació per cromatografia de bescanvi catiònic

Per a portar a terme la purificació de la proteïna HMGA1a hem fet servir la cromatografia de bescanvi catiònic (v. 3.2.4). Inicialment es va partir de les condicions descrites per Reeves i Nissen [81]. Segons aquests autors, la proteïna HMGA1a recombinant es pot purificar amb una columna de bescanvi catiònic del tipus MACRO-PREP 50S de la marca comercial Bio-Rad, amb l'ús de dissolucions de $\text{KH}_2\text{PO}_4/\text{KCl}$ pH=7 amb concentracions creixents de KCl. En el nostre laboratori hem fet servir una columna de 20 cm de longitud i 1,5 cm de diàmetre intern amb un volum empaquetat de 30 ml de la resina MACRO-PREP 25S (Bio-Rad). Hem realitzat diferents cromatografies amb l'ús de les dissolucions de $\text{KH}_2\text{PO}_4/\text{KCl}$ pH=7, i variant les concentracions de KCl durant el gradient.

Un problema que es va fer evident des del principi a l'utilitzar les dissolucions de ($\text{KH}_2\text{PO}_4/\text{KCl}$) durant la cromatografia de bescanvi catiònic és que la proteïna HMGA1a s'elueix a elevades concentracions de KCl. El problema que ens suposava era doble. Per una banda, la presència d'elevats continguts d'aquestes sals dificulta tant el procés de carregar les mostres als gels com la nitidesa de les bandes resultants (i, per tant, la correcta identificació de les bandes). D'altra banda, cal fer diàlisi molt més extensives per tal d'assegurar-nos l'eliminació d'aquestes sals (requisit necessari per a poder dur a terme els assajos cristal·logràfics).

Per aquest motiu es va assajar també la purificació de l'HMGA1a utilitzant la mateixa columna però amb tampons Tris/NaCl pH=7, que evitaven aquests problemes, i utilitzant concentracions creixents de NaCl per a l'elució. Justament la proteïna HMGA1a feta servir pels assajos de cristal·lització que es presenten en aquest treball va ser obtinguda d'una cromatografia realitzada amb aquestes condicions i amb la que es va purificar la forma HMGA1a truncada. Aquesta forma difereix de la forma sencera en què no presenta part de la cua àcida, però sí conserva els tres dominis *AT-hook*. Així, aquesta forma truncada de la proteïna HMGA1a ens permet estudiar igualment les interaccions dels *AT-hooks* amb oligonucleòtids rics en ATs i a més a més presenta més càrrega neta positiva (més afinitat pel DNA).

La mostra carregada en aquesta cromatografia era un extracte enriquit en proteïna HMGA1a recombinant obtingut a partir d'un litre de cultiu segons el procediment explicat a l'apartat anterior. El següent esquema mostra els diferents passos seguits durant la cromatografia. El flux va ser d'1 ml/min i es van recollir fraccions de 3 ml.



A la fig. 4.3 es mostra el cromatograma obtingut d'aquesta columna. L'anàlisi electroforètica dels pics detectats a 220 nm va permetre identificar que l'HMGA1a eluïa al pic principal (v. fig. 4.3). Pel que fa als pics secundaris corresponents a les fraccions de la 5 fins a la 30, els corresponents gels de poliacrilamida-SDS van mostrar com no contenen proteïna HMGA1a (no mostrat). És important destacar que els pics del cromatograma corresponents a la proteïna sencera i a la truncada es solapen, no obtenint així una separació perfecta entre les dues. Al gel de la fig. 4.3 podem comprovar com la proteïna sencera elueix abans que la truncada en les cromatografies de bescanvi catiònic, ja que està globalment carregada menys positivament. Així, les primeres fraccions eluïdes, de la 39 a la 42, contenen proteïna sencera però amb força impureses. Les fraccions 43 i 44 contenen tant proteïna sencera com truncada. I, per últim, la proteïna HMGA1a truncada apareix majoritàriament a partir de la fracció 45.

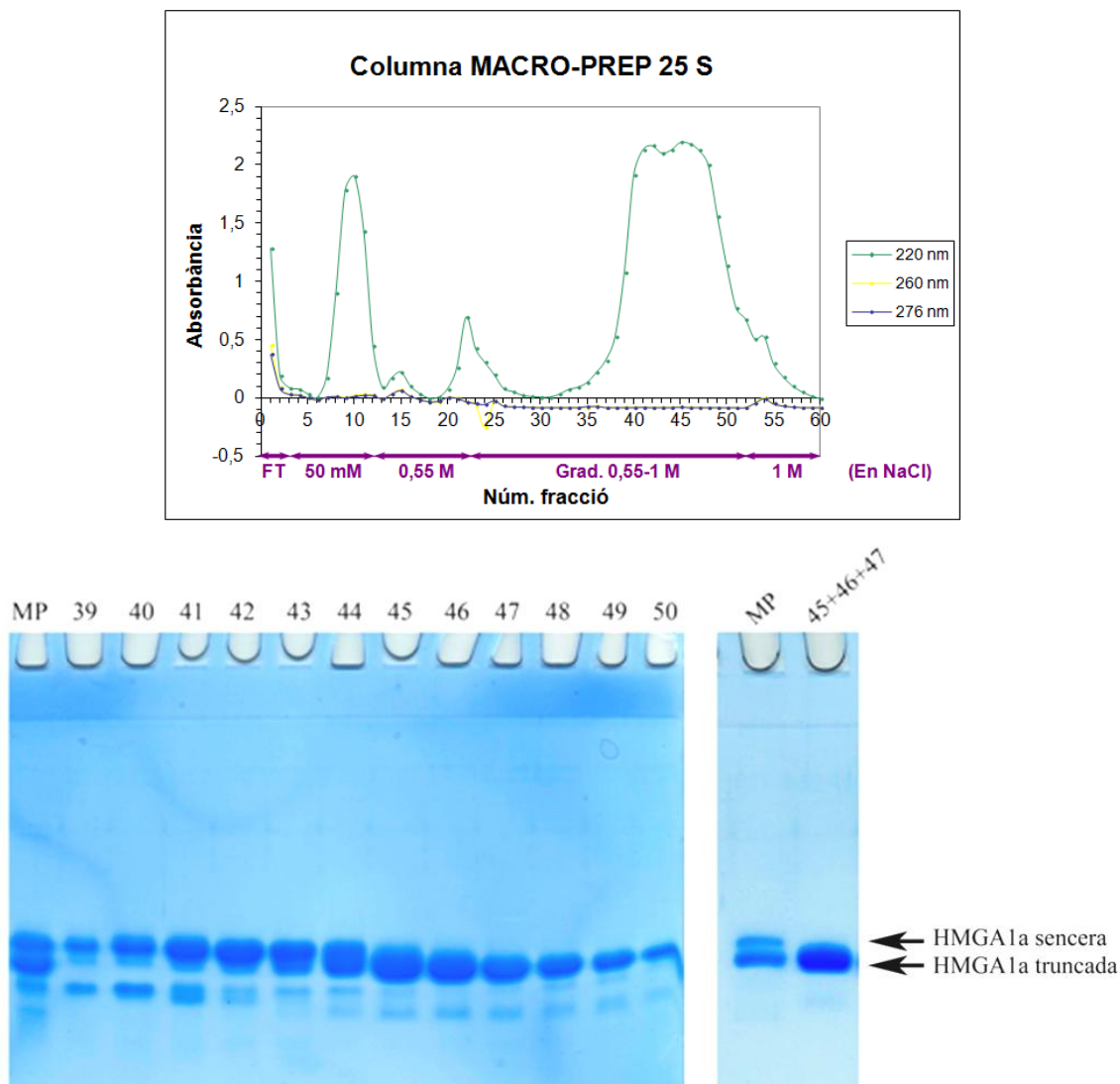


Fig. 4.3. Cromatograma i gels de poliacrilamida-SDS corresponents a la columna MACRO-PREP 25S (Tris/NaCl pH=7) per a la proteïna HMGA1a.

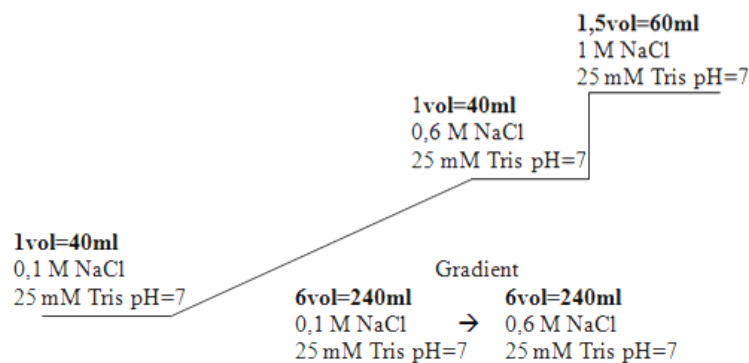
MP, mostra de partida carregada a la columna. El gel de la dreta corresponent a les fraccions 45+46+47 un cop reunides i dialitzades.

Les fraccions 45, 46 i 47 es van reunir i dialitzar. Com s'observa al gel de la dreta de la fig. 4.3, aquestes fraccions contenen HMGA1a truncada força pura. Aquesta mostra de proteïna HMGA1a es va concentrar amb dispositius de concentració (v. 3.5.1) per tal de fer-la servir per a assajos cristal·logràfics. La concentració i el volum d'aquesta mostra es detalla a l'apartat 4.1.3.

4.1.1.2. Purificació de la proteïna HMGA1b

Purificació per cromatografia de bescanvi catiònic

L'extracte enriquit en proteïna HMGA1b procedent d'una expressió (2 litres de cultiu) realitzada de forma anàloga a la de l'HMGA1a va ser purificat per cromatografia de bescanvi catiònic. Per tal d'intentar millorar l'eficàcia de la purificació en aquesta ocasió es va fer servir una resina diferent, la CM52 (carboximetil cel·lulosa), amb els mateixos tampons Tris/NaCl pH=7 i aplicant un gradient més suau. L'esquema d'aquesta cromatografia es mostra a continuació:



Es va fer circular per la columna un flux constant d'1 ml/min i es van recollir fraccions de 4 ml. A la fig. 4.4 podem observar el cromatograma corresponent a aquesta cromatografia i el gel de poliacrilamida-SDS obtingut a l'analitzar electroforèticament les fraccions que presentaven major absorbància a 220 nm.

L'aplicació d'un gradient més llarg i el canvi de resina va permetre una millor separació dels pics d'absorbància corresponents a la proteïna sencera (eluïda primer en el gradient) i la proteïna truncada (eluïda en segon lloc). No obstant això, les fraccions amb contingut majoritari de proteïna HMGA1b sencera (de la fracció 73 a la 80 en el gel) presenten certes impureses d'altres proteïnes per sota d'aquesta banda. Per a aquest motiu i per a assegurar-nos que no quedessin traces de la forma truncada es va decidir repurificar les fraccions de la 72 a la 80 d'aquesta cromatografia.

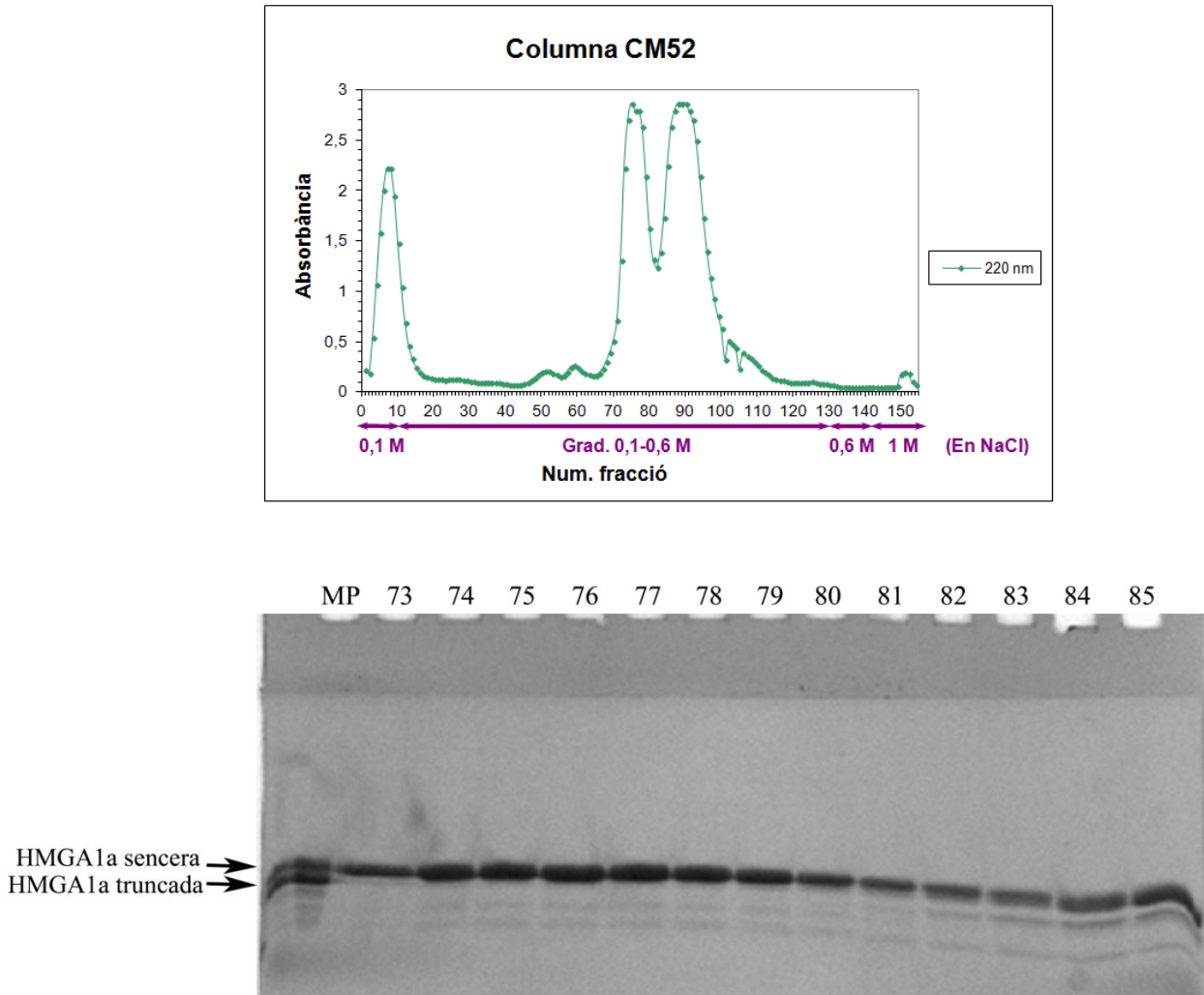


Fig. 4.4. Cromatograma i gel SDS-PAGE corresponents a la columna CM52 (Tris/NaCl pH=7) per a la proteïna HMGA1b.

MP, mostra de partida carregada a la columna.

L'esquema dels tampons per a la cromatografia és igual que l'anterior però vam fer servir 0,2 M NaCl/Tris pH=7 com a tampó d'equilibrat i primer tampó del gradient. No és va aconseguir separar la proteïna truncada de la sencera (v fig. 4.5). Així, les primeres fraccions eluïdes (de la 42 a la 55 en aquests gels) contenen tant proteïna sencera com truncada juntament amb d'altres impureses.

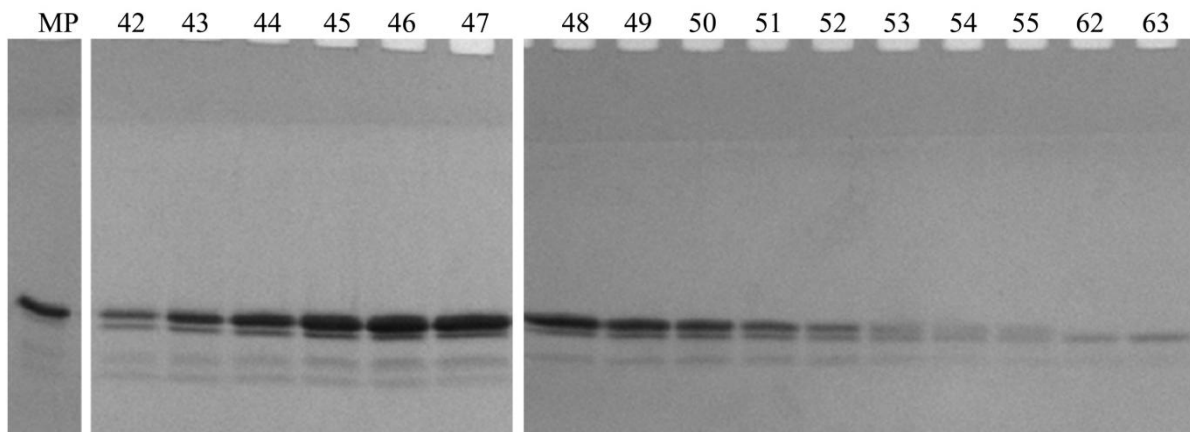


Fig. 4.5. Gels SDS-PAGE corresponents a la columna de repurificació CM52 (Tris/NaCl pH=7).

MP, mostra de partida carregada a la columna.

Malgrat les dues cromatografies addicionals (amb la columna MACRO-PREP aquest cop) que es van realitzar posteriorment, no es va arribar a obtenir una mostra prou pura ni de proteïna HMGA1b sencera ni de proteïna HMGA1b truncada. Les fraccions amb proteïna sencera majoritària presenten una banda inferior de proteïna truncada amb d'altres impureses mentre que les fraccions amb proteïna truncada majoritària s'obtenen en molt poca quantitat.

Davant d'aquestes dificultats per a la obtenció de la proteïna HMGA1b pura es va creure oportú la construcció d'un mutant lliure de la cua àcida, el fragment HMGA1b(1-79). D'aquesta manera s'esperava evitar la coexpressió amb forma truncades a més d'obtenir una forma amb major potencial d'unió al DNA.

4.1.1.3. Obtenció de la proteïna HMGA1b(1-79): mutagènesi, expressió i purificació

Mutagènesi dirigida

Per a obtenir la proteïna HMGA1b(1-79) es va partir del plasmidi que codifica la forma sencera HMGA1b i es va introduir un codó stop (TAG) en la posició corresponent a l'aminoàcid 80 (GAG) mitjançant mutagènesi dirigida. D'aquesta manera, al sintetitzar-se la proteïna s'aturarà el procés en l'aminoàcid 80 obtenint així el mutant de l'aminoàcid 1 al 79 (amb els tres *AT-hooks* però sense la cua àcida) (v. fig. 4.1). Un cop dissenyats i adquirits els *primers* adients, vam obtenir el mutant HMGA1b(1-79) utilitzant el *QuickChangeTM Site-Directed Mutagenesis Kit* (Stratagene) (v. 3.2.1).

Es va aïllar el DNA plasmídic mutat (v. 0) i es va portar a seqüenciar al Parc Científic de Barcelona per a comprovar que la mutació s'hagués introduït correctament i que no s'hagués produït cap canvi addicional en la seqüència nucleotídica. El cromatograma obtingut de la mostra portada a seqüenciar és el que es mostra a la fig. 4.6. Segons l'assignació automàtica de bases que fa el programa veiem que el triplet original GAG ha passat a TAT i no al codó stop TAG. És a dir, a més de la mutació de la primera G a T, que era l'únic canvi que volíem introduir, segons la interpretació del programa l'altra G també ha passat a T, fet inesperat.

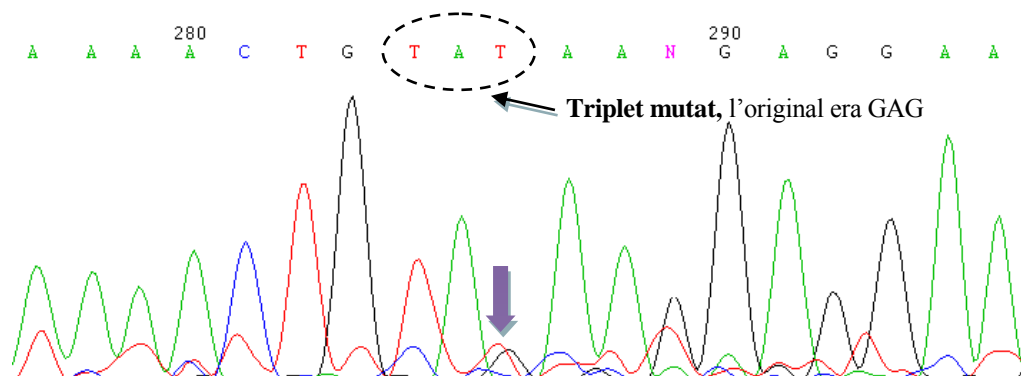


Fig. 4.6. Cromatograma de la seqüenciació del DNA plasmídic mutat.

Observem el fragment de DNA on s'hauria d'haver produït la mutació i encerclat el triplet mutat.

No obstant això, si ens fixem bé en el cromatograma, la tercera base del codó mutat presenta dos pics d'alçada similar corresponents a una base T i a una base G (fletxa a la fig. 4.6) fet que pot haver originat un error en la interpretació automàtica. Tenint en compte que la base original era la G i que el pic de T és similar als altres pics de T que hi ha al *background* del cromatograma, es va considerar que la interpretació correcta hauria de ser la G, tal i com esperàvem. A més a més, en el cas que el triplet mutat fos realment TAT tindríem una tirosina en la posició 80 i a l'expressar la proteïna obtindríem la seqüència sencera amb aquest aminoàcid 80 mutat. L'expressió de la proteïna va confirmar que aquest no era el cas (v. més endavant).

Expressió en E. coli

Hem fet diferents expressions del mutant HMGA1b(1-79) seguint el protocol explicat en l'apartat 3.2.2. Després de realitzar la transformació en cèl·lules de la soca Rosetta(DE3)pLysS d'*Escherichia coli* amb el DNA plasmídic modificat per mutagènesi dirigida, aquestes es van fer créixer en medi de cultiu LB i a continuació es va iniciar la síntesi de la proteïna amb la inducció amb IPTG. Una vegada transcorregudes les tres hores corresponents a la inducció es va centrifugar el cultiu. El sediment bacterià es va tractar per tal d'extreure la proteïna HMGA1b(1-79) de l'interior dels bacteris i eliminar part de les impureses d'origen bacterià (v. 3.2.3). Al final

del tractament vam obtenir un extracte enriquit en la proteïna d'interès que vam analitzar en un gel de poliàcrilamida-SDS juntament amb les mostres a l'inici i al final de la inducció (v. fig. 4.7).

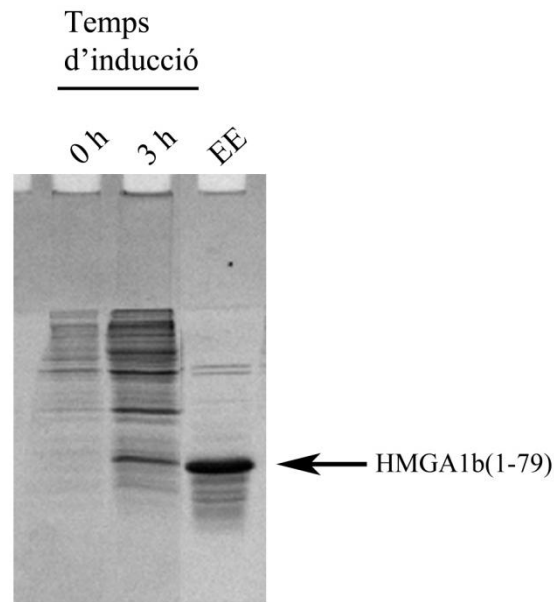


Fig. 4.7. Gel de poliàcrilamida-SDS de la inducció i de l'extracte enriquit de la proteïna HMGA1b(1-79).

EE, extracte enriquit obtingut al final del tractament del sediment bacterià.

Com hem explicat abans, la seqüenciació no demostrava al 100% que la mutació s'hagués produït correctament. Al carregar, però, la proteïna mutada juntament amb un control de la proteïna HMGA1b original en un gel de poliàcrilamida-SDS, es pot comprovar com, efectivament, la forma mutada presenta una major mobilitat electroforètica, tal i com correspon a una proteïna de menor mida (v. fig. 4.8).

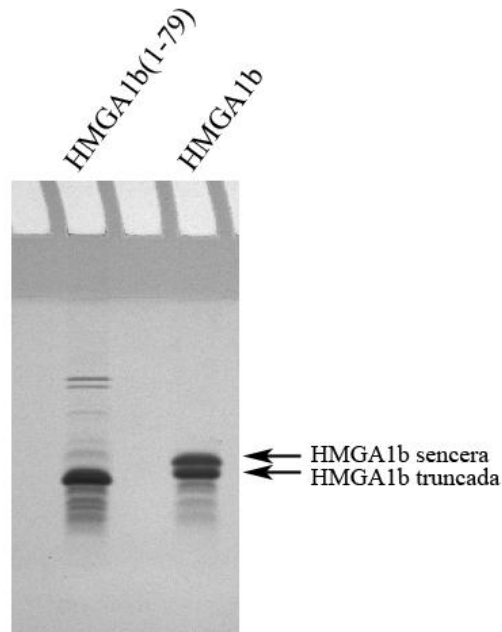
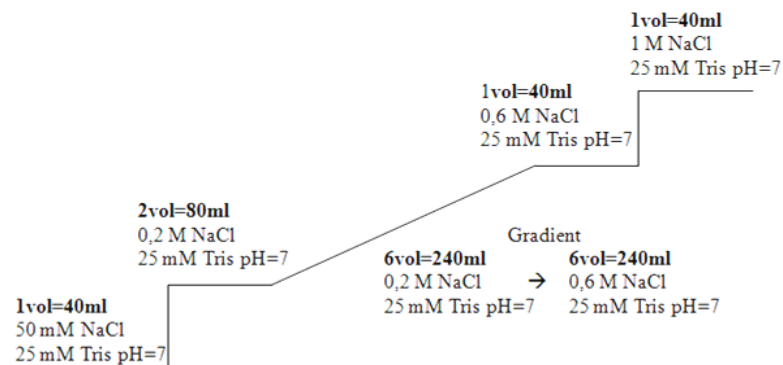


Fig. 4.8. Gel de poliacrilamida-SDS per a comparar la forma HMGA1b(1-79) obtinguda per mutagènesi dirigida amb l'HMGA1b original.

Les dues bandes intenses que s'observen en el carril HMGA1b corresponen a la forma sencera i truncada que s'obtenen durant l'expressió de l'HMGA1b (v. més detalls al text).

Purificació per cromatografia de bescanvi catiònic

Per a la purificació en cromatografia de bescanvi catiònic del fragment HMGA1b(1-79) es van seguir unes condicions similars a les utilitzades per a la forma sencera, l'HMGA1b. Es va treballar amb la columna CM52, amb tampons Tris/NaCl pH=7, un flux constant d'1 ml/min, fraccions eluïdes de 4 ml, però amb un perfil de concentracions una mica diferent. A continuació es mostra aquest esquema:



A la fig. 4.9 es presenta una de les diferents cromatografies realitzades per a la purificació de la proteïna HMGA1b(1-79) i el seu corresponent gel de poliacrilamida-SDS. Com amb aquest fragment ja no tenim el problema de la coexpressió amb formes truncades, les fraccions eluïdes d'aquesta cromatografia es mostren força pures.

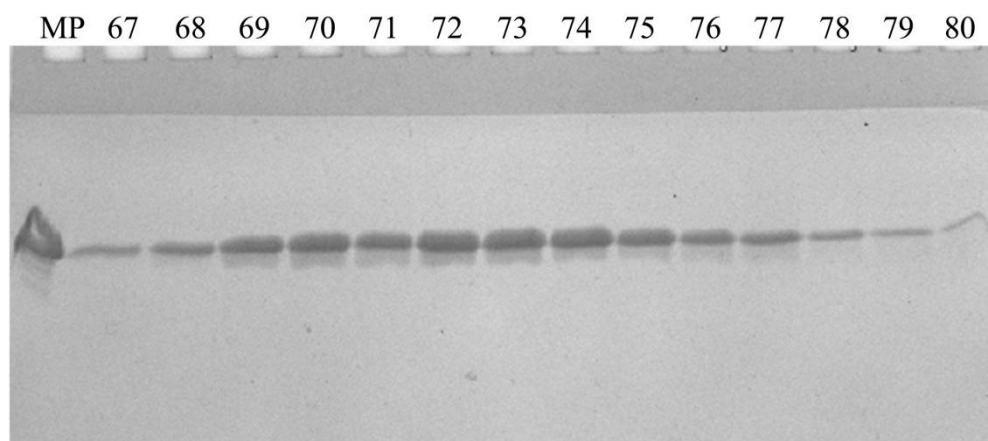
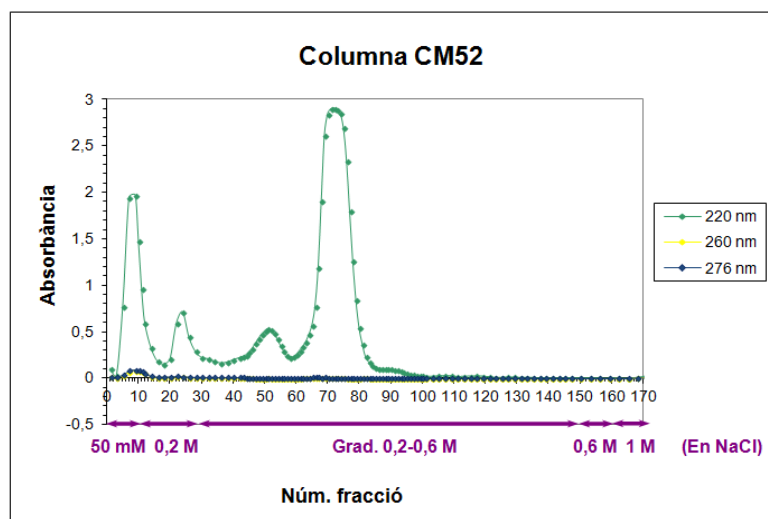


Fig. 4.9. Cromatograma i gel SDS-PAGE corresponents a la columna CM52 (Tris/NaCl pH=7) per a la proteïna HMGA1b(1-79).

Al gel es mostren les fraccions corresponents al pic principal del cromatograma. MP, mostra de partida.

Purificació per cromatografia d'exclusió molecular

Com a l'últim pas de poliment en la purificació del fragment HMGA1b(1-79) es va recórrer a la cromatografia d'exclusió molecular utilitzant una columna Superdex 75 10/300 GL en un sistema d'FPLC AKTA Purifier (GE Healthcare). Les mostres més pures eluïdes de la columna de bescanvi catiònic van ser reunides i concentrades fins a una concentració d'uns 20 µg/µl (~2

mM). Per tal de tenir una bona relació entre el volum de la mostra i de la columna i optimitzar així la resolució, va ser necessari dur a terme diferents cromatografies, carregant cada cop un petit volum de la mostra concentrada. Per això, es van carregar volums d'entre 40 i 50 μ l de la mostra concentrada en successives cromatografies d'exclusió molecular utilitzant un loop de 100 μ l. Es van eluir fraccions de 0,5 ml fent circular 1,5 volums de columna (36 ml) del tampó amb un flux de 0,5 ml/min.

Un exemple representatiu d'un dels cromatogrames de la columna Superdex 75 en el sistema FPLC per a la proteïna HMGA1b(1-79) s'observa la fig. 4.10. Per a cadascuna de les múltiples cromatografies realitzades es va obtenir un resultat molt similar.

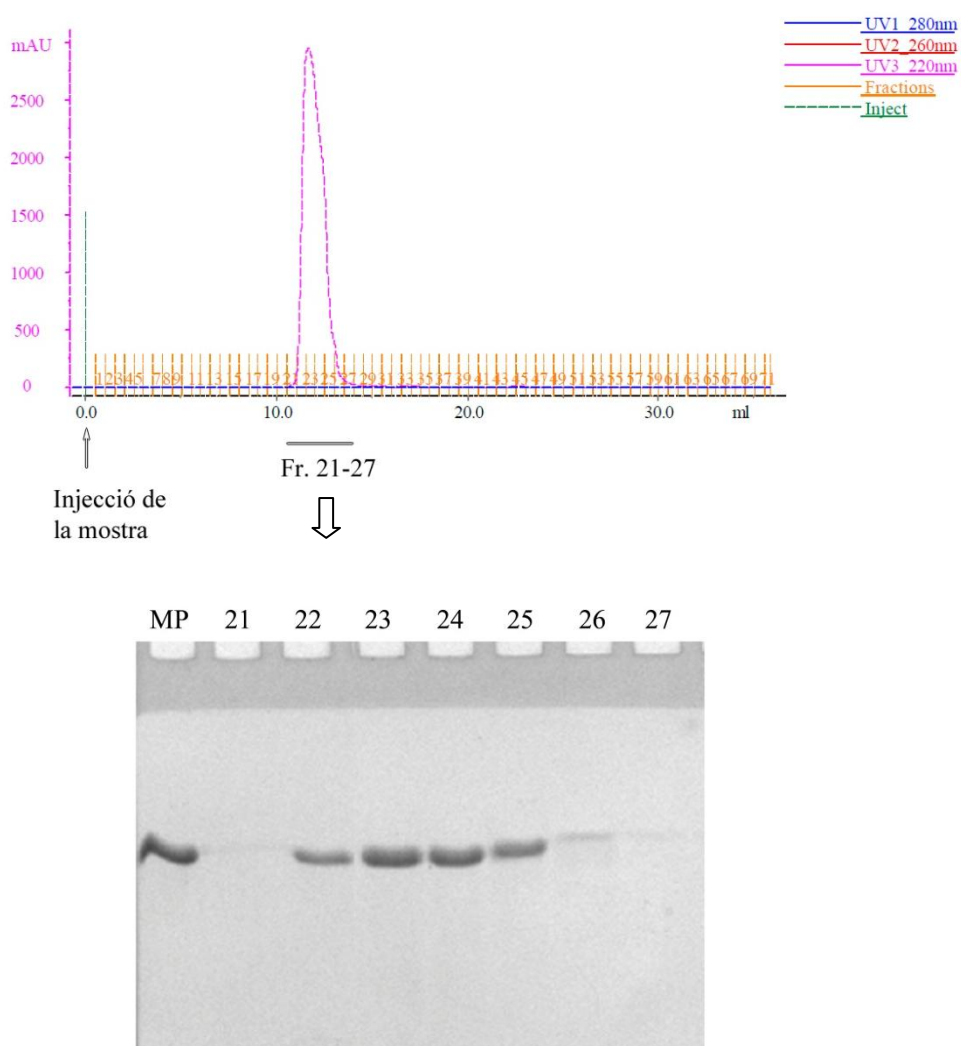


Fig. 4.10. Cromatograma i gel SDS-PAGE corresponents a la columna Superdex 75 per a l'HMGA1b(1-79).

MP, mostra de partida carregada a la columna.

Un cop reunides i concentrades les fraccions més pures eluïdes de les diferents columnes d'exclusió molecular es van obtenir els tres estocs de mostres concentrades que es mostren a l'anàlisi electroforètica de la fig. 4.11. Les concentracions d'aquests estocs es troben especificades a la taula 4.1 de l'apartat 4.1.3.

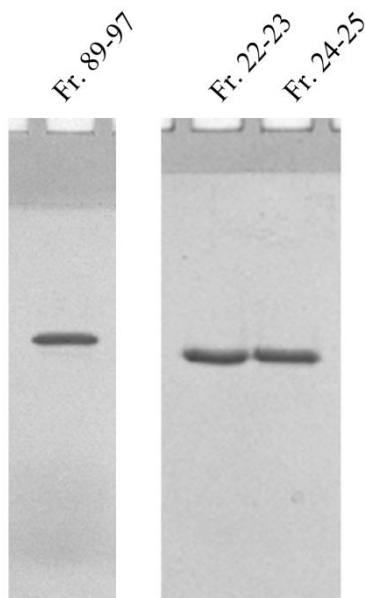


Fig. 4.11. Gel SDS-PAGE de les fraccions pures i concentrades de la proteïna HMGA1b(1-79) obtingudes després de la repurificació per exclusió molecular.

4.1.2. Obtenció de les proteïnes HMGB

La proteïna HMGB1 de mamífer conté dos dominis HMG-box, el Box A i el Box B, i una cua acídica a l'extrem C-terminal (v. fig. 4.12). Tots dos dominis HMG-box presenten una estructura tridimensional en forma de L constituïda per tres hèlixs alfa, però presenten petites diferències geomètriques i de seqüència (fig. 1.12) que els hi confereixen propietats d'unió al DNA característiques. Per tal d'abordar la cristal·lització de la proteïna HMGB1 es va decidir treballar amb els dos dominis per separat (Box AS, Box BS) i un fragment (Box ABS) contenint els dos dominis i el *linker* (fragment entre tots dos dominis) (fig. 4.12). Al tractar-se de constructes de seqüència més curta i sense la cua acídica s'espera que siguin més fàcilment cristal·litzables que la forma sencera de la proteïna. Per a facilitar-ne la seva purificació els fragments Box AS, Box BS i Box ABS es van expressar fusionats amb la proteïna Glutatió S-Transferasa en l'extrem N-terminal. En el cas de la proteïna NHP6A, que conté només un HMG-box i una cua N-terminal carregada positivament (fig. 4.12), es va treballar directament amb la forma sencera.

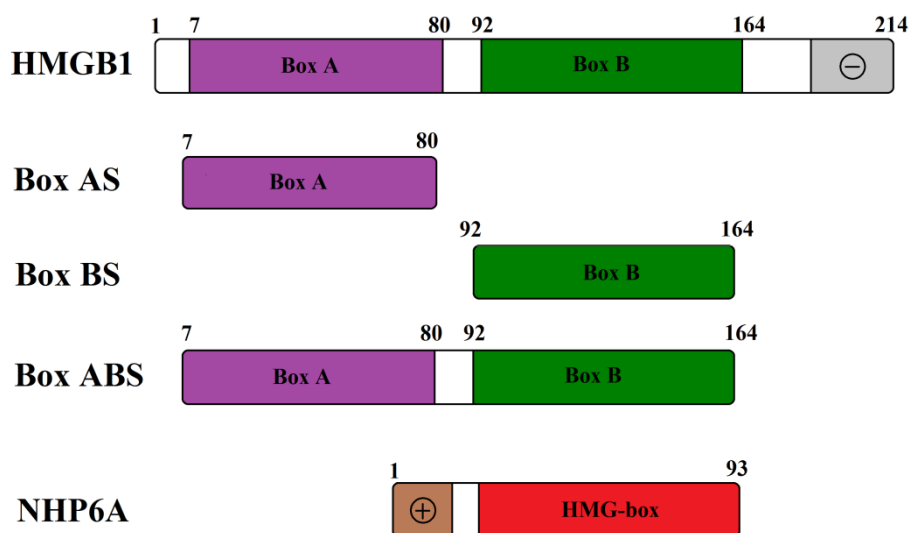


Fig. 4.12. Esquema de la proteïna HMGB1 (i els seus fragments) i de la proteïna NHP6A.

Els dominis Box A i Box B es representen en lila i verd respectivament, mentre que el domini HMG-box de l'NHP6A en vermell. En gris la regió acídica C-terminal de l'HMGB1 i en marró la regió bàsica N-terminal de l'NHP6A. Per a facilitar la comparació entre estructures resoltes no s'ha comptat la metionina inicial en la numeració dels residus de la proteïna HMGB1 (per analogia amb la numeració de l'estructura 1CKT). Els constructes tenen una denominació amb una S (*short*) al final per a diferenciar-los d'unes altres formes disponibles al laboratori que contenien a més a més els residus que flanquegen els dominis (formes L, *long*). Les masses moleculars de les proteïnes amb les que hem treballat es recullen a l'annex A.

Tot seguit es mostren els resultats obtinguts en els diferents passos de la purificació de les proteïnes HMG-box: Box AS (v. 4.1.2.1), Box BS (v. 4.1.2.2), Box ABS (v. 4.1.2.3) i NHP6A (v. 4.1.2.4). Els estocs finals de proteïnes pures i concentrades aconseguits estan recollits a la taula 4.1 de l'apartat 4.1.3.

4.1.2.1. Expressió i purificació de la proteïna HMGB1 Box AS

Expressió en E. coli

Al llarg del treball es van dur a terme diferents expressions del fragment Box AS realitzades tal i com es descriu a l'apartat 3.3.1. A la fig. 4.13 s'observa un exemple d'una expressió on s'han analitzat electroforèticament les mostres del cultiu bacterià corresponents a temps 0, 1 h 30' i 3 h post-inducció. La banda de la proteïna HMGB1 Box AS migra amb una massa molecular aparent de ~35000 degut a que s'expressa fusionada amb la Glutació S-transferasa (~26000 Da).

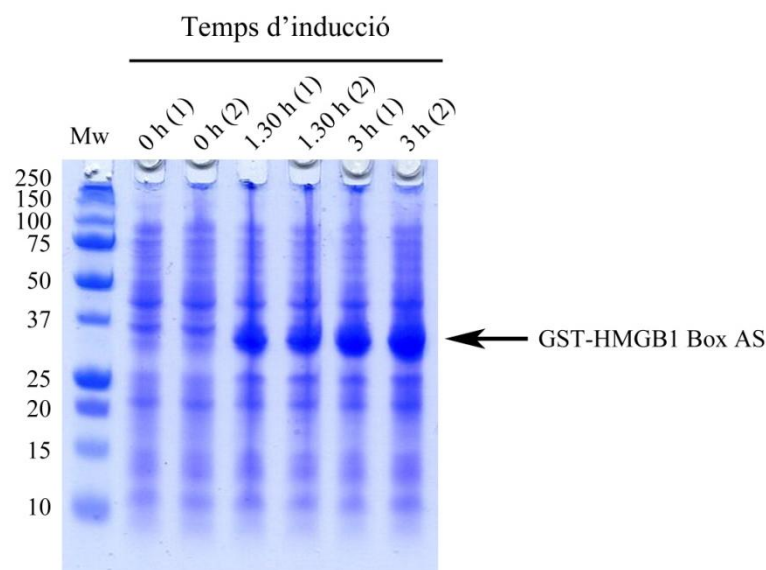


Fig. 4.13. Gel SDS-PAGE de les mostres analitzades durant la inducció de la proteïna HMGB1 Box AS.

El cultiu es va repartir en dos erlenmeyers (1 i 2) que es van controlar individualment. Mw, marcador de massa molecular (de l'anglès *molecular weight*).

Purificació per cromatografia

a) Cromatografia d'afinitat

Un cop la proteïna de fusió GST-HMGB1 s'ha extret de l'interior dels bacteris mitjançant la sonicació de les cèl·lules (v. 3.3.2), la mostra es centrifuga i el sobrenadant es tracta en una cromatografia d'afinitat utilitzant una columna amb grups glutatió on s'unirà la GST-Box AS expressada (v. 3.3.3). Després d'una incubació de mínim dues hores de la mostra en la resina de Glutació Sefarosa, s'elueix el *flow through* (FT), el qual contindrà les proteïnes no retingudes a la matriu. La proteïna de fusió GST-HMGB Box AS queda unida a la resina a través de la glutatió S-transferasa (GST) del seu extrem N-terminal. Per a poder eluir la nostra proteïna de la columna es fa una digestió amb l'enzim trombina, que talla la proteïna de fusió just després de la cua GST. Per tant, a continuació de fer un rentat de la columna amb el tampó adient per tal que actuï

la trombina (ThW), s'afegeix l'enzim i es deixa actuar tota la nit amb agitació a 4°C (v. 3.3.3). D'aquesta manera en els rentats posteriors (FTc, EIW1, EIW2 i EIW3) recuperarem la proteïna HMGB1 Box AS ara ja lliure de la cua GST (v. fig. 4.14).

A la fig. 4.14 podem veure el gel corresponent a la purificació per cromatografia d'afinitat de la proteïna GST-HMGB1 Box AS. La mostra de partida (MP) presenta una banda més intensa corresponent a la nostra proteïna de fusió i un gran nombre d'altres bandes corresponents a d'altres proteïnes d'origen bacterià. Aquestes altres proteïnes són principalment eliminades en les primeres fraccions eluïdes de la columna Glutatió Sefarosa 4B (FT1, FT2, W i ThW) ja que no s'uneixen a aquesta. Al gel SDS-PAGE podem observar que les fraccions FT1 i FT2 presenten també una mica de proteïna GST-HMGB1 Box AS degut a que s'ha sobrecarregat la columna. Un cop s'ha realitzat l'escissió de la cua GST amb la trombina, la proteïna HMGB1 Box AS (~8900 Da) és eluïda en les fraccions FTc, EIW1, EIW2 i EIW3 mentre que la cua GST segueix unida a la matriu de la columna (com es pot veure en el carril "beads" de la fig. 4.14).

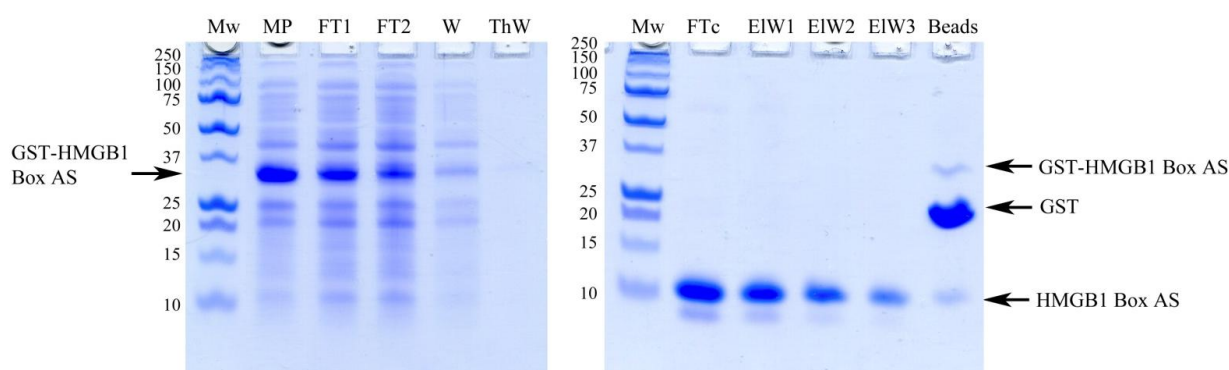


Fig. 4.14. Gel SDS-PAGE corresponent a la purificació GST per a l'HMGB1 Box AS.

Mw, marcador de massa molecular (de l'anglès *molecular weight*); Mp, mostra de partida carregada a la columna; FT, *flow through*; W, primer rentat de la columna (*wash*); ThW, rentat amb el tampó adient per a la digestió amb la trombina (*thrombin wash*); FtC, tampó eluït després de l'escissió (*flow through after cleavage*); EIW, fraccions eluïdes (*Elution wash*); Beads, mostra de les micropartícules que componen la matriu de la columna usada.

b) Cromatografia de bescanvi catiònic

Les fraccions FtC, EIW1, EIW2 i EIW3, que contenen aproximadament uns 15 mg de proteïna Box AS amb una mica d'impureses corresponents a formes truncades d'aquesta proteïna, es van reunir i es van carregar en una columna de bescanvi catiònic HiTrap SP FF en un sistema FPLC (AKTA Purifier, GE Healthcare). La mostra, d'uns 20 ml, va ser filtrada a 0,2 µm i carregada mitjançant tres injeccions amb un loop de 10 ml. La cromatografia va ser realitzada amb un flux constant de 2 ml/min, recollint fraccions de 2 ml i a una temperatura de 4°C. Es va aplicar un

gradient de 45 min per passar del tampó A (100 mM NaCl, 10 mM Tris pH=7,9, 1 mM EDTA, 1 mM DTT) fins al tampó B (1 M NaCl, 10 mM Tris pH=7,9, 1 mM EDTA, 1 mM DTT).

A la fig. 4.15 podem veure el cromatograma corresponent a una cromatografia d'aquest tipus per al fragment Box AS i un gel d' SDS corresponent al pic d'absorbàncies del cromatograma. Les mostres corresponents als pics de les injeccions de la MP no presentaven proteïna Box AS. En l'anàlisi electroforètica del pic d'absorbància a 280 nm obtingut durant el gradient s'observa com la forma truncada del Box AS apareix a l'inici del pic mentre que el fragment Box AS pròpiament s'elueix a continuació. La proteïna Box AS s'obté molt pura a partir de la fracció 24.

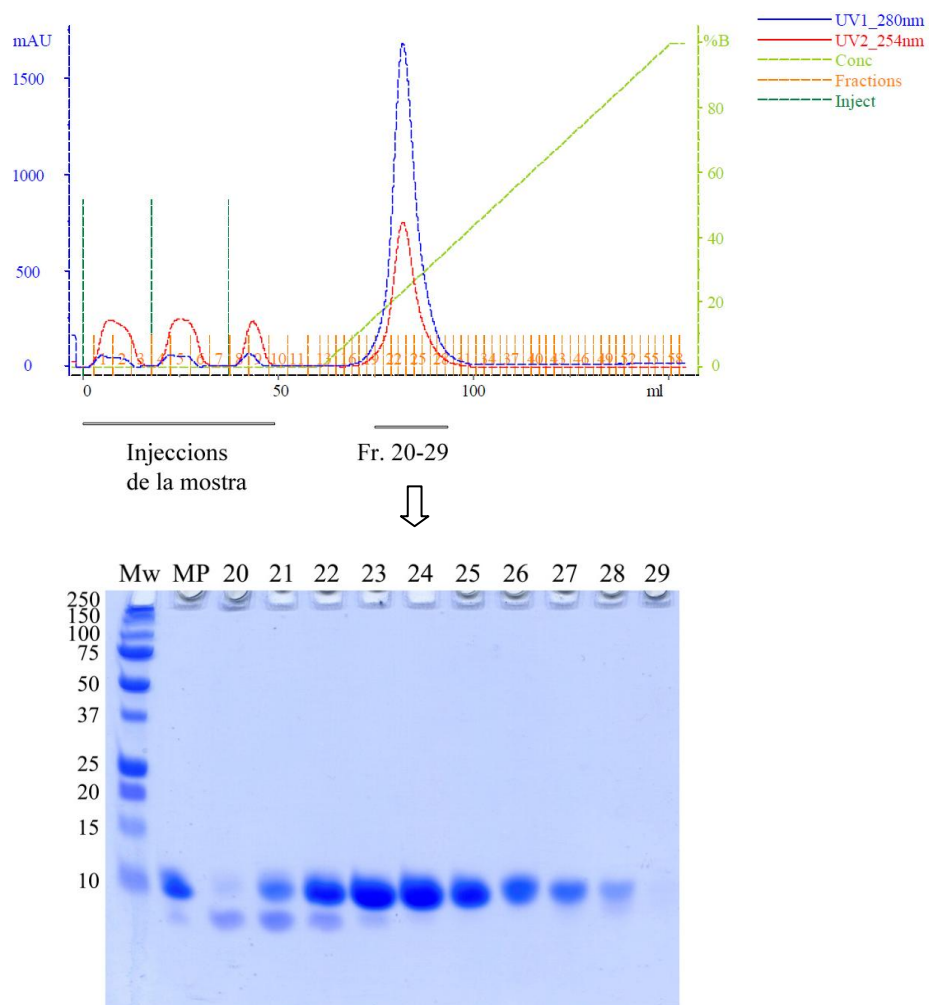


Fig. 4.15. Cromatograma i gel d' SDS corresponents a la columna HiTrap SP FF per a l'HMGB1 Box AS.

Mw, marcador de massa molecular (de l'anglès *molecular weight*); MP, mostra de partida carregada a la columna.

c) Cromatografia d'exclusió molecular

Les fraccions eluïdes de la columna de bescanvi que presentaven majoritàriament proteïna Box AS van ser sotmeses a cromatografia d'exclusió molecular en una columna HiLoad Superdex 75 (120 ml empaquetats). La mostra de partida (uns 10 mg de proteïna) va ser concentrada fins a ~1 ml i va ser filtrada just abans de ser carregada en un loop de 2 ml. L'elució es va dur a terme amb un tampó 150 mM NaCl, 50 mM HEPES pH=7,4, 1 mM DTT a un flux constant d'1 ml/min a 4°C. Degut a les dimensions de la columna només es van començar a recollir fraccions (d'1 ml de volum) després de passats 60 ml d'elució.

A la fig. 4.16 es mostra un exemple de cromatograma i el seu corresponent gel de poliacrilamida-SDS d'una cromatografia d'exclusió molecular pel fragment Box AS. En el cromatograma observem un pic no simètric degut a l'elució del fragment Box AS juntament amb una certa proporció de la seva forma truncada al final del pic. La forma truncada presenta menor massa molecular i per tant queda més retinguda a la columna, tal i com queda reflectit al gel (on apareix a partir de la fracció 23).

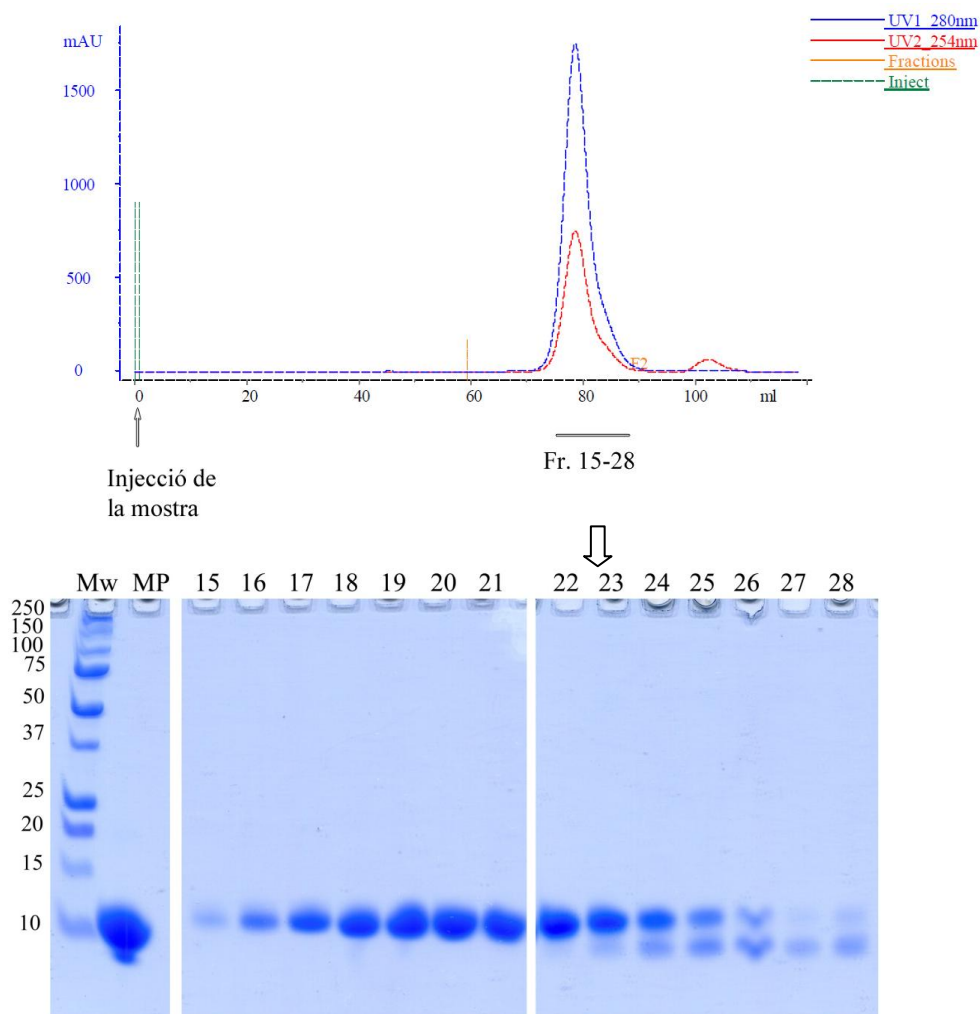


Fig. 4.16. Cromatograma i gel SDS-PAGE corresponents a la columna Superdex 75 per a l'HMGB1 Box AS.

Les fraccions de la 17 a la 22 d'aquesta columna HiLoad Superdex 75, un cop reunides i concentrades, s'han fet servir pels assajos cristal·logràfics. Amb els mateixos dispositius de concentració s'ha bescanviat el tampó per 50 mM NaCl, 25 mM HEPES pH=7,4, 1 mM DTT. A la fig. 4.17 es recullen els diferents estocs de proteïnes pures i concentrades obtinguts de la mateixa manera que s'acaba de detallar. A la taula 4.1 de l'apartat 4.1.3 s'especificuen les concentracions i volums d'aquests estocs concentrats.

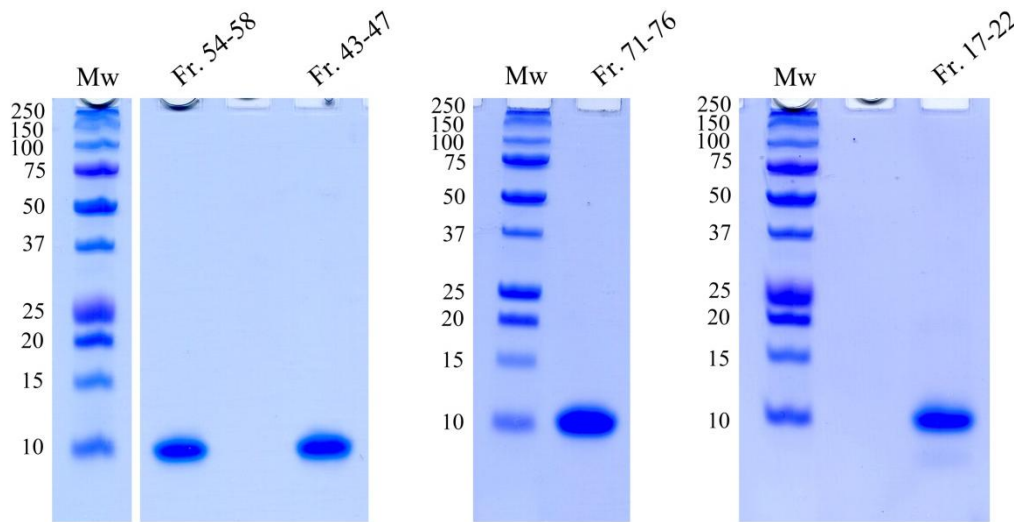


Fig. 4.17. Gel SDS-PAGE de les fraccions pures i concentrades obtingudes per a l'HMGB1 Box AS.

4.1.2.2. Expressió i purificació de la proteïna HMGB1 Box BS

El fragment Box BS s'ha expressat i purificat segons el protocol d'obtenció de les proteïnes HMGB (v. apartat 3.3) i aplicant les mateixes condicions descrites a l'apartat anterior per al fragment Box AS.

Expressió en E. coli

A la fig. 4.18 s'observa un gel d'SDS corresponent a les mostres de cultiu obtingudes durant una de les induccions de l'expressió de la proteïna de fusió GST-HMGB1 Box BS. Observem com augmenta la quantitat corresponent a la proteïna GST-HMGB1-Box BS, amb una massa molecular de ~34000 (~26000 Da corresponents a la GST més ~8400 Da al Box BS), a mesura que progressa el temps d'inducció.

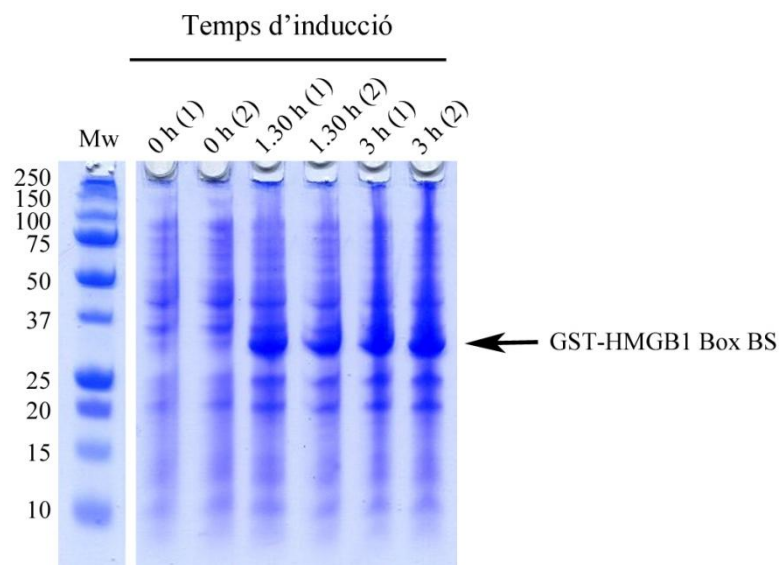


Fig. 4.18. Gel SDS-PAGE de les mostres analitzades durant la inducció per a la proteïna HMGB1 Box BS.

Els carrils s'han numerat com a la fig. 4.13.

Purificació per cromatografia

a) Cromatografia d'afinitat

Després de realitzar la lisi de les cèl·lules per sonicació, a partir de l'extracte clarificat (MP de la columna) s'ha realitzat el procediment de purificació per cromatografia d'afinitat. Al gel observem les proteïnes d'origen bacterià descartades en les primeres fraccions eluïdes de la columna de glutatió sefarosa (FT1, FT2 i W) i les fraccions que contenen la proteïna d'interès Box BS un cop tallada la cua GST (FTc, EIW1, EIW2 i EIW3) (fig. 4.19). Igualment, s'ha analitzat electroforèticament una petita mostra de la resina (*beads*) com a control.

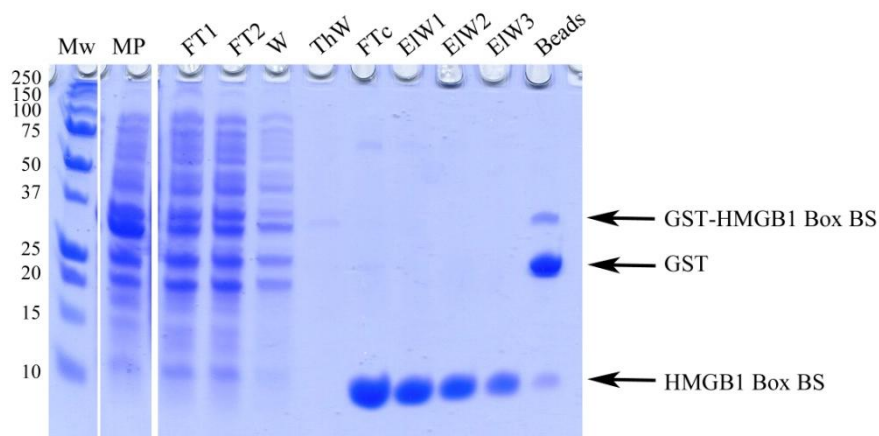


Fig. 4.19. Gel SDS-PAGE corresponent a la purificació GST per a l'HMGB1 Box BS.

Els carrils s'han numerat com a la fig. 4.14.

El fragment Box BS ha estat el que ha presentat una millor purificació de les quatre proteïnes HMGB amb les que hem treballat. Podem veure com obtenim unes fraccions amb unes bandes molt intenses corresponents a aquesta proteïna i no s'observa a penes cap altre tipus d'impuresa.

b) Cromatografia de bescanvi catiònic

De la mateixa manera que en el cas de la Box AS, les fraccions contenint el fragment Box BS (FTc, EIW1, EIW2, EIW3) es van carregar a una columna HiTrap. El cromatograma de la cromatografia de bescanvi catiònic mostra un pic corresponent a l'elució del fragment Box BS a l'inici del gradient (v. fig. 4.20).

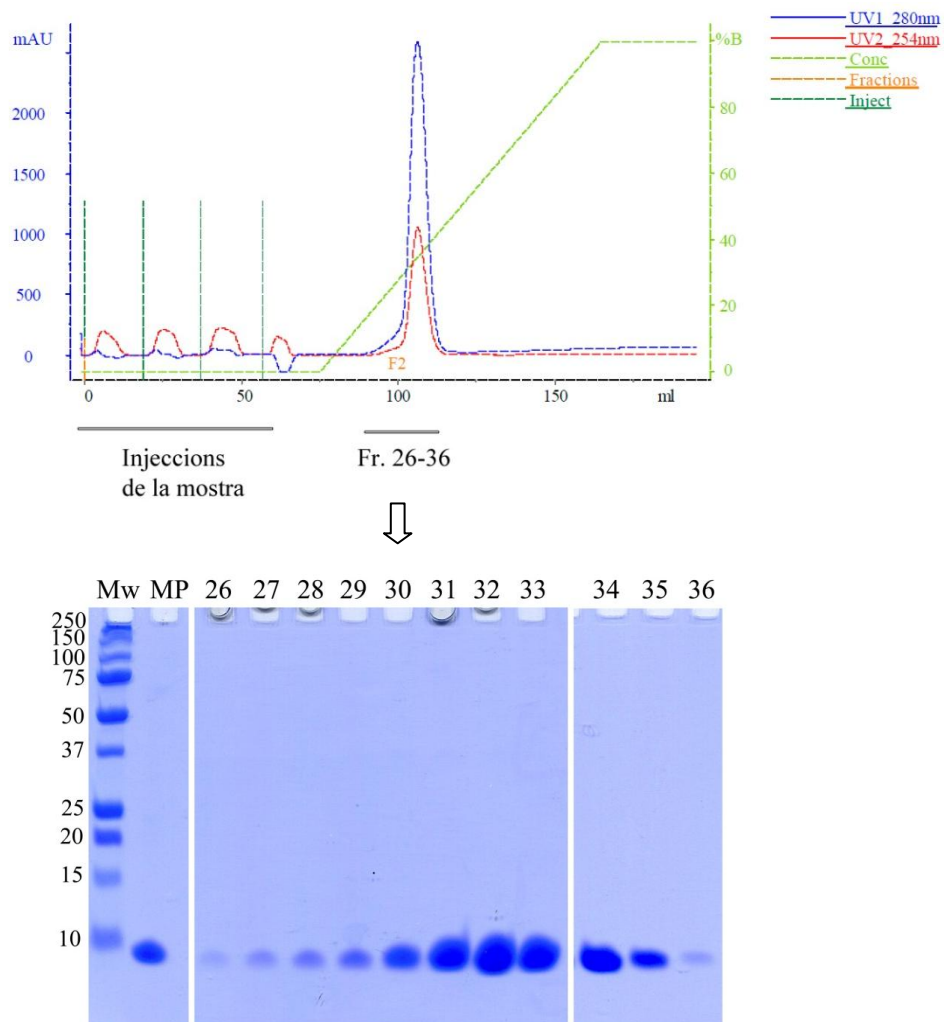


Fig. 4.20. Cromatograma i gel SDS-PAGE corresponents a la columna HiTrap SP FF per a l'HMGB1 Box BS. Mw, marcador de massa molecular; MP, mostra de partida carregada a la columna.

c) *Cromatografia d'exclusió molecular*

Tot i obtenir en aquest pas ja una mostra de proteïna Box BS molt pura de la fracció 26 a la 36, hem realitzat una columna d'exclusió molecular per assegurar-nos de tenir un estoc de proteïna pur i homogeni. A la fig. 4.21 es mostra el perfil d'absorbàncies per a la columna HiLoad Superdex 75 i el gel corresponent. En aquest cas es van ajuntar les fraccions de la 22 a la 30 (la zona central del pic) per a fer servir per a cristal·lografia.

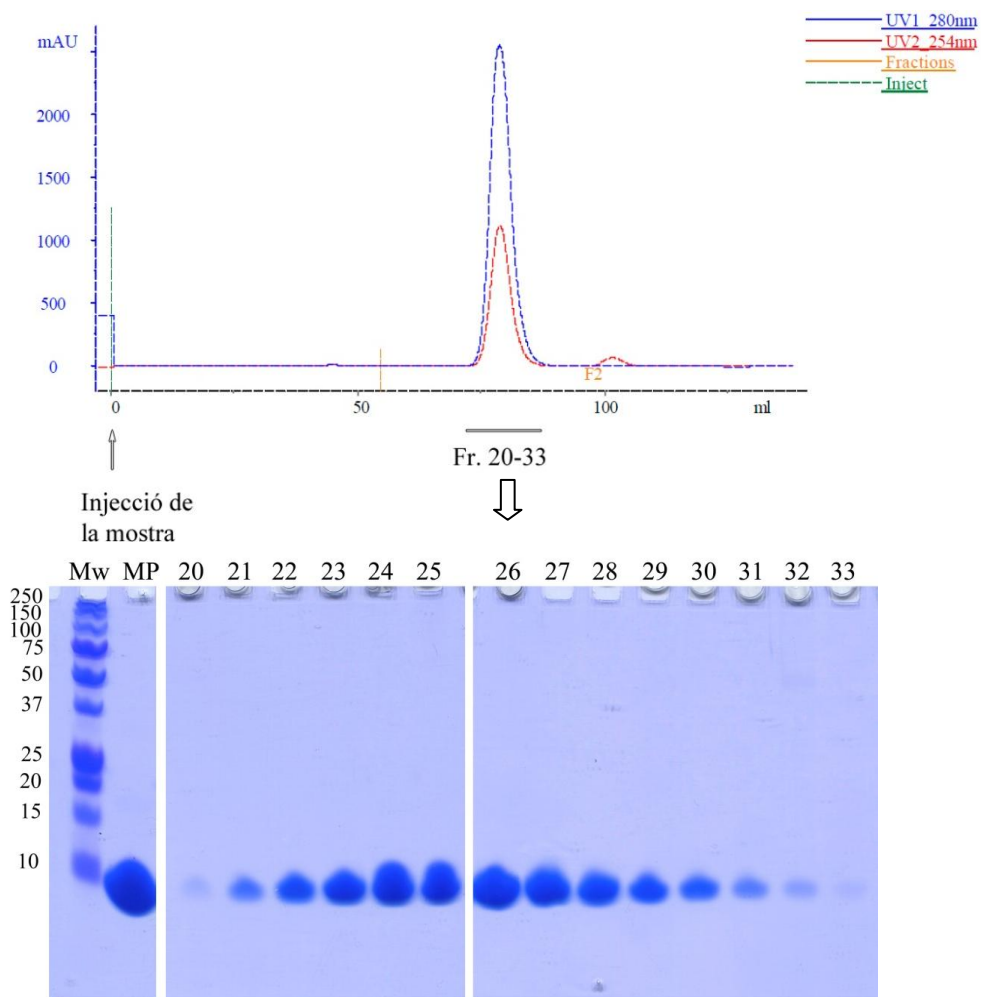


Fig. 4.21. Cromatograma i gel SDS-PAGE corresponents a la columna Superdex 75 per a l'HMGB1 Box BS.

Es van obtenir dos estocs de proteïnes molt purs de dues expressions diferents (de 2 litres de cultiu cadascuna). A la fig. 4.22 es mostren les fraccions analitzades electroforèticament un cop ajuntades, dessalades i concentrades.

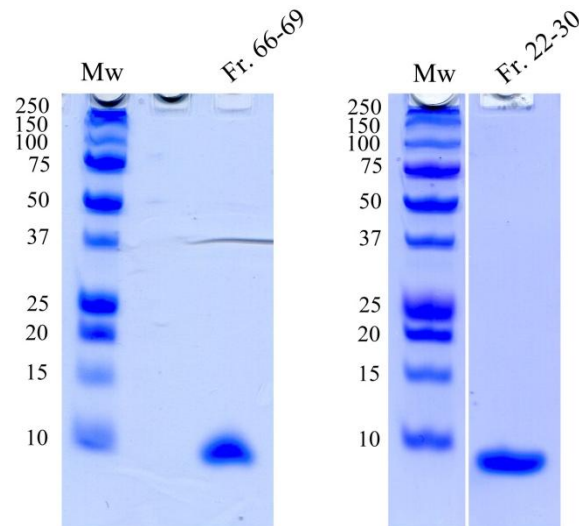


Fig. 4.22. Gel SDS-PAGE de les fraccions pures i concentrades obtingudes per a l'HMGB1 Box BS.

4.1.2.3. Expressió i purificació de la proteïna HMGB1 Box ABS

S'ha obtingut i purificat el fragment Box ABS seguint els passos descrits a l'apartat 3.3 i amb les mateixes condicions que s'han descrit per al fragment Box AS (v. apartat 4.1.2.1).

Expressió en E. coli

A la fig. 4.23 s'observa l'increment de la concentració del fragment Box ABS durant una expressió per a tres temps diferents d'inducció. En aquest cas, la banda de la proteïna GST-HMGB1-Box ABS presenta una massa molecular de ~44500 (~26000 Da corresponents a la GST i ~18500 Da al fragment Box ABS).

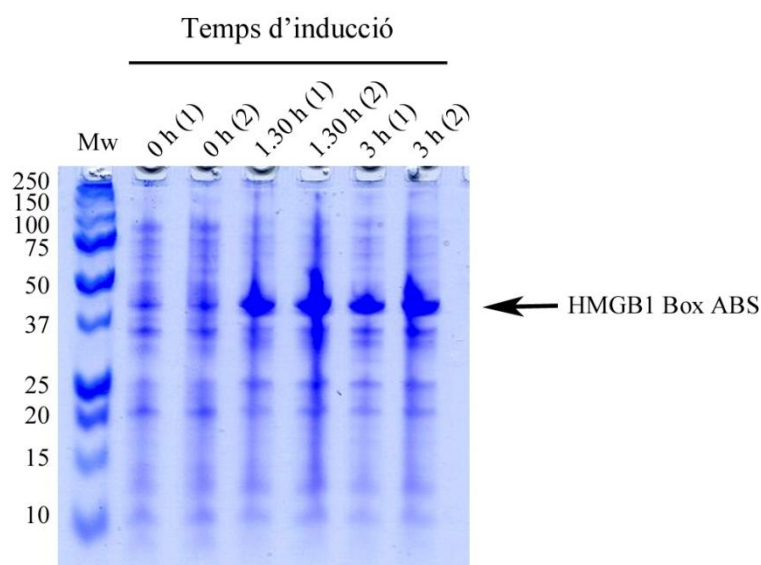


Fig. 4.23. Gel SDS-PAGE de les mostres analitzades durant la inducció de la proteïna HMGB1 Box ABS.

Els carrils s'han numerat com a la fig. 4.13.

Purificació per cromatografia

a) Cromatografia d'afinitat

A la fig. 4.24 es visualitzen les mostres analitzades electroforèticament obtingudes en els diferents passos de la cromatografia d'afinitat. La proteïna de fusió GST-HMGB1 Box ABS, de ~44500 Da de massa molecular, correspon a la banda més intensa de la mostra de partida (MP) carregada a la columna. La proteïna Box ABS (~18500 Da) sense la cua GST és eluïda de la columna en les fraccions FTc, EIW1, EIW2 i EI3 un cop la trombina ha realitzat l'escissió de la proteïna GST. En aquest cas es podria haver millorat el rendiment d'obtenció d'aquesta proteïna ja que podem observar restes de proteïna Box ABS en la resina un cop acabada la purificació (*beads*, al gel).

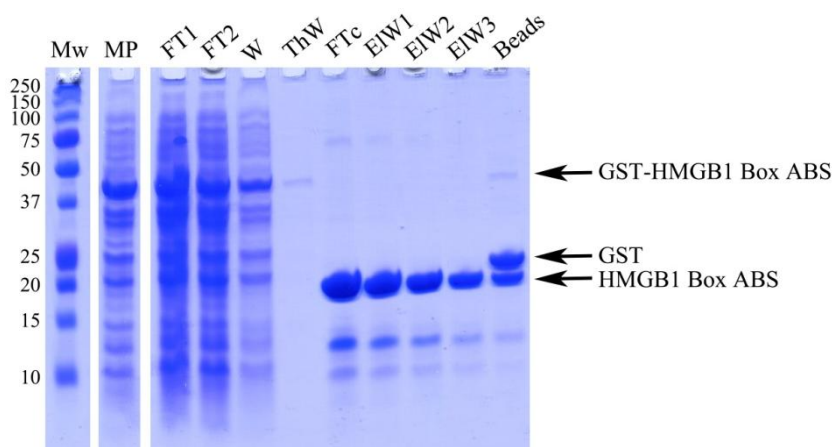


Fig. 4.24. Gel SDS-PAGE corresponent a la purificació GST per a l'HMGB1 Box ABS.

Els carrils s'han numerat com a la fig. 4.14.

b) Cromatografia de bescanvi catiònic

Anàlogament als casos anteriors, les fraccions amb proteïna Box ABS van ser reunides i repurificades primer per cromatografia de bescanvi catiònic i després per cromatografia d'exclusió molecular. A la fig. 4.25 s'observa un exemple de cromatografia de bescanvi catiònic i a la fig. 4.26 una cromatografia d'exclusió molecular per a la proteïna HMGB1 Box ABS.

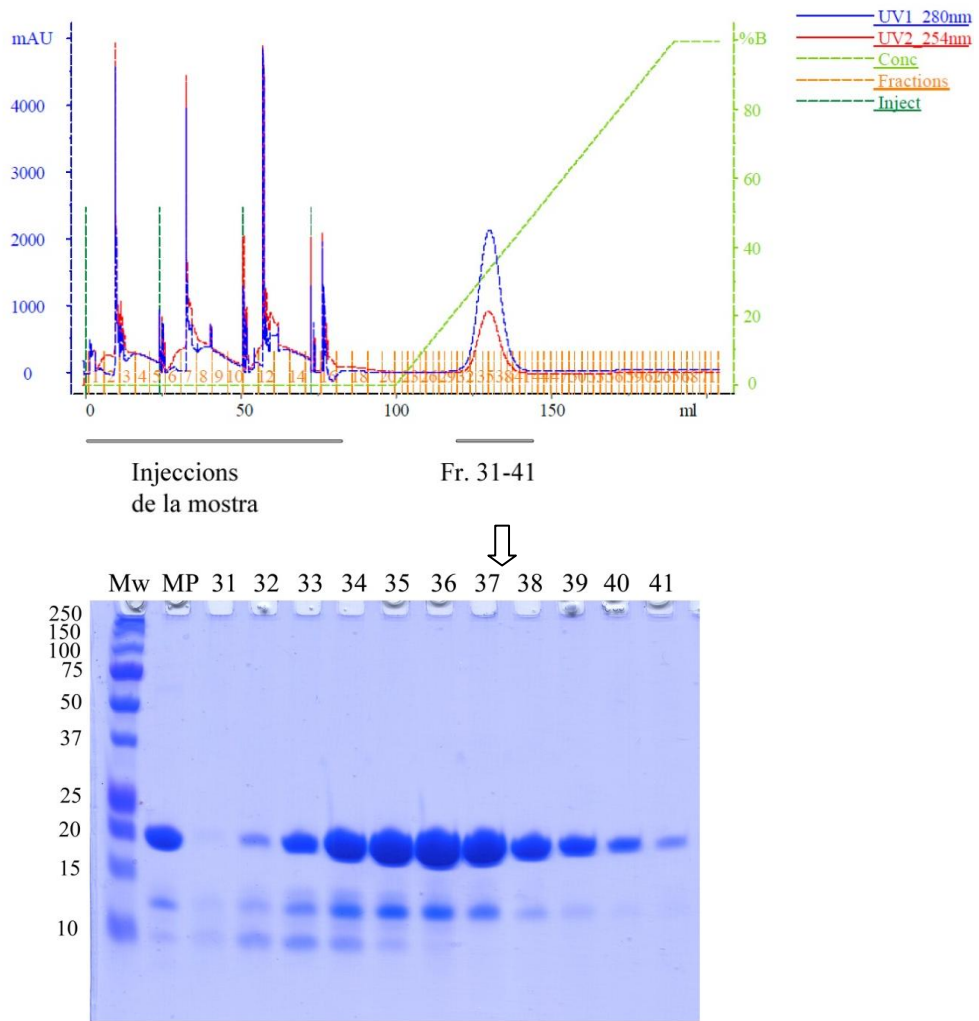
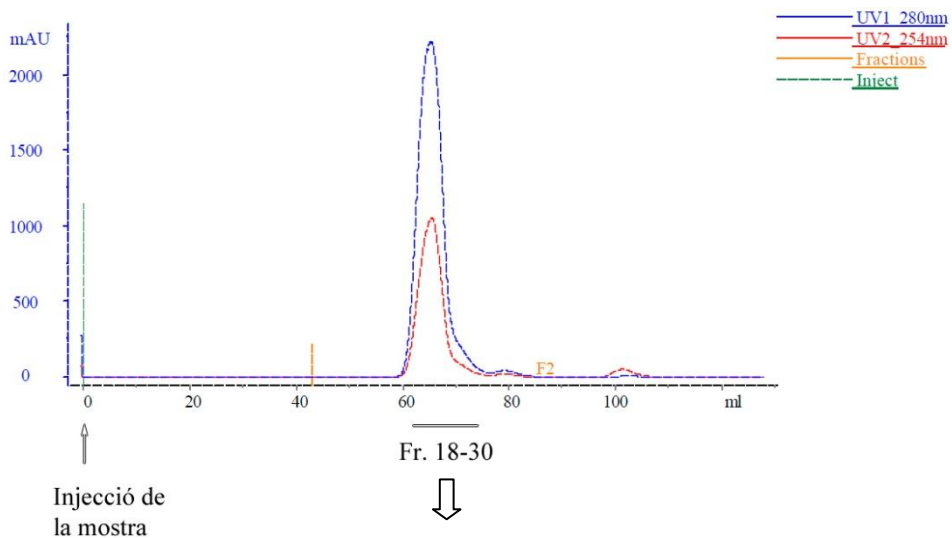


Fig. 4.25. Cromatograma i gel SDS-PAGE corresponents a la columna HiTrap SP per a l'HMGB1 Box ABS.
 Mw, marcador de massa molecular; MP, mostra de partida carregada a la columna.

c) Cromatografia d'exclusió molecular



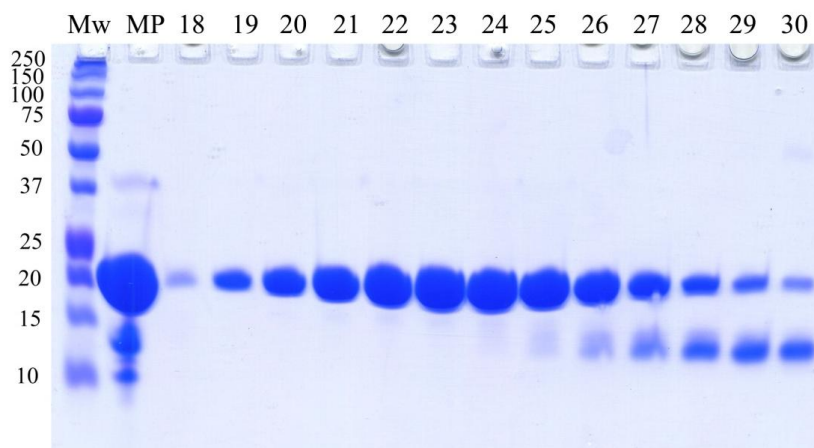


Fig. 4.26. Cromatograma i gel SDS-PAGE corresponents a la columna Superdex 75 per a l'HMGB1 Box ABS.

Com observem a la figura anterior es van poder obtenir fraccions del fragment Box ABS molt purs en les primeres fraccions eluïdes de la columna Superdex 75. En aquest cas en concret, es van ajuntar i concentrar les fraccions de la 19 a la 23.

De les diferents expressions i les seves corresponents cromatografies d'afinitat, de bescanvi catiònic i d'exclusió molecular hem obtingut tres estocs de fraccions pures un cop reunides i concentrades (v. fig. 4.27 i taula 4.1).

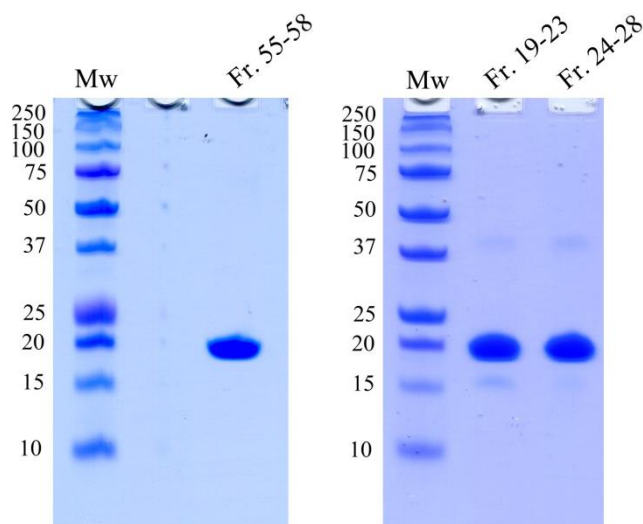


Fig. 4.27. Gel SDS-PAGE de les fraccions pures i concentrades obtingudes per a l'HMGB1 Box ABS.

4.1.2.4. Expressió i purificació de la proteïna NHP6A

Expressió en E. coli

Per a obtenir mostres de proteïna NHP6A pures, en primer lloc aquesta va ser expressada segons el protocol descrit a l'apartat 3.4.1. Cal destacar que, a diferència del casos anteriors, la proteïna NHP6A (proteïna HMGB de llevat) es va expressar en la seva forma nativa (~10800 Da), és a dir, sense fusionar amb una cua GST (v. fig. 4.28).

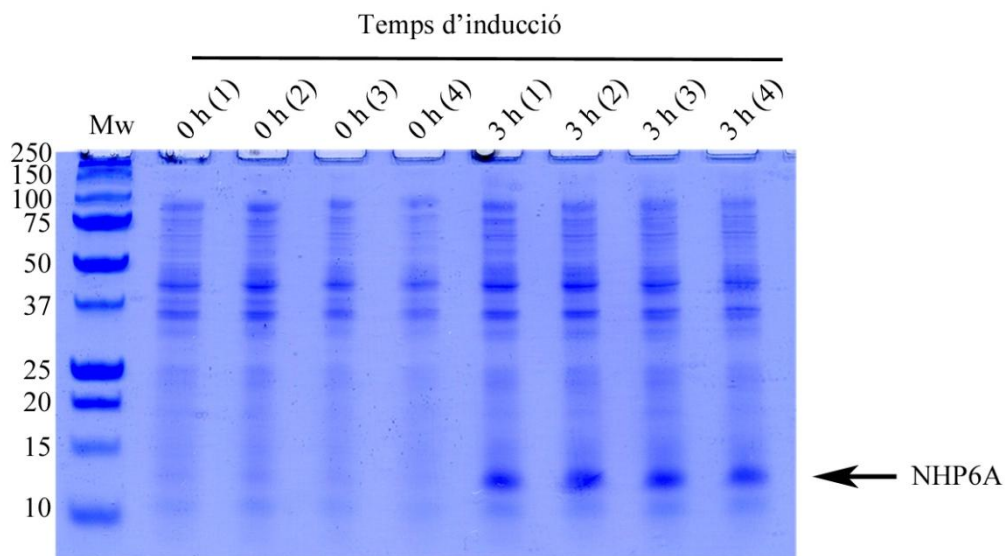


Fig. 4.28. Gel SDS-PAGE de les mostres analitzades durant la inducció de la proteïna NHP6A.

Els carrils s'han numerat com a la fig. 4.13.

La mostra resuspesa del sediment bacterià obtingut al final de l'expressió va ser tractada per sonicació i a continuació va ser centrifugada. El sobrenadant resultant (o extracte clarificat) va ser sotmès a una precipitació fraccionada amb sulfat d'amoni (v. 3.4.3). D'aquesta manera es van obtenir tres precipitats a diferents graus de saturació de sulfat d'amoni (30, 60 i 100%) i l'últim sobrenadant. Aquestes mostres van ser analitzades electroforèticament per tal de visualitzar-ne el contingut en la proteïna d'interès (fig. 4.29). Podem veure en el gel que la proteïna NHP6A precipita a 60% i 100% de saturació en sulfat d'amoni. Així aquestes mostres van ser dialitzades per a continuar purificant mentre que la mostra a 30% de saturació i el sobrenadant final van ser descartats.

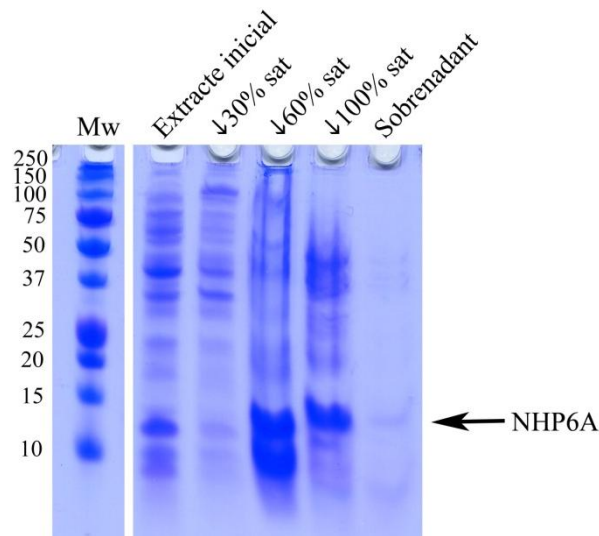


Fig. 4.29. Gel SDS-PAGE corresponent a la precipitació amb sulfat d'amoni de l'NHP6A.

Purificació per cromatografia

a) Cromatografia de bescanvi aniònic

Un cop dialitzades, les mostres corresponents als precipitats al 60% i al 100% de saturació en $(\text{NH}_4)_2\text{SO}_4$ (uns 12 ml) van ser carregades en una columna de bescanvi aniònic, concretament en una columna DEAE. En aquest tipus de columna, amb una resina contenint grups carregats positivament, seran eluides en el FT (*flow through*) aquelles proteïnes bàsiques que no queden retengudes a la columna, com és el cas de la proteïna NHP6A. En canvi, les proteïnes carregades negativament i els àcids nucleics, retinguts inicialment en la columna, s'eluiran a l'aplicar una major concentració de sals. Al gel de la fig. 4.30 observem com les fraccions amb mostra no retinguda a la columna presenten proteïna NHP6A.

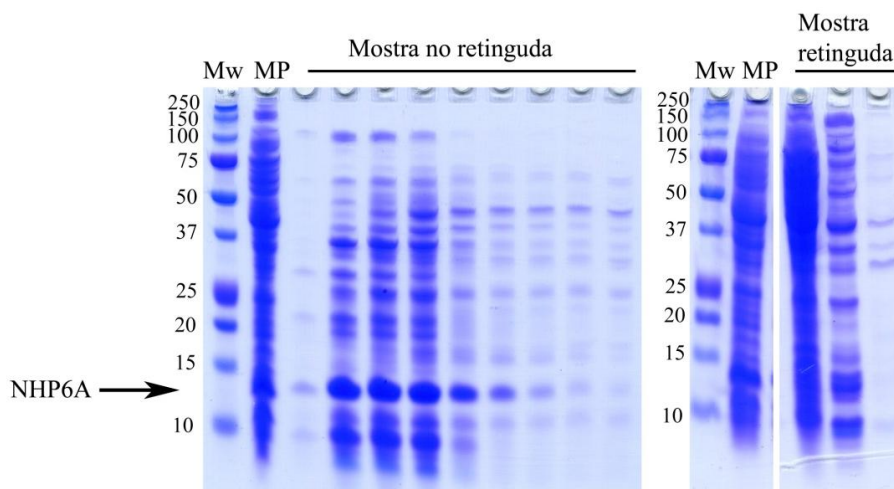


Fig. 4.30. Gel SDS-PAGE corresponent a la columna DEAE per a l'NHP6A. Mw, marcador de massa molecular; MP, mostra de partida carregada a la columna.

b) Cromatografia de bescanvi catiònic

Les fraccions corresponents al FT (*flow through*) de la columna DEAE van ser ajuntades i concentrades i es van carregar en una columna HiTrap SP FF de bescanvi catiònic (v. fig. 4.31).

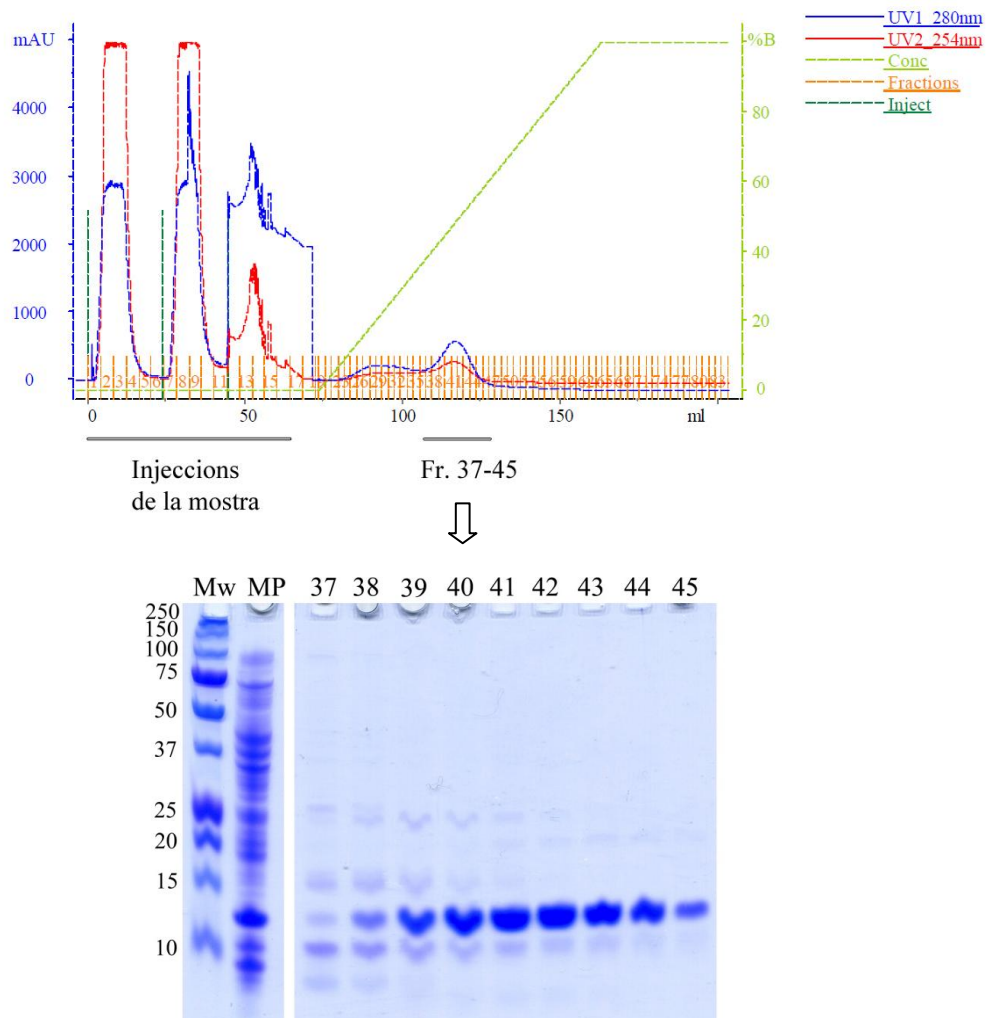


Fig. 4.31. Cromatograma i gel SDS-PAGE corresponents a la columna HiTrap SP FF per a l'NHP6A.

Mw, marcador de massa molecular; MP, mostra de partida carregada a la columna.

Al cromatograma es poden observar uns pics força grans corresponents a les proteïnes eluïdes al carregar la MP a la columna i que no presenten afinitat per la resina. Al gel de la fig. 4.31 observem les fraccions del cromatograma corresponents al pic principal del gradient i que contenen proteïna NHP6A. Tot i no estar del tot pures podem veure la gran eficiència d'aquesta columna en eliminar gran part de les proteïnes d'origen bacterià.

c) *Cromatografia d'exclusió molecular*

Finalment es va realitzar una cromatografia d'exclusió molecular per tal d'eliminar les proteïnes contaminants, visualitzades com bandes menys intenses per sobre (major massa molecular) i per sota (menor massa molecular) de la banda corresponent a l'NHP6A. A la fig. 4.32 podem veure com el pic més pronunciat correspon a la proteïna NHP6A, tot i que les primeres fraccions d'aquest pic contenen també impureses de major massa molecular. El pic més petit que s'observa a continuació correspon a les impureses amb menor massa molecular. En aquesta ocasió es van concentrar les fraccions de la 33 a la 38 (mostra més pura) i les fraccions 31 i 32 que presenten unes petites impureses.

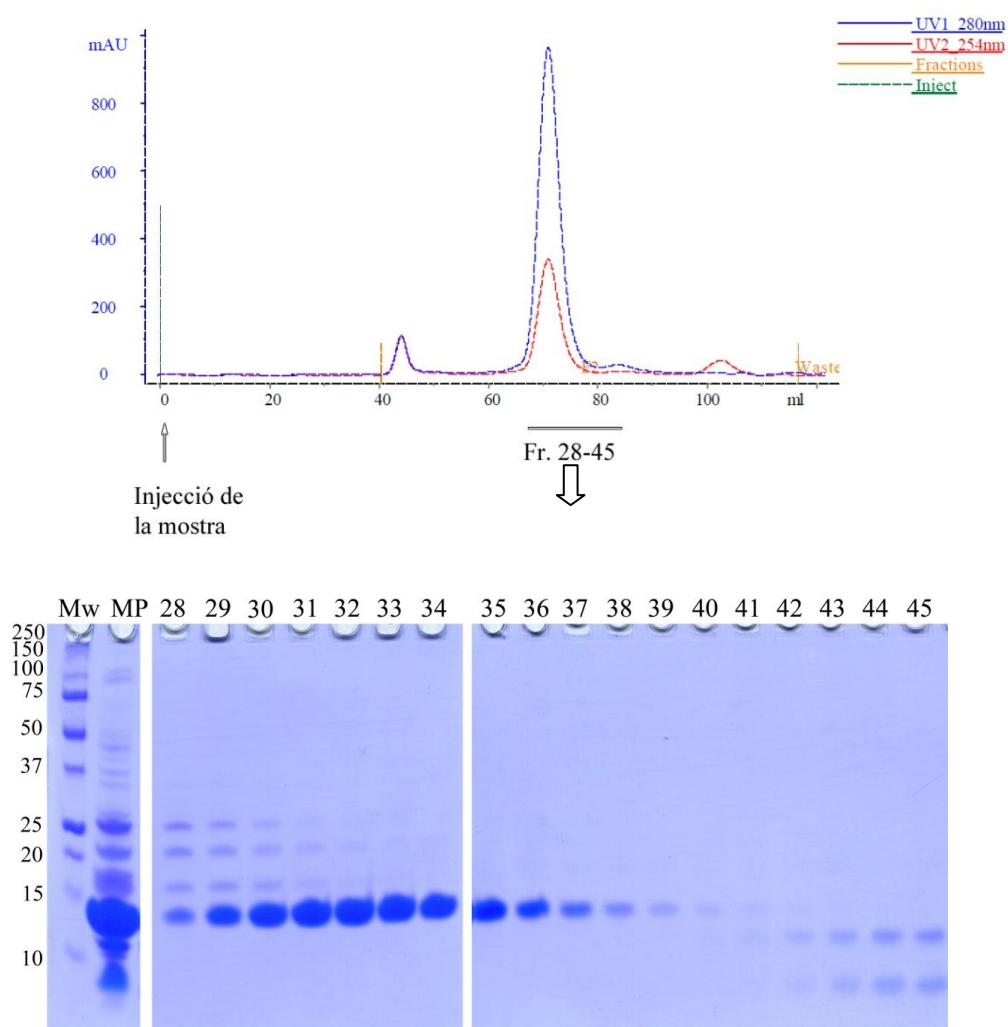


Fig. 4.32. Cromatograma i gel SDS-PAGE corresponents a la columna Superdex 75 per a l'NHP6A.

Mw, marcador de massa molecular; MP, mostra de partida carregada a la columna.

A la fig. 4.33 observem un gel de poliacrilamida-SDS on s'observen les fraccions finals concentrades 33-38 i 31-32. Les concentracions i volums d'aquests estocs concentrats s'especifiquen a la taula 4.1 de l'apartat 4.1.3.

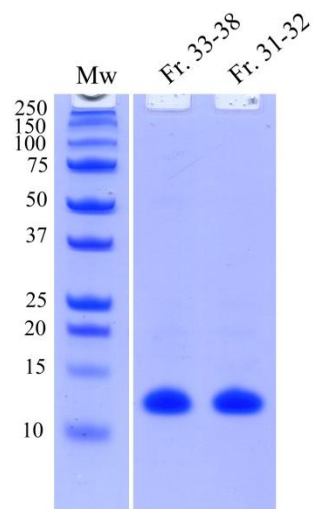


Fig. 4.33. Gel SDS-PAGE de les fraccions pures i concentrades obtingudes de l'NHP6A.

4.1.3. Resum de les proteïnes obtingudes per als assajos cristal·logràfics

Un cop les proteïnes HMGA i HMGB van ser expressades i purificades (v. apartats 4.1 i 4.1.2, respectivament) les mostres que presentaven un major grau de puresa van ser dialitzades i concentrades (v. apartat 3.5) per a fer servir en els assajos de cristal·lografia. A partir del valor d'absorbància de la mostra concentrada es va determinar la seva concentració. A la taula següent es resumeixen les concentracions i els volums que es van obtenir per a les diferents proteïnes purificades al llarg d'aquest treball.

Taula 4.1. Fraccions pures i concentrades de les proteïnes HMGA i HMGB pels assajos cristal·logràfics.

Nom proteïna	Nom fracció	Concentració (mM)	Concentració ($\mu\text{g}/\mu\text{l}$)	Volum (μl)	Tampó
HMGA1a*	Fr. 45+46+47	2,89	33,7	80	10 mM Tris pH= 7,5
HMGA1b(1-79)	Fr. 89-97	2,16	18,8	80	50 mM NaCl 10 mM Tris pH= 7,5
	Fr. 22-23	1,77	15,4	100	
	Fr. 24-25	2,87	24,9	100	
HMGB1 Box AS	Fr. 71-76	7,40	66,6	15	50 mM NaCl 25 mM HEPES pH= 7,4 1 mM DTT
	Fr. 54-58	4,25	38,8	20	
	Fr. 43-47	3,40	31,0	20	
	Fr. 17-22	5,99	53,5	65	
HMGB1 Box BS	Fr. 66-69	10,2	86,0	15	50 mM NaCl 25 mM HEPES pH= 7,4 1 mM DTT
	Fr. 22-30	9,48	80	100	
HMGB1 Box ABS	Fr. 55-58	1,2	21,8	20	50 mM NaCl 25 mM HEPES pH= 7,4 1 mM DTT
	Fr. 24-28	3,08	56,9	45	
	Fr. 19-23	5,24	97,0	65	
NHP6A	Fr. 31-32	4,39	47,4	24	50 mM NaCl 25 mM HEPES pH= 7,4 1 mM DTT
	Fr. 33-38	5,75	62,1	30	
	Fr. 21-24+24	3,26	35,2	30	

*Aquesta mostra correspon a la forma truncada de la proteïna HMGA1a.

Totes les formes de les proteïnes HMGB van ser purificades a la *University of Colorado* excepte l'estoc Fr. 21-24+24 que va acabar de ser purificat al laboratori de Barcelona.

4.2. Assajos cristal·logràfics

Un cop hem obtingut les diferents proteïnes HMG i els seus fragments prou purs i concentrats, el següent pas és la seva cristal·lització amb oligonucleòtids de seqüències riques en ATs (v. apartat 3.8). L'obtenció de cristalls corresponents a aquests complexos i la seva posterior difracció de raigs X ens pot permetre conèixer la seva estructura atòmica (v. apartat 3.9).

D'aquesta manera podrem estudiar les interaccions presents entre els *AT-hooks* de la proteïna HMGA1 i establir les condicions que governen en l'especificitat que els *AT-hooks* tenen per les zones riques en A i T. Anàlogament, amb la cristal·lització dels fragments de la proteïna HMGB1 i la proteïna NHP6A amb oligonucleòtids, podrem tenir un major coneixement de la seva forma d'interaccionar amb el DNA i quins canvis poden provocar en aquest. A més a més l'estudi de les interaccions de les *HMG-boxes* amb el DNA ens pot aportar una informació molt valuosa sobre la diferència de reconeixement d'unió al DNA que presenten les diferents proteïnes que contenen aquest domini.

4.2.1. Complexos HMGA-DNA i pèptids *AT-hook*-DNA

Es van realitzar diferents proves cristal·logràfiques amb les proteïnes HMGA (HMGA1a truncada i HMGA1b(1-79)) i tres pèptids corresponents als *AT-hooks* d'aquestes. La cristal·lització d'aquests mateixos oligonucleòtids amb pèptids corresponents als dominis *AT-hook* és també molt important, ja que al ser seqüències molt més curtes podrien ser més fàcilment cristal·litzables i ens aportarien també informació de com interaccionen específicament amb zones riques en ATs. S'ha treballat amb els pèptids sintètics corresponents al segon i tercer *AT-hook* de la proteïna HMGA1a (KRPRGRP i RKPRGRPCK) així com també amb un pèptid corresponent a la regió comú d'aquests dos dominis (PRGRP) (v. fig. 4.34).

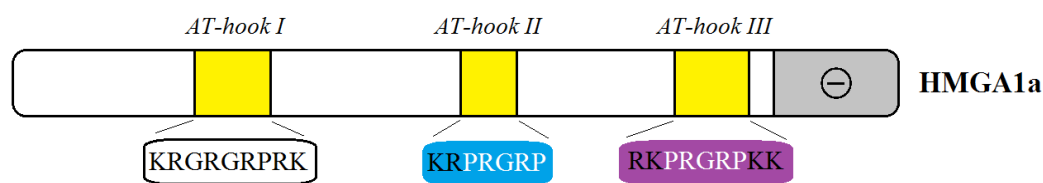


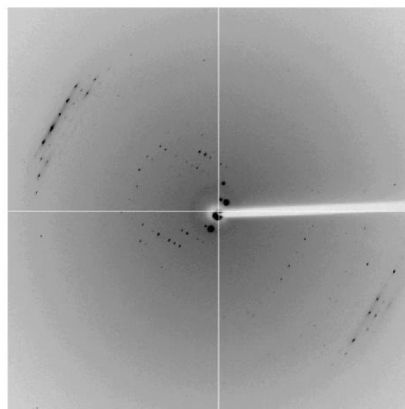
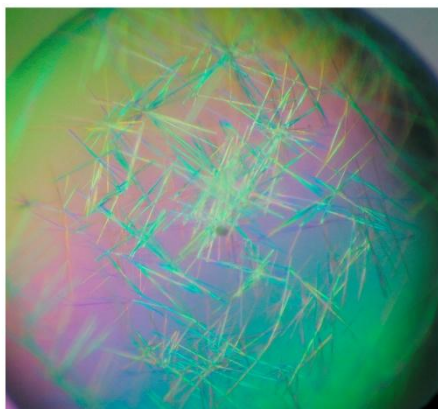
Fig. 4.34. Seqüència aminoacídica dels *AT-hooks* de la proteïna HMGA1a.

Els tres dominis *AT-hook* apareixen en groc i la regió carboxi-terminal en gris. Ombrejat en blau es troba el pèptid KRPRGRP corresponent a l'*AT-hook* II i ombrejat en violeta el pèptid RKPRGRPCK corresponent a l'*AT-hook* III. En lletres blanques es destaca la regió comú d'aquests dos *AT-hooks* (PRGRP).

Es van realitzar dues línies d'experiments. Per una banda es van voler optimitzar els resultats obtinguts en proves de cristal·lització prèvies per a la proteïna HMGA1a truncada i el pèptid PRGRP realitzades durant el PFC [128]. En tots dos casos es van obtenir uns cristalls que difractaven força bé fins a una resolució d'aproximadament 3,2 Å però la resolució de la seva estructura només va permetre la localització dels oligonucleòtids i no de cap fragment de la seqüència aminoacídica corresponent a la proteïna o al pèptid (v. fig. 4.35).

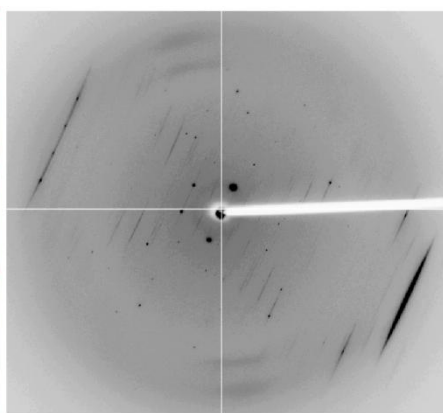
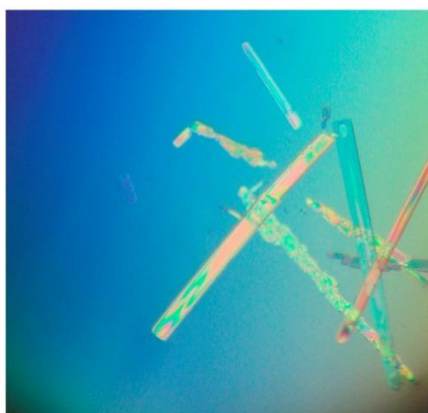
D'altra banda es van iniciar proves de cristal·lització amb dos pèptids més i amb el fragment HMGA1b(1-79). Així es va cristal·litzar l'oligonucleòtid CGCATTAAATGCG amb els pèptids KRPRGRP i RKPRGRPCK, amb unes seqüències de dos i quatre aminoàcids més llargues que el primer pèptid (v. fig. 4.34) i la proteïna HMGA1b(1-79) amb els oligonucleòtids AATAATTATT i GGGAAATTTCCC. És per aquest motiu que en el primer grup de proves es van realitzar unes gotes amb unes condicions més concretes mentre que en el segon es va fer un escombrat de condicions més ampli provant diferents tampons de cristal·lització.

a



Gota 11
0,6 mM HMG1a truncada
0,3 mM CGCATTAAATGCG
20 mM MgCl ₂ , 0,25 mM NiCl ₂
25 mM NaCac pH=6
2,5% MPD*
Pou: 33% MPD
Dades del cristall
P32
a=45,15; b=45,16; c=98,41 Å
α=90°; β=90°; γ=120°

b



Gota 47
1,3 mM PRGRP
0,3 mM AATAATTATT
20 mM MgCl ₂ , 0,5 mM NiCl ₂ , 1 mM NaCl
25 mM NaCac pH=6
2,5% MPD*
Pou: 31% MPD
Dades del cristall
C2
a=45,15; b=45,16; c=98,41 Å
α=90°; β=90°; γ=120°

Fig. 4.35. Imatges, condicions de cristal·lització i diagrames de difracció dels millors resultats obtinguts en assajos previs [128]. a) HMG1a truncada + CGCATTAAATGCG. b) Pèptid PRGRP + AATAATTATT.

*Corresponent al tampó Matrix n°15 (veure annex D). Cristalls difractats al sincrotró de Grenoble (línia BM-16): a) angle d'oscil·lació 1°, temps d'exposició 15 s, distància al detector 280 mm i longitud d'ona 0,9783 Å. b) Igual excepte, angle d'oscil·lació 3° i temps d'exposició 10 s.

La formació dels complexos entre les proteïnes HMGA i els oligonucleòtids fets servir en els assajos cristal·logràfics d'aquest treball s'ha visualitzat mitjançant anàlisis electroforètiques en gels EMSA (v. 3.10.1). A la fig. 4.36 observem els gels EMSA corresponents a la proteïna HMG1b(1-79) amb els oligonucleòtids CGCATTAAATGCG, AATAATTATT i GGGAAATTTCCC. Podem veure com en els tres casos es visualitza un retard en la migració de les bandes de DNA degut a la formació del complex a l'incrementar la concentració de proteïna.

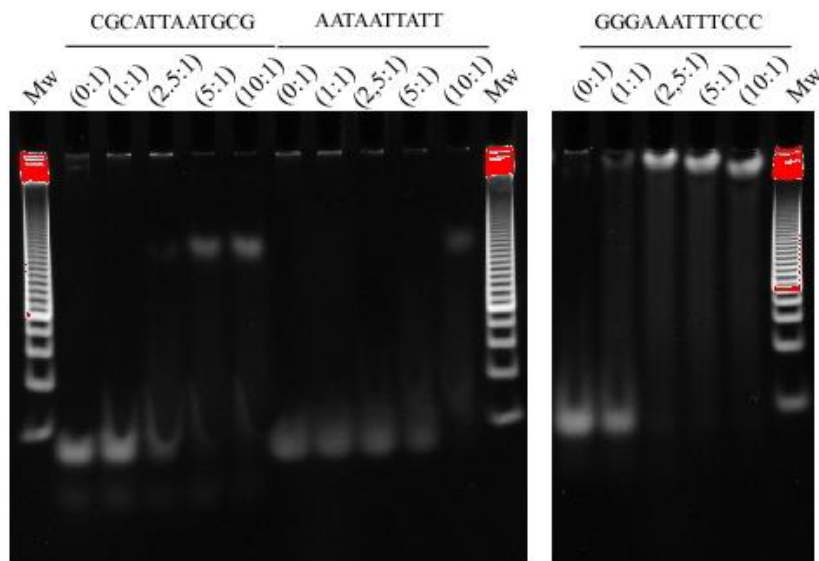


Fig. 4.36. Assaigs EMSA per a la formació de complexos entre l'HMGA1b(1-79) i diferents oligonucleòtids. S'indica la ràtio molar (Proteïna:DNA) per a cada mostra carregada al gel. El marcador de massa molecular (Mw) presenta bandes de 20 pb a 500 pb en increments de 20 pb.

Resum de les condicions de cristal·lització

El conjunt dels assajos cristal·logràfics de les proteïnes HMGA i els pèptids *AT-hook* amb oligonucleòtids rics en ATs es resumeixen a la taula 4.2. S'han utilitzat tant tampons comercials de la casa Hampton Research (Natrix HR2-116) (v. annex D) com variacions dels tampons que han donat millors resultats (principalment amb sals de $MgCl_2$). La majoria dels tampons presenten concentracions baixes de sals, per tal d'evitar la competència d'aquestes amb els grups carregats positivament de la proteïna en la unió a l'oligonucleòtid. Hem triat aquells tampons amb uns pHs d'entre 6 i 7,5 per a treballar en unes condicions pròximes a les fisiològiques. Sempre s'ha intentat assajar diferents relacions molars entre la part proteica i el DNA per a variar també un altre factor de cristal·lització. La temperatura de treball ha estat de 4°C en tots els casos. A continuació es presenten aquests resultats amb més detall.

Taula 4.2. Resum dels assajos cristal·logràfics per a les proteïnes HMGA i els pèptids *AT-hook* amb DNA.

Proteïna o pèptid	Oligonucleòtid	Ràtio P:D (mM)	Condicions			n° gotes	Cristalls	Difracció
			Sal	pH	Precip			
HMGA1a truncada	CGCATTAAATGCG	0,6:0,3 0,6:0,6 0,3:0,6	Natrix (n° 15)			4	Agulles fines i llargues	Poques taques, resol. 3Å
		0,6:0,3	MgCl ₂ (+NiCl ₂)	6	MPD	4	Agulles fines i llargues	-
HMGA1b(1-79)	AATAATTATT	0,35:1 0,3:0,1 0,4:0,2	Natrix (n° 7, 10, 15, 16, 22, 25, 31, 34, 35, 42, 43)			28	-	-
		0,35:1 0,3:0,1 0,4:0,2	MgCl ₂ ; CaCl ₂ ; NaCl; -	6; 6,5, 7; 7,5	MPD	26	Cristall líquid	-
	GGGAAATTTCCC	0,6:0,2 0,4:0,2	Natrix (n° 7, 10, 15, 16, 22, 25, 31, 34, 35, 42, 43)			14	-	-
		0,6:0,2 0,4:0,2	MgCl ₂ ; CaCl ₂ ; NaCl;-	6; 6,5, 7; 7,5	MPD	26	-	-
PRGRP (<i>AT-hook</i> II i III)	AATAATTATT	1,3:0,3	Natrix (n° 15 i 42)			3	Varetes fines i llargues	Poques taques
		1,3:0,3	MgCl ₂	6	MPD	3	Agulles grans	-
KRPRGRP (<i>AT-hook</i> II)	CGCATTAAATGCG	0,4:0,2 0,2:0,2 0,1:0,2	Natrix (n° 7, 10, 15, 22, 25, 31, 42, 43)			24	Agulles; Varetes	Difracció amb desordre
		0,4:0,2 0,2:0,2 0,1:0,2	MgCl ₂	6; 7,5	MPD	15	Agulles; Varetes	Difracció amb desordre
RKPRGRP (<i>AT-hook</i> III)	CGCATTAAATGCG	0,4:0,2 0,2:0,2 0,1:0,2	Natrix (n° 7, 10, 15, 22, 25, 31, 42, 43)			24	Agulles; Varetes	Difracció amb desordre
		0,4:0,2 0,2:0,2 0,1:0,2	MgCl ₂	6; 7,5	MPD	15	Agulles; Varetes	Difracció amb desordre

La ràtio P:D és la ràtio molar proteïna o pèptid:DNA. Tots els assajos s'han realitzat a 4°C. Precip., precipitant.

Resultats de cristal·lització

a) HMGA1-DNA

En tots els assajos cristal·logràfics realitzats per a la proteïna HMGA1a truncada i l'oligonucleòtid CGCATTAAATGCGC es van obtenir cristalls però no es va aconseguir millorar els resultats obtinguts en els assajos previs (v. fig. 4.35). En aquestes gotes s'han format principalment moltes agulles molt fines i llargues (v. fig. 4.37 esquerra). S'han congelat tres cristalls corresponents a la condició a 20 mM MgCl₂, 0,5 mM NiCl₂, 25 mM NaCac pH=6 amb MPD com a precipitant (Natrix n°15); dos d'ells a una relació molar proteïna:DNA d'1:2 i l'altre a 1:1. La difracció d'aquest últim cristall (R3-1) va ser la millor i es mostra a la fig. 4.37. Observem com el diagrama de difracció presenta molt poques taques ben definides. Aquests diagrames de difracció presenten una distribució de taques similar al cristall dels estudis previs

compatibles amb una cel·la unitària equivalent ($a=45 \text{ \AA}$; $b=45 \text{ \AA}$; $c=98 \text{ \AA}$; $\alpha=90^\circ$; $\beta=90^\circ$; $\gamma=120^\circ$) en el grup espacial P32.

Així doncs, hem pogut reproduir les condicions de cristal·lització del cristall previ obtenint cristalls amb la mateixa morfologia però no hem obtingut unes imatges de difracció de major qualitat per a poder analitzar la seva estructura tridimensional.

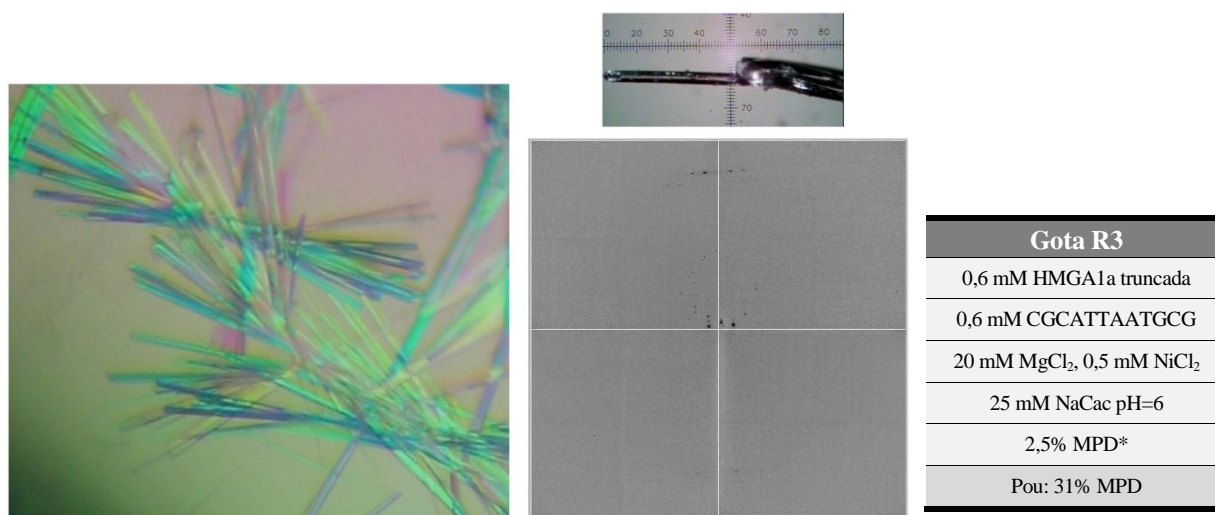


Fig. 4.37. Imatge, condicions de cristal·lització i diagrama de difracció d'un cristall corresponent a la proteïna HMGA1a truncada i l'oligonucleòtid CGCATTAAATGCG (gota R3, cristall R3-1).

*Corresponent al tampó Matrix n°15 (veure annex D). Cristall difractat al sincrotró de Grenoble (línia BM-16): angle d'oscil·lació 3° , temps d'exposició 10 s, distància al detector 275 mm i longitud d'ona $0,9795 \text{ \AA}$.

En la cristal·lització de la proteïna HMGA1b(1-79) amb els dos oligonucleòtids (AATAATTATT i GGGAAATTTCCC) s'han obtingut cristalls líquids i esferulites en les condicions assajades. Els cristalls líquids presenten pròpiament desordre i per tant no s'ha congelat cap d'ells. Tot i així s'ha pogut observar que, igualment que per a l'HMGA1a truncada, els millors resultats s'obtenen amb la sal MgCl₂.

b) Pèptids AT-hook II i AT-hook III-CGCATTAAATGCG

Tant en la cristal·lització de l'oligonucleòtid CGCATTAAATGCGC amb el pèptid corresponent a l'AT-hook II (KRPRGRP) com en el cas de l'AT-hook III (RKPRGRPCK), s'han format cristalls en totes les gotes realitzades (v. taula 4.2) a excepció de les que contenen el tampó Matrix número 25 en las que teníem precipitat amorf. En tots dos casos hem vist que els cristalls es formaven molt ràpidament (a baixa concentració de MPD o PEG).

S'han congelat un total de 29 cristalls. Aquests són principalment varetes que presenten puntes afilades (v. fig. 4.38 com exemple) o bé puntes rectes. Les gotes corresponents a aquests cristalls principalment contenen sals de $MgCl_2$ a diferents concentracions, es troben a pH 6 o 7,5 i tenen MPD com a precipitant. També s'han congelat dos cristalls d'una gota a 40 mM de $MgAc$ i pH=6.5 (un pel pèptid KRPRGRP i un altre pel RKPRGRP) i també un altre a 2,5 mM de $MgSO_4$ i pH=6 per aquest últim. S'ha vist com el precipitant MPD presenta millors resultats que el PEG.

Si ens fixem en la relació pèptid/oligonucleòtid no es veu una preferència clara per cap de les tres relacions i depèn de la condició de la gota. Si comparem la cristal·lització del pèptid *AT-hook II* amb la del pèptid *AT-hook III* amb l'oligonucleòtid CGCAT₂A₂TGCG hem observat que els cristalls formats amb el primer tendeixen a tenir millor forma (costats uniformes, major grandària i única direcció de creixement).

Els resultats de la difracció d'aquests cristalls no van ser els esperats. Tot i que aquests cristalls presentaven una bona forma i birrefringència no han difractat gaire bé (v. fig. 4.38 i fig. 4.39). Les reflexions dels diagrames de difracció no són punts, fet que ens fa pensar que el complex es troba desordenat dins del cristalls. A la fig. 4.38 observem un diagrama de difracció corresponent a un cristall en forma de vareta afilada del pèptid *AT-hook II* i a la fig. 4.39 un altre per un cristall de l'*AT-hook III*. Podem observar el desordre que presenten les molècules dins del cristall per l'obtenció de taques mal definides i en més d'una direcció.

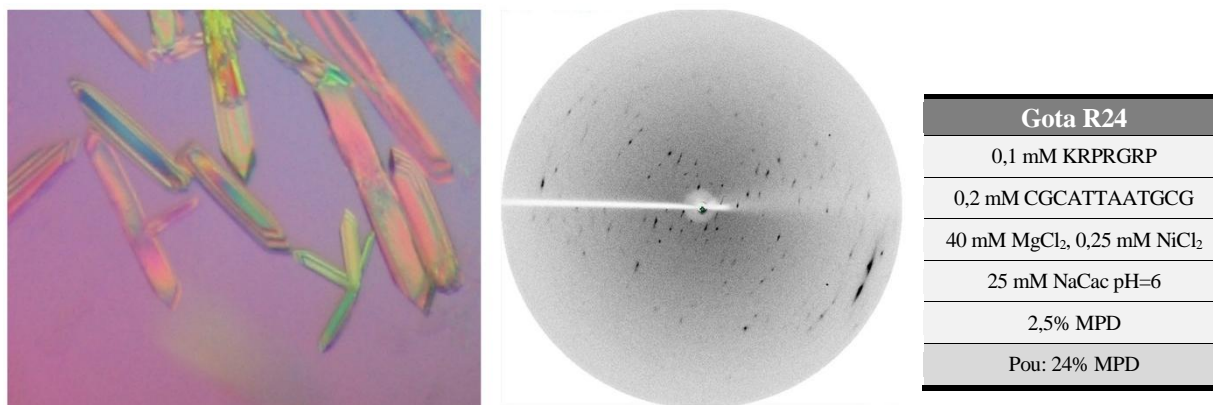


Fig. 4.38. Imatge, condicions de cristal·lització i diagrama de difracció d'un cristall corresponent al pèptid KRPRGRP (*AT-hook II*) i l'oligonucleòtid CGCATTAAATGCG (gota R24, cristall R24-1).

Cristall difractat a la Plataforma Automatitzada de Cristal·lografia del PCB: angle d'oscil·lació 2° , temps d'exposició 300 s, distància al detector 300 mm i longitud d'ona 1,542 Å.

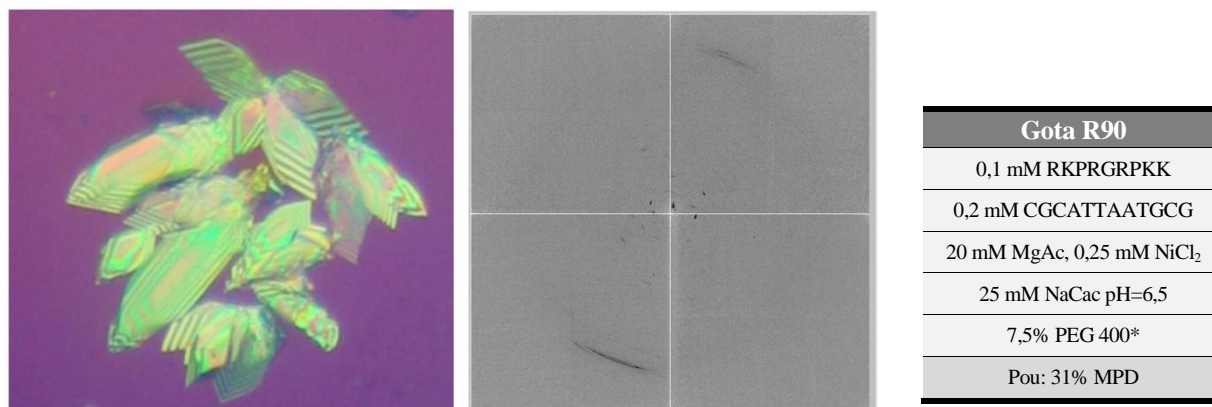


Fig. 4.39. Imatge, condicions de cristal·lització i diagrama de difracció d'un cristall corresponent al pèptid **RKPRGRPKK** (*AT-hook III*) i l'oligonucleòtid **CGCATTAAATGCG** (gota R90, cristall R90-1).

*Corresponent al tampó Natrix n°22 (veure annex D). Cristall difractat al sincrotró de Grenoble (línia BM-16): angle d'oscil·lació 5°, temps d'exposició 10 s, distància al detector 275 mm i longitud d'ona 0,9795 Å.

En un primer moment es va pensar que els cristalls presentaven desordre degut a què aquests havien crescut ràpidament i les molècules no s'havien ordenat dins del cristall. Però, al veure que tots els cristalls obtinguts amb els dos pèptids i tant els cristalls difractats a la plataforma cristal·logràfica (Parc Científic de Barcelona) com al Sincrotró de Grenoble presentaven el mateix problema es va sospitar de la integritat de mostra d'oligonucleòtid utilitzada. Efectivament, a l'analitzar en un gel tipus EMSA la dilució de la mostra de DNA feta servir pels assajos cristal·logràfics es va veure que presentava degradació (fig. 4.40a). Es pot observar com apareix una banda de la mateixa intensitat i per sota de la corresponent a l'oligonucleòtid original. A la fig. 4.40b podem observar la millor de les difraccions obtingudes per aquests cristalls i, tot i obtenir forces taques i fins a uns 3 Å, segueixen presentant desordre.

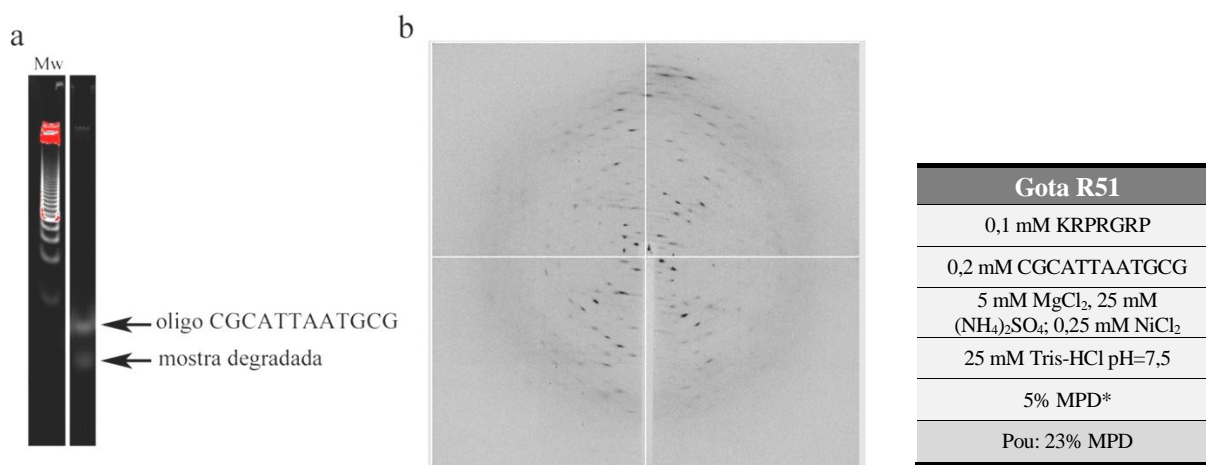


Fig. 4.40. a. Gel EMSA de l'oligonucleòtid **CGCATTAAATGCG**. b. Condicions de cristal·lització i diagrama de difracció d'un cristall corresponent al pèptid **KRPRGRP** (*AT-hook II*) i aquest oligonucleòtid (cristall R51-1).

*Corresponent al tampó Natrix n°43 (veure annex D). Cristall difractat al sincrotró de Grenoble (línia BM-16): angle d'oscil·lació 12°, temps d'exposició 12 s, distància al detector 275 mm i longitud d'ona 0,9795 Å.

El fet que, tot i estar degradat, s'han pogut obtenir múltiples cristalls amb aquest oligonucleòtid i els pèptids *AT-hook* II i III, fa molt recomanable repetir en futurs assajos condicions similars de cristal·lització però amb un nou estoc d'oligonucleòtid CGCATTAATGCG.

c) PRGRP-AATAATTATT

Hem obtingut cristalls en les 6 gotes realitzades. Principalment són varetes molt grans però de forma irregular. És possible que aquests cristalls s'hagin format força ràpidament i per això les estructures formades no posseeixen prou ordre. Només hem congelat una vareta fina de la condició que presentava menys sals (5 mM MgCl₂ amb una petita concentració de NiCl₂). En el seu diagrama de difracció observem molt poques taques ben definides a baixa resolució (v. fig. 4.41).

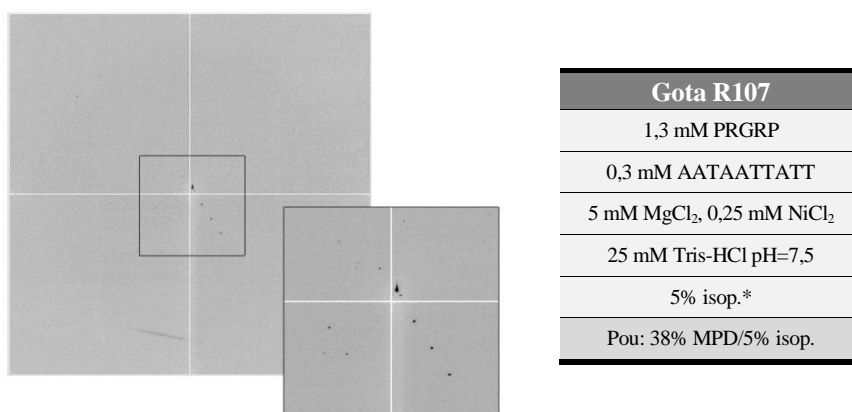


Fig. 4.41. Imatge, condicions de cristal·lització i diagrama de difracció d'un cristall corresponent al pèptid PRGRP i l'oligonucleòtid AATAATTATT (gota R107, cristall R107-1).

*Corresponent al tampó Natrx n°42 (veure annex D). Cristall difractat al sincrotró de Grenoble (línia BM-16): angle d'oscil·lació 5°, temps d'exposició 10 s, distància al detector 275 mm i longitud d'ona 0,9795 Å.

Alguns dels diagrames de difracció corresponents a les proves cristal·logràfiques pel pèptid PRGRP i l'oligonucleòtid AATAATTATT que havíem realitzat amb anterioritat [128] s'han tractat en el present estudi. Tot i que els diagrames de difracció obtinguts dels cristalls de la gota 45 (fig. 4.35) eren difícils de processar, un dels cristalls d'aquesta càpsula era clarament un cristall únic i no presentava una estructura tan complexa. Aquest cristall és el 45-2 i presenta el diagrama de difracció que es mostra a la fig. 4.42. Tot i que la resolució màxima d'aquest cristall no és molt alta (aproximadament 3.1 Å) i que tenim poques imatges de difracció (15 diagrames cada 12°) hem pogut obtenir un model d'estructura aproximat d'aquest cristall estudiant els diagrames de difracció (v. apartat 4.3).

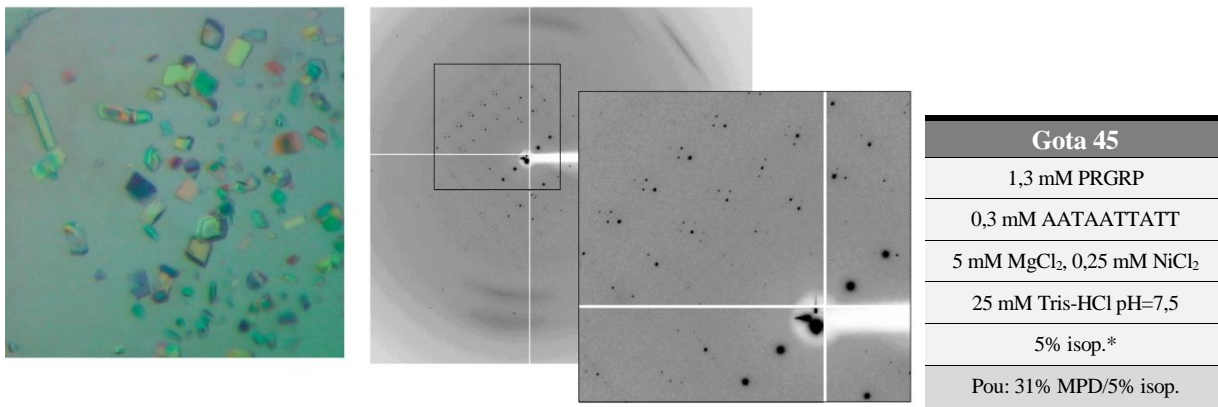


Fig. 4.42. Imatge, condicions de cristal·lització i diagrama de difracció d'un cristall corresponent al pèptid PRGRP i l'oligonucleòtid AATAATTATT (gota 45, cristall 45-2).

*Corresponent al tampó Natrix n°42 (veure annex D). Cristall difractat al sincrotró de Grenoble (línia BM-16): angle d'oscil·lació 12°, temps d'exposició 80 s, distància al detector 250 mm i longitud d'ona 0,9840 Å.

Electroforesi de cristalls

En alguns casos per tal de comprovar el contingut en proteïna o pèptid dels cristalls realitzats es van analitzar electroforèticament els cristalls obtinguts un cop difractats (v. apartat 3.6.2).

Per una banda es va comprovar el contingut proteic en dos cristalls corresponents a la proteïna HMGA1a truncada, el R3-1 i el R4-1. Totes dues gotes tenien les mateixes condicions de cristal·lització (20 mM MgCl₂, 25 mM NaCac pH=6 i 0,5 mM NiCl₂) i presentaven el mateix tipus de cristalls, agulles llargues i fines (v. fig. 4.37), encara que la relació molar en la primera era 1:1 i en la segona 1:2. Podem observar en el gel SDS PAGE de la fig. 4.43 com s'observen unes bandes força intenses en tots dos casos. No obstant això al gel s'aprecia més d'una banda en les mostres dels cristalls. La de major massa molecular correspon a la proteïna de partida, l'HMGA1a truncada, mentre que les que es situen per sota són formes més truncades d'aquesta. Si comparem els dos cristalls, la pitjor situació és la corresponent al cristall R3-1 que va ser congelat passat més temps des de la seva aparició que l'R4-1. Així, degut al temps passat a la gota, tot i les precaucions preses, en aquests casos la proteïna apareix degradada. Tot i així, es pot verificar la presència de proteïna en aquests cristalls.

D'altra banda, en el cas dels cristalls de pèptids va ser necessari realitzar un gel amb una major concentració de poliacrilamida i una fixació diferent (v. 3.6.2) ja que són cadenes d'aminoàcids molt curtes. A la fig. 4.43 es presenta el resultat obtingut pels cristalls de l'oligonucleòtid CGCATTAATGCG i el pèptid RKPRGRPCK (*AT-hook* III). En aquest gel es van carregar també controls de l'estoc inicial de pèptid de l'ordre de nanograms. Observem que per als 3 cristalls analitzats en aquest gel només un (carril R72-1) mostra una banda corresponent a l'AT-

hook III. Aquest cristall correspon a la condició analitzada amb menor concentració de sals (2,5 mM MgCl₂, 0,25 NiCl₂ i 25 mM Tris pH=7,5). Això ja és d'esperar ja que els cations de les sals treuen els llocs d'unió del DNA i per tant, aquests cations competeixen amb els *AT-hooks* de la proteïna. No obstant això, cal tenir present que els diagrames de difracció de tots aquests cristalls presentaven desordre per degradació parcial de la mostra d'oligonucleòtid i per tant la unió dels *AT-hooks* a aquest es veu afectada. Pot ser per aquest motiu que no observem contingut peptídic en algun altre dels cristalls analitzats o bé perquè la quantitat de pèptid no arriba a ser detectable en aquesta tinció.

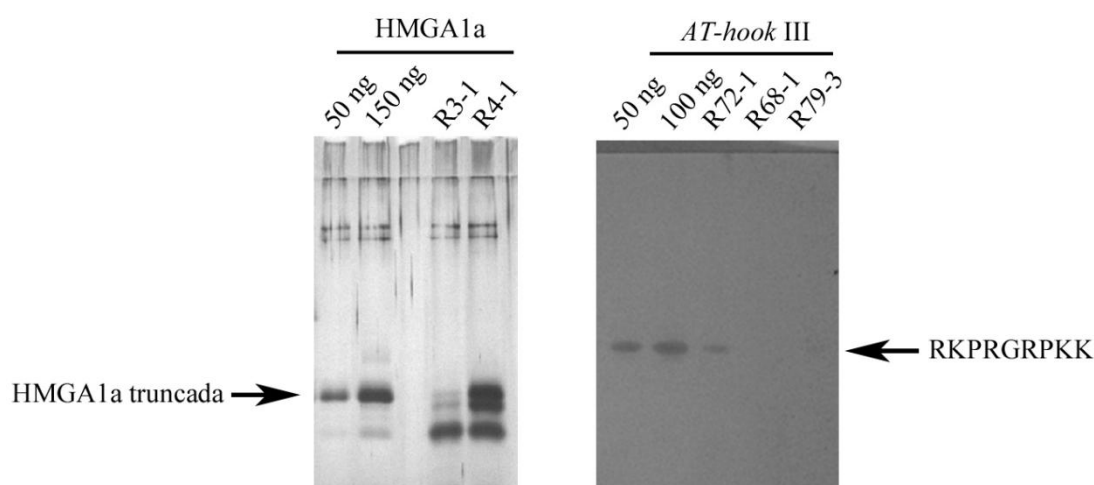


Fig. 4.43. Gels de poliacrilamida dels cristalls de la proteïna HMGA1a i del pèptid *AT-hook III*.

Gel d' SDS dels cristalls de la proteïna HMGA1a truncada amb l'oligonucleòtid CGCATTAATGCG (esquerra) i gel d'àcid acètic-urea d'alguns dels cristalls obtinguts del pèptid *AT-hook III* i aquest mateix oligonucleòtid (dreta). Es van carregar controls de la proteïna (50 i 150 ng) i del pèptid (50 i 100 ng). Tots dos es van tenyir amb plata.

Conclusions

En l'optimització de la cristal·lització de la proteïna HMGA1a truncada amb l'oligonucleòtid CGCATTAATGCG es van poder reproduir els mateixos cristalls però cap de les seves difraccions van ser millors que les corresponents a les proves anteriors. En els assajos de cristal·lització de la proteïna HMGA1b(1-79) no es va obtenir cap resultat positiu. La cristal·lització dels pèptids corresponents a l'*AT-hook II* i *III* de la proteïna va permetre l'obtenció de diversos cristalls amb diferents morfologies, però, els seus diagrames de difracció no van ser els esperats degut a la integritat de la mostra d'oligonucleòtid utilitzada. Un dels cristalls corresponents a les proves cristal·logràfiques prèvies pel pèptid PRGRP i l'oligonucleòtid AATAATTATT presentava una bona difracció la qual s'analitza a l'apartat 4.3.

4.2.2. Complexos HMGB1 Box AS-DNA

En els assajos cristal·logràfics dels complexos del domini Box AS (v. fig. 4.12) amb DNA es va treballar amb 6 oligonucleòtids diferents en total: GTTATTAATAAC, AATAAATTTATT, ATATCGATAT, AATAATCGCGATTATT i dues variants del segon (ATAAATTTATTA i AATAAATTTATTT). Amb l'elecció d'aquestes dues variants amb extrems no roms (fig. 4.44) es va voler optimitzar l'obtenció dels cristalls de la Box BS amb l'oligonucleòtid original (com es veurà a l'apartat 4.2.3) i de pas es va intentar cristal·litzar-los també amb la Box AS i l'NHP6A (v. apartat 4.2.5).

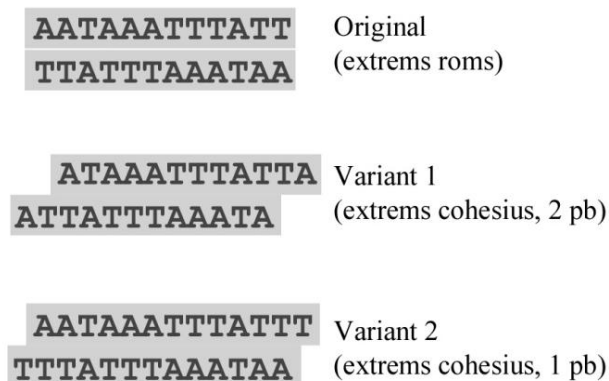


Fig. 4.44. Variants amb extrems no roms de l'oligonucleòtid AATAAATTTATT.

Per a visualitzar la formació del complex entre la proteïna i el DNA es van realitzar assajos del canvi en la mobilitat electroforètica abans de realitzar els assajos cristal·logràfics (v. 3.10.1 per més detalls). A la fig. 4.45 podem observar els gels EMSA corresponents a la formació del complex entre el domini Box A i els oligonucleòtids utilitzats per a cristal·lografia. Veiem el retard a unes relacions molars (proteïna:oligonucleòtid) d'entre 2:1 i 4:1 depenent del DNA. Cal recordar que en les nostres condicions aquests assajos són qualitius i per tant no podem comparar afinitats d'unió ja que treballem en condicions properes a la saturació (la concentració de DNA és de l'ordre de la K_D). No obstant això podem confirmar la unió entre proteïna i DNA.

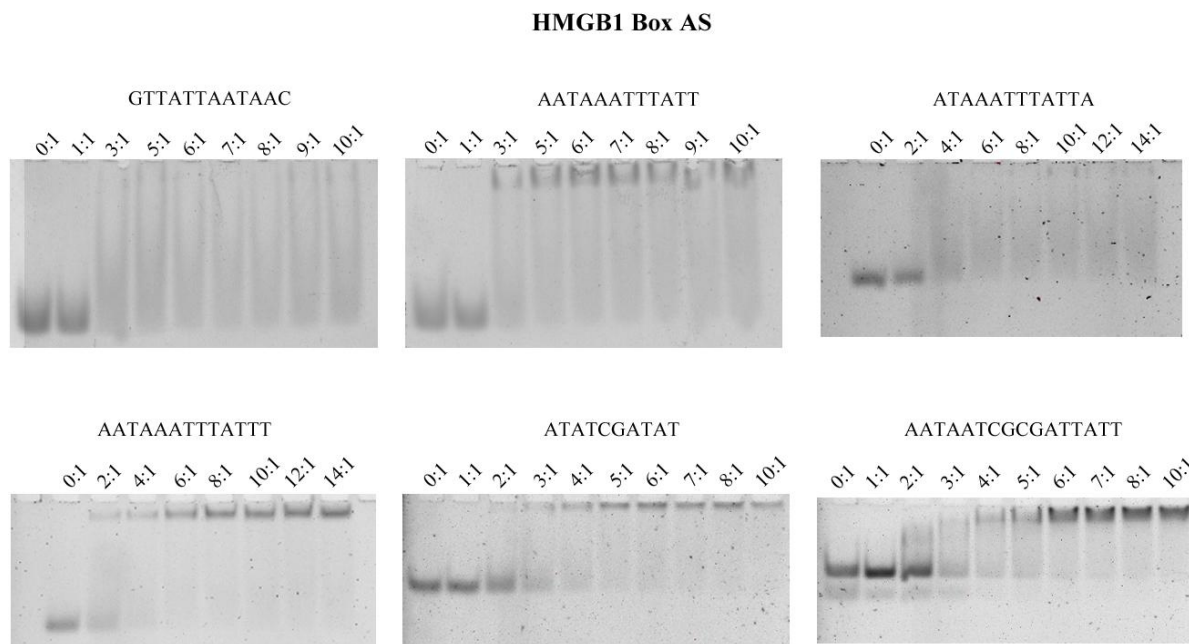


Fig. 4.45. Assaigs EMSA per a la formació de complexos entre la Box AS i diferents oligonucleòtids.

En els carrils s'indica la relació molar proteïna:DNA corresponent a cada incubació.

Resum de les condicions de cristal·lització

A la taula 4.4 es resumeixen els diferents assajos cristal·logràfics realitzats tant a la Universitat de Colorado com al laboratori de Barcelona. En aquesta taula s'ha destacat el tipus de cristall obtingut i la difracció pels millors resultats obtinguts de cada sèrie.

Els primers assajos de cristal·lització per als fragments de l'HMGB1 i l'NHP6A es van començar a realitzar al grup de la Dra. Churchill a la Universitat de Colorado. Degut a l'àmplia experiència del grup en la cristal·lització de la proteïna HMG-box de la *Drosophila*, l'HMGD, es va decidir abordar l'obtenció de cristalls dels complexos de les proteïnes HMGB tenint en compte les condicions de cristal·lització de l'HMGD. Churchill i col·laboradors varen resoldre l'estructura de dues formes d'aquesta proteïna, el fragment HMGD74 i el mutant HMGD-M13A (tots dos amb només el domini HMG-box) [39][45]. A la taula 4.3 es mostren les condicions de cristal·lització d'aquests dos cristalls corresponents a la proteïna HMGD [129][39][45].

Taula 4.3. Condicions de cristal·lització de la proteïna HMGD [129][39][44].

HMGD74	HMGD-M13A
1,9 mM Proteïna	1,2 mM Proteïna
1,5 mM GCGATATCGC	1,2 mM GCGATATCGC
-	10 mM NaCl
8 mM Tris-HCl, pH=7,85	5 mM MES-Na, pH=5,25
9,2% PEG 4000	7,6% PEG 3350
Pou: 23% PEG 4000 20 mM Tris-HCl, pH=7,85	Pou: 32% PEG 3350

Així, les primeres proves de cristal·lització de l'HMGB1 Box AS (v. taula 4.4), així com també del Box BS, del Box ABS i de l'NH6A (veure els apartats a continuació) es van realitzar aplicant unes condicions similars a les de cristal·lització de l'HMGD. Principalment, es va treballar sense addició de sals (a part de les corresponents al tampó) i amb el precipitant polimèric PEG 3350. Com ja s'ha comentat, la presència de cations obstaculitza la interacció entre la proteïna i el DNA, dificultant així l'obtenció de cristalls del complex. També es van aplicar uns rangs de pHs de 6 fins a 8. Així mateix es va treballar amb una concentració final de proteïna i DNA una mica major que en el cas de les HMGAs. Aquest cop, però, es va preferir treballar amb un excés de proteïna (relació molar (2:1)) per a facilitar-ne la formació del complex. En tots els casos s'ha incubat el complex, és a dir la barreja de la proteïna i el DNA, prèviament a la realització de les gotes. Sempre es van mantenir les gotes a 4°C, excepte en una sèrie d'assajos (indicat a la taula 4.4).

A més a més de realitzar aquests assajos cristal·logràfics sense sals addicionals, es van dur a terme proves cristal·logràfiques fent servir la matriu comercial JCGS Core Suite I (veure annex D). Aquesta matriu conté 96 condicions de cristal·lització diferents (variacions en les sals, el pH i els precipitants) i està dissenyada especialment per a l'obtenció de cristalls de proteïna. Ja a Barcelona, es van repetir algunes de les condicions "sense sal" i es va provar la cristal·lització del Box AS amb alguns tampons Natrix HR2-116 (v. taula 4.4 i annex D).

A continuació s'explicarà amb més detall la cristal·lització de la Box AS amb cada oligonucleòtid.

Taula 4.4. Resum dels assajos cristal·logràfics per a la proteïna HMGB1 Box AS.

Proteïna	Oligonucleòtid	Ràtio P:D (mM)	Condicions			nº gotes	Cristalls	Difracció
			Sal	pH	Precip.			
HMGB1 Box AS Colorado	GTTATTAATAAC	0,8;0,4	-	6; 6,5; 7; 7,5; 8	PEG3350 MPD	84	Mini-cristalls	-
	AATAAATTTATT	0,8;0,4	Matriu JCSG** (totes les 96 condicions)			96	Cristalls petits	No difracten
		0,8;0,4	-	6; 6,5; 7; 7,5; 8	PEG3350 MPD	90	-	-
		0,8;0,4* 0,6;0,3*	-	6; 6,5; 7; 7,5; 8	MPD	40 (4 i 18°C)	-	-
HMGB1 Box AS Barcelona	AATAAATTTATT	0,8;0,4	Natrix (nº 4, 7, 10, 15, 16, 22, 25, 31, 34, 35, 36, 42, 43, 47)			22	Mini-varetes; Làmines	Bona fins ~3Å
		0,8;0,4	-;	6; 6,5; 7; 7,5; 8	PEG3350	5	-	-
		0,8;0,4	MgAc	5,5; 6; 6,5	MPD	4	Làmines	Bona fins ~3Å
	ATAAATTTATTA (v1)	0,8;0,4	-	6; 6,5; 7; 7,5; 8	PEG3350	5	Mini-cristalls	-
	AATAAATTTATTT (v2)	0,8;0,4	-	6; 6,5; 7; 7,5; 8	PEG3350	5	-	-
	ATATCGATAT	0,8;0,4	Natrix (nº 7, 10, 15, 16, 42, 43)			8	Agulles; Varetes planes; Cristall gran	Bones difraccions fins ~2Å
		0,8;0,4	-	6; 7; 8	PEG3350 MPD	6	Agulles	Bona fins ~2,5Å
	AATAATCGCGATTATT	0,8;0,4	Natrix (nº 7, 10, 15, 16, 42, 43)			6	Mini-varetes	-
0,8;0,4		-	6; 7; 8	PEG3350 MPD	6	Mini-varetes	-	

*Es van realitzar gotes barrejant diferents proporcions del tampó i del complex [(3:2)(2:1)(1:1)(1:2)(2:3)], a la taula només es mostren les concentracions finals per a les barreges (1:1). **Gotes preparades automàticament amb robot.

Els oligonucleòtids ATAAATTTATTA (v1) i AATAAATTTATTT (v2) són variants de l'oligonucleòtid AATAAATTTATT, presenten igual seqüència nucleotídica a excepció dels extrems (v. text). La ràtio P:D és la ràtio molar proteïna:DNA. Tots els assajos s'han realitzat a 4°C (en cas contrari està indicat a la taula).

Resultats de la cristal·lització

Les proves cristal·logràfiques s'han realitzat tant en plaques com en càpsules. En plaques s'indica en primer lloc el nombre de la placa i tot seguit la posició de la gota en ella (per exemple, RS01-A1). Les càpsules estan nomenades amb el nom de la proteïna o fragment i el nombre de gota (per exemple, AS01). Les gotes són del tipus assentada a les plaques i del tipus penjada a les càpsules.

a) *HMGB1 Box AS - GTTATTAATAAC*

En la cristal·lització de la proteïna HMGB1 Box A amb l'oligonucleòtid GTTATTAATAAC s'ha treballat amb unes condicions sense sals addicionals a totes les gotes preparades. Com a precipitant s'ha fet servir tant PEG 3350 com MPD. En el primer cas, no hem obtingut cap cristall en cap dels assajos, només s'ha observat precipitat i separació de fases. Utilitzant l'MPD com a precipitant, s'han obtingut cristalls molt petits en tres gotes i precipitat cristal·lí en una altra. Les quatre són la mateixa condició però a diferents pHs (10 mM NaCac o bé Tris-HCl i 5% MPD). A major (10%) o menor (2,5%) concentració de precipitant a la gota no s'han obtingut cristalls, tampoc a les gotes amb major concentració (25 mM) del tampó. A la imatge de la fig. 4.46 s'observen aquests petits cristalls d'una de les gotes juntament amb una taula amb les condicions de cristal·lització d'aquesta. Aquests cristalls no han crescut més i no se n'ha pogut pescar cap degut a la seva mida.

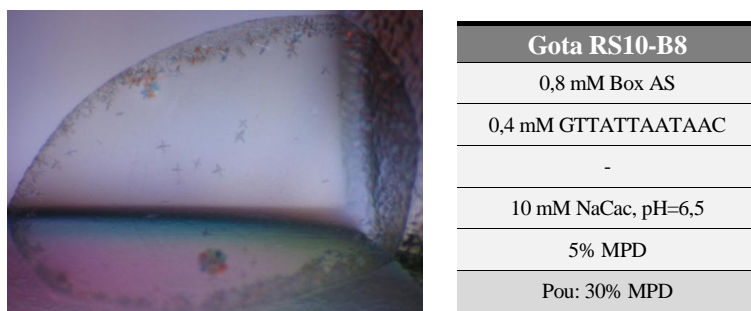


Fig. 4.46. Imatge i condicions de cristal·lització de la gota RS10-B8 corresponent a la proteïna Box AS i a l'oligonucleòtid GTTATTAATAAC.

b) *HMGB1 Box AS - AATAAATTTATT*

S'ha obtingut precipitat cristal·lí i esferulites en les millors gotes amb l'oligonucleòtid AATAAATTTATT en les condicions sense sals tant amb PEG 3350 com amb MPD. En algunes gotes amb PEG de les esferulites han crescut eriçons d'agulles curtes però no presentaven un bon aspecte.

La cristal·lització d'aquest mateix oligonucleòtid amb la matriu comercial JCGS Core Suite I (veure annex D) va produir dos *hits* de les 96 condicions possibles. Van créixer cristalls petits en les dues condicions de la matriu que es mostren a la fig. 4.47 i la fig. 4.48. Un dels cristalls amb millor forma de la gota RS13-H4 va ser difractat però el diagrama de difracció no presentava taques, probablement per la seva reduïda mida. En canvi els cristalls de la gota RS13-C4 no es van arribar a pescar ja que eren massa petits, no presentaven costats rectes i podrien ser cristall líquid.

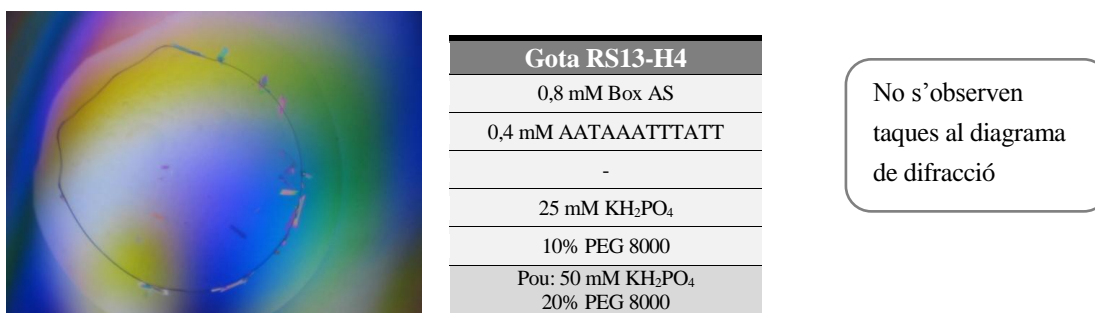


Fig. 4.47. Imatge i condicions de cristal·lització de la gota RS13-H4 corresponent a la proteïna Box AS i a l'oligonucleòtid AATAAATTTATT.

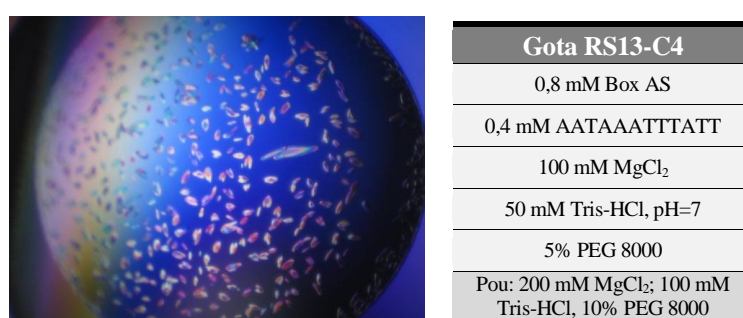


Fig. 4.48. Imatge i condicions de cristal·lització de la gota RS13-C4 corresponent a la proteïna Box AS i a l'oligonucleòtid AATAAATTTATT.

En els assajos amb les condicions de la matriu Natrix HR2-116 es van obtenir dos resultats positius. Per una banda uns cristalls amb forma de vareta petits a la gota AS19 (v. fig. 4.49) que no varen ser congelats deguts a la seva mida i a què no acabaven d'estar definits a les puntes.

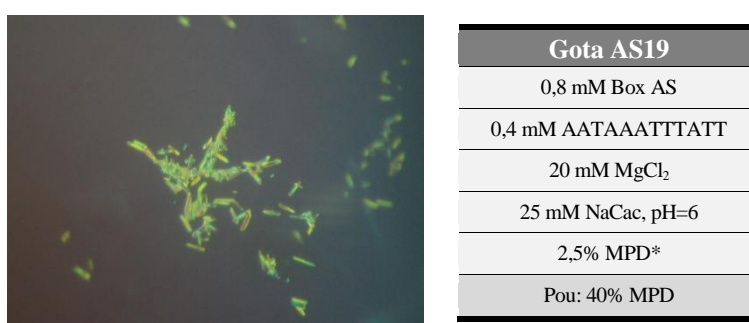


Fig. 4.49. Imatge i condicions de cristal·lització de la gota AS19 corresponent a la proteïna Box AS i a l'oligonucleòtid AATAAATTTATT.

*Corresponent al tampó Natrix n°15.

D'altra banda a la gota AS20 va créixer una agrupació de cristalls en forma de flor formada per unes làmines amb unes puntes molt rectes i molt birefringents (v. fig. 4.50). Amb molta cura es van separar les làmines una a una i es van congelar un total de 8 cristalls.

Es va intentar optimitzar l'obtenció d'aquestes làmines tan fines per tal d'aconseguir cristalls més separats i més gruixuts. Amb aquesta finalitat es van realitzar rèpliques, repeticions variant el pH, diferents gotes amb sembres de la gota mare (*seedings*) i també alguns dels cristalls es van congelar amb crioprotecció. Es van obtenir aquest mateix tipus de cristalls en totes les gotes realitzades, en aquesta ocasió, amb més d'un apilament de làmines a les gotes. Es van congelar 11 cristalls més corresponents a aquestes optimitzacions. Gairebé tots els cristalls presentaven una difracció de fins 3,2 Å, però en la majoria es podia apreciar la presència de més d'un cristall a la difracció. Probablement perquè separar perfectament aquestes làmines apilades presentava moltes dificultats. Tot i així, un dels cristalls de la gota mare (AS20-1) mostrava un diagrama que podria correspondre a un cristall únic. El seu diagrama de difracció es mostra a la fig. 4.50.

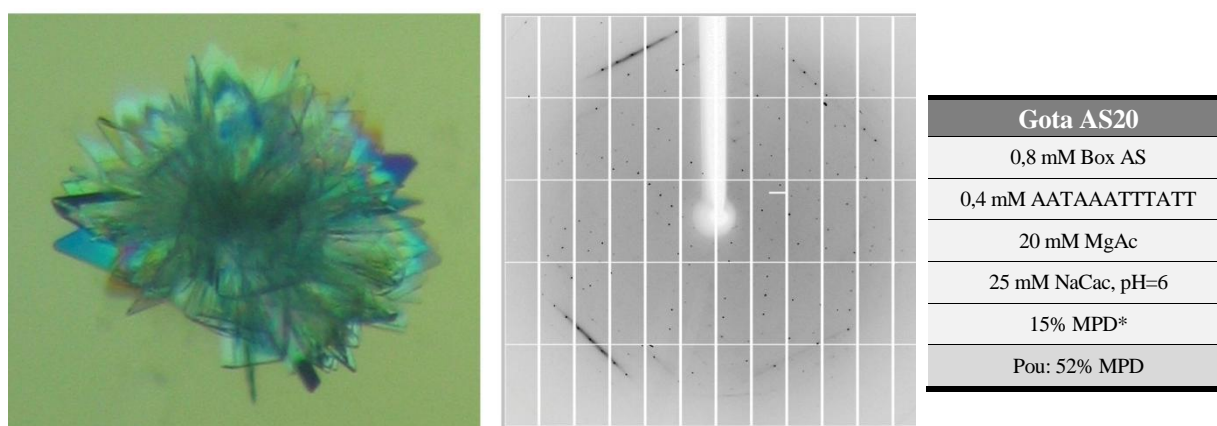


Fig. 4.50. Imatge, condicions de cristallització i diagrama de difracció d'un cristall corresponent al domini Box AS i l'oligonucleòtid AATAAATTTATT (gota AS20, cristall AS20-1).

*Corresponent al tampó Natrix n°16 (v. annex D). Cristall difractat al sincrotró ALBA (línia Xaloc): angle d'oscil·lació 2°, temps d'exposició 2 s, distància al detector 600 mm i longitud d'ona 0,9795 Å.

En el diagrama de difracció del cristall AS20-1 observem com es veuen dos *stackings* intensos corresponents al DNA (línies a la part superior i part inferior esquerra del diagrama). Aquest fet indica dues direccions en les que es troben posicionades les cadenes de DNA. Més endavant, a l'apartat 4.3, es presenten els resultats de l'anàlisi dels diagrames de difracció corresponents a aquests cristalls.

c) *HMGB1 Box AS - variants AATAAATTTATT*

Amb les variants de l'oligonucleòtid AATAAATTTATT v1 i v2 (v. fig. 4.44) no s'ha obtingut cap cristall prometedor. Només es van obtenir en una de les 10 gotes realitzades uns cristalls molt petits en forma de creu (en presència de 25 mM NaCac pH=6,5 i PEG 3350).

d) *HMGB1 Box AS - ATATCGATAT*

Els resultats de cristal·lització de l'oligonucleòtid ATATCGATAT amb el Box AS van ser els millors obtinguts al llarg d'aquest treball. Han crescut cristalls en 4 de les condicions Natrix HR2-116 i en una gota amb tampó sense sals i MPD. En total s'han congelat 13 cristalls corresponents a aquest conjunt d'assajos cristal·logràfics.

Les 6 condicions Natrix provades (v. taula 4.4) contenen el catió Mg^{+2} i només no varen cristal·litzar les proves que presentaven isopropanol a la seva composició (v. annex D). Així s'han obtingut cristalls dels tampons Natrix número 10, 15, 16 i 43. En la gota que contenia PEG 4000 com a precipitant van créixer uns cristalls petits en forma de varetes curtes no aptes per a congelar (v. fig. 4.51). En canvi, en les altres tres gotes que presentaven MPD es van obtenir agulles o varetes suficientment grans i amb molta birefringència que van poder ser pescats.

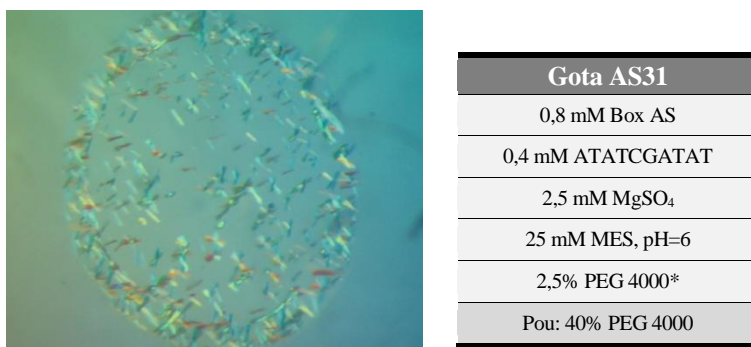


Fig. 4.51. Imatge i condicions de cristal·lització de la gota AS31 corresponent a la proteïna Box AS i a l'oligonucleòtid ATATCGATAT.

*Corresponent al tampó Natrix n°10.

A la fig. 4.52 es pot observar el cristall que va créixer a la gota AS32 (20 mM MgCl₂, 25 mM NaCac pH=6 i MPD, Natrix 15). D'aquesta gota es va obtenir, inicialment, aquest únic cristall de 1000 x 200 μm de mida. Tot i ser un cristall ideal, molt gran i birefringent amb puntes rectes i únic, no va presentar una bona difracció. Al corresponent diagrama de difracció de la fig. 4.52 podem veure com s'obtenen només taques a baixa resolució però molt ordenades.

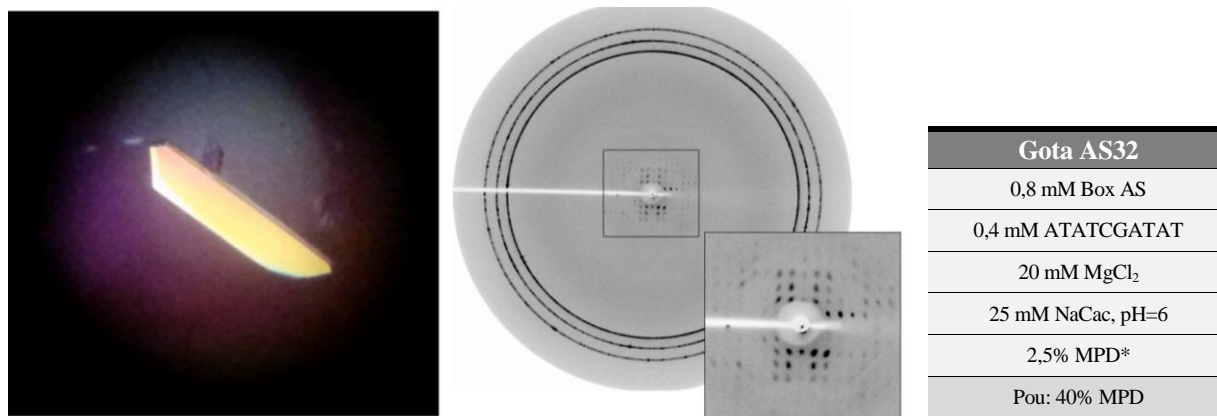


Fig. 4.52. Imatge, condicions de cristal·lització i diagrama de difracció d'un cristall corresponent al domini Box AS i l'oligonucleòtid ATATCGATAT (gota AS32, cristall AS32-1).

*Corresponent al tampó Natrix n°15 (v. annex D). Cristall difractat a la Plataforma Automatitzada de Cristal·lografia del PCB: angle d'oscil·lació 5°, temps d'exposició 600 s, distància al detector 300 mm i longitud d'ona 1,542 Å.

Aquest cristall va ser retornat a la gota i a continuació van aparèixer unes agulles molt birefringents i suficientment gruixudes per a ser congelades i difractades. La difracció d'aquests cristalls va ser molt bona arribant-se a observar taques a 1,8 Å de resolució. A la fig. 4.53 podem observar cristalls tipus agulla obtinguts en aquesta gota (eren menys i presentaven millor aspecte al ser pescats) i la difracció corresponent al cristall AS32-3.

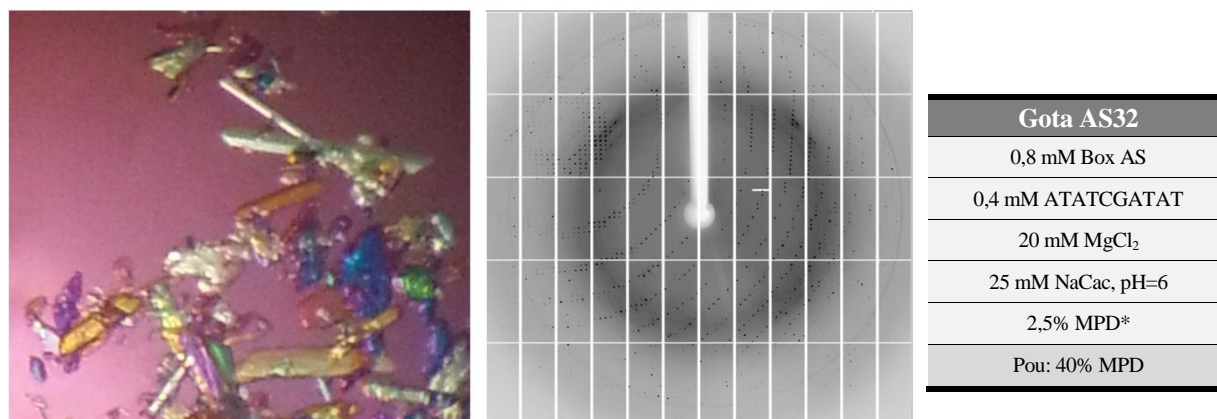


Fig. 4.53. Imatge, condicions de cristal·lització i diagrama de difracció d'un cristall corresponent al domini Box AS i l'oligonucleòtid ATATCGATAT (gota AS32, cristall AS32-3).

*Corresponent al tampó Natrix n°15 (v. annex D). La fotografia va ser realitzada un cop pescats els cristalls i, per tant, aquests no presenten tan bon aspecte. Cristall difractat al sincrotró ALBA (línia Xaloc): angle oscil·lació 1°, temps d'exposició 4 s, distància al detector 425 mm i longitud d'ona 0,9795 Å.

La repetició exacta d'aquesta condició va produir un resultat similar, per una part van créixer uns cristalls grans en forma de vareta i al voltant unes agulles més fines. A la fig. 4.54 es mostren aquests cristalls i un dels diagrames de difracció obtinguts per a un fragment del cristall gran pescat de la gota, on podem observar també una molt bona resolució. Cal comentar que els diagrames de difracció d'aquesta gota (AS67, v. fig. 4.54) i de l'original (AS32, v. fig. 4.53) presenten un perfil de taques diferent corresponents a unes cel·les unitàries també diferents com s'explicarà amb més detall més endavant (v. apartat 4.3). Se sap que en les macromolècules la forma externa del cristall (morfologia del cristall) està relacionada amb l'estructura interna (empaquetament de les molècules). Normalment l'obtenció de cristalls de diferents morfologies sol estar relacionada amb la naturalesa de la sal, el pH o la temperatura però es poden obtenir dues morfologies en una mateixa condició si es treballa en el límit de cristal·lització de dues formes de cristalls [130].

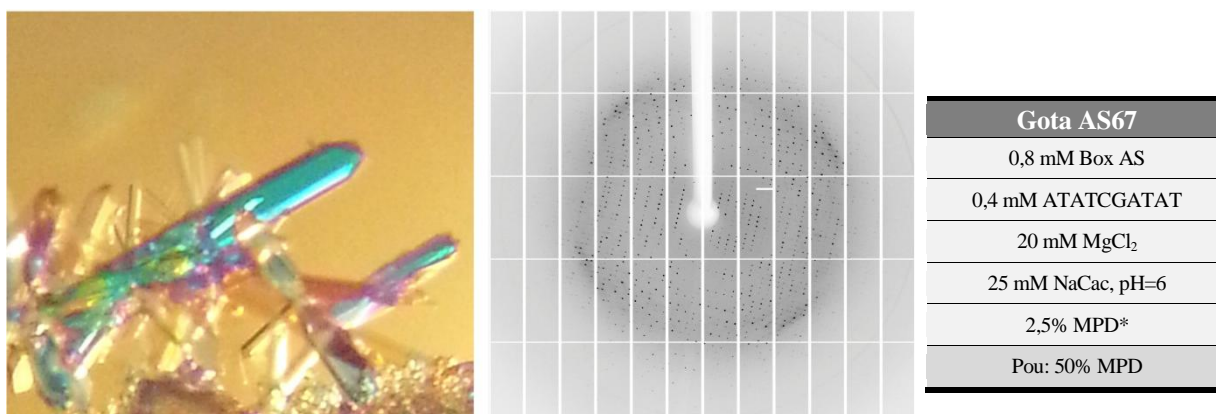


Fig. 4.54. Imatge, condicions de cristal·lització i diagrama de difracció d'un cristall corresponent al domini Box AS i l'oligonucleòtid ATATCGATAT (gota AS67, cristall AS67-1).

*Corresponent al tampó Matrix n°15 (v. annex D). La fotografia va ser realitzada un cop pescats els cristalls i, per tant, aquests no presenten tan bon aspecte. Cristall difractat al sincrotró ALBA (línia Xaloc): angle d'oscil·lació 1°, temps d'exposició 2 s, distància al detector 450 mm i longitud d'ona 0,9795 Å.

En canvi, vam observar que una altra morfologia cristal·logràfica trobada a la gota AS35, on van créixer uns cristalls en forma de varetes planes (làmines en forma de vareta) (v. fig. 4.55) donaven el mateix tipus de difracció (és a dir, la mateixa cel·la) que el cristall AS32-3. Aquesta gota té en comú amb la gota AS32 la sal de clorur de magnesi, però a menor concentració i amb 25 mM de sulfat d'amoni i pH diferent. Aquest cristall difractava igualment fins a una molt bona resolució (taques fins 1,8 Å) (v. fig. 4.55).

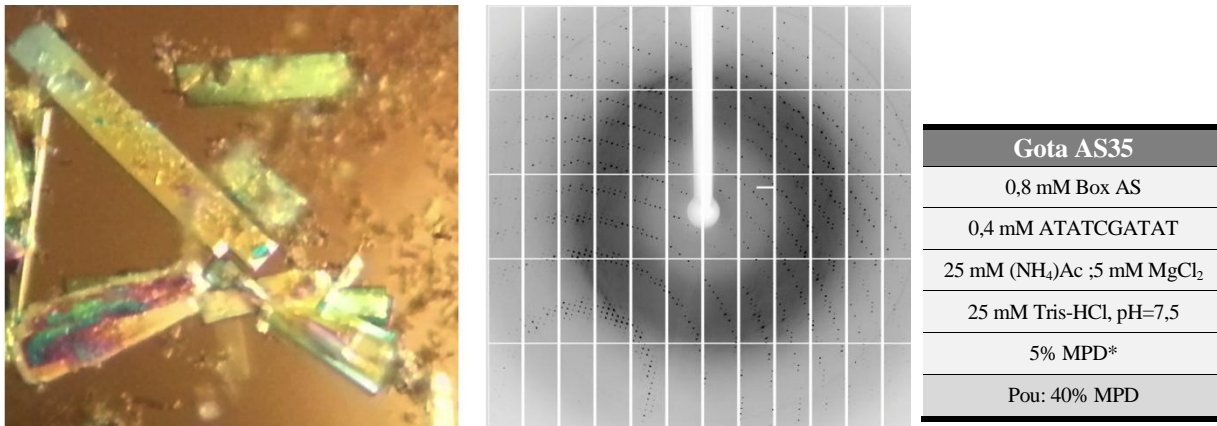


Fig. 4.55. Imatge, condicions de cristal·lització i diagrama de difracció d'un cristall corresponent al domini Box AS i l'oligonucleòtid ATATCGATAT (gota AS35, cristall AS35-2).

*Corresponent al tampó Matrix n°43 (v. annex D). La fotografia va ser realitzada un cop pescats els cristalls i, per tant, aquests no presenten tan bon aspecte. Cristall difractat al sincrotró ALBA (línia Xaloc): angle d'oscil·lació 1°, temps d'exposició 4 s, distància al detector 450 mm i longitud d'ona 0,9795 Å.

Finalment, també es van obtenir uns cristalls prometedors amb una condició sense sals i MPD com a precipitant. En aquesta gota es van formar unes agulles molt brillants i amb molt bona forma que també van ser pescades i difractades. La condició d'aquesta gota presenta el mateix tampó, pH i precipitant que la gota AS32 (v. fig. 4.53) però aquest cop sense l'addició de les sals de clorur de magnesi. A la fig. 4.56 s'observa una imatge de la gota, ja amb un pitjor aspecte, un cop pescats els cristalls. Els cristalls pescats eren agulles com les assenyalades amb fletxes a la imatge. El diagrama de difracció que s'observa presenta una resolució de 2,5 Å i a diferència dels diagrames anteriors no s'observen unes taques més intenses en la zona de l'*stacking* del DNA (v. apartat 4.3 per a més detalls).

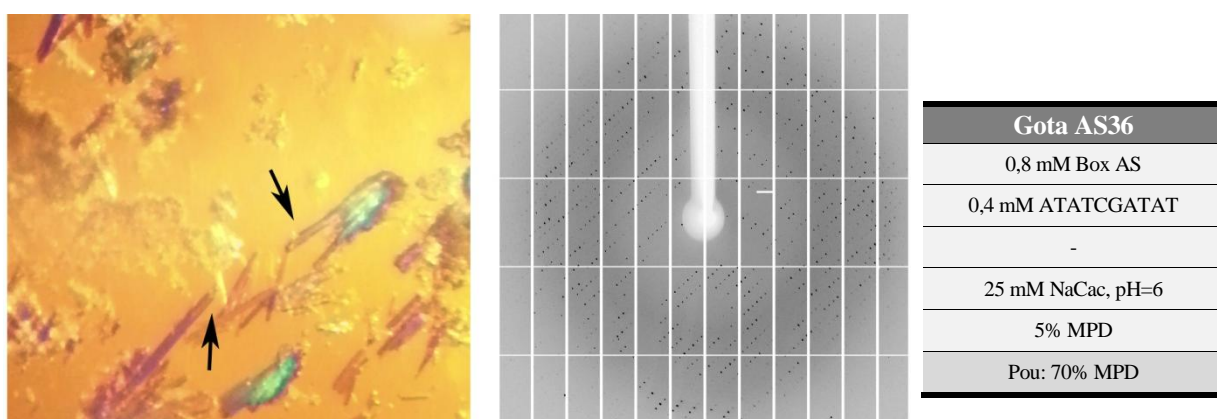


Fig. 4.56. Imatge, condicions de cristal·lització i diagrama de difracció d'un cristall corresponents al domini Box AS i l'oligonucleòtid ATATCGATAT (gota AS36, cristall AS36-1).

La fotografia va ser realitzada un cop pescats els cristalls i, per tant, aquests no presenten tan bon aspecte. Cristall difractat al sincrotró ALBA (línia Xaloc): angle d'oscil·lació 2°, temps d'exposició 2 s, distància al detector 600 mm i longitud d'ona 0,9795 Å.

e) *HMGB1* Box AS - AATAATCGCGATTATT

Les proves de cristal·lització corresponents a la proteïna Box A amb l'oligonucleòtid AATAATCGCGATTATT en les condicions sense sals varen presentar separació de fases (amb precipitant MPD) i precipitat cristal·lí (amb PEG 3350). Així no es va congelar cap cristall corresponent a aquestes gotes.

En les gotes realitzades amb aquest oligonucleòtid amb la matriu Natrix només es van obtenir uns cristalls petits en forma de varetes curtes (Natrix nº15, 20 mM MgCl₂, 25 mM NaCac pH=6 i MPD) que no es van arribar a pescar.

Electroforesi de cristalls

Per a verificar la presència de proteïna als cristalls s'ha procedit a analitzar electroforèticament una selecció dels cristalls més representatius obtinguts en la cristal·lització del domini Box AS. En el gel de poliacrilamida SDS de la fig. 4.57 observem bandes corresponents a la proteïna Box AS per a tots els cristalls analitzats, confirmant així la presència proteica als cristalls.

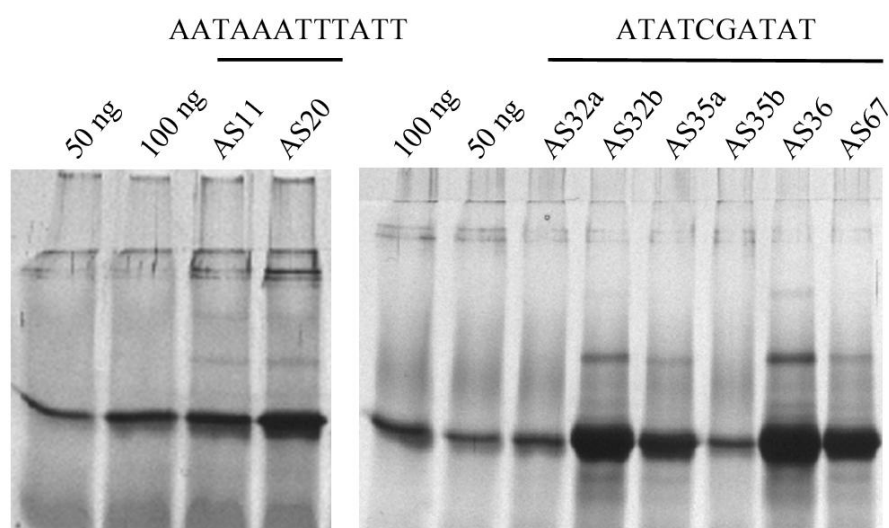


Fig. 4.57. Gel poliacrilamida-SDS dels cristalls de la proteïna Box AS.

S'han carregat al gel controls de proteïna a 50 i 100 ng.

Cristalls amb l'oligonucleòtid AATAAATTTATT: AS11, eriçó d'agulles; AS20, làmines (v. fig. 4.50).

Cristalls amb l'oligonucleòtid ATATCGATAT: AS32a, agulla (v. fig. 4.53); AS32b, fragment del cristall gran (v. fig. 4.52); AS35a, 3-4 varetes làmina (v. fig. 4.55); AS35b, 1 vareta làmina (v. fig. 4.55); AS36, grup d'agulles juntes (v. fig. 4.56); AS67, grup d'agulles juntes (v. fig. 4.54).

Conclusions

Els millors resultats d'aquest treball s'han observat amb aquest fragment, el Box AS. S'han obtingut bones difraccions tant amb l'oligonucleòtid AATAAATTTATT així com amb l'ATATCGATAT. En el primer cas es van formar unes làmines molt fines difícils de separar entre elles i per tant va ser complicat obtenir diagrames de difracció per a un cristall únic. Les varetes i agulles obtingudes per a l'ATATCGATAT no presentaven aquestes dificultats i mostraven uns diagrames de difracció fins a 1,8 Å de resolució. Es van calcular tres cel·les unitàries diferents per als cristalls ATATCGATAT-Box AS com es veurà a l'apartat 4.3. Cal destacar que una mateixa condició de cristallització va presentar dues morfologies de cristalls (varetes grans i agulles) amb cel·les diferents mentre que dos cristalls de gotes amb condicions i formes diferents (agulles i varetes planes) van presentar el mateix empaquetament.

4.2.3. Complexos HMGB1 Box BS-DNA

La cristal·lització del fragment Box BS (v. fig. 4.12) amb els mateixos 6 oligonucleòtids s'ha realitzat de forma paral·lela a les proves de cristal·lització amb el Box AS. Per aquest motiu, s'han aplicat unes condicions de cristal·lització similars (v. 4.2.2), com veurem a continuació. Abans però de realitzar els assajos de cristal·lització s'ha comprovat la formació dels complexos en gels EMSA.

A la fig. 4.58 podem veure els assajos de canvi de mobilitat electroforètica corresponents a la proteïna Box BS amb els diferents oligonucleòtids fets servir per a cristal·lografia. Aquest tipus d'assaig ens permet visualitzar la formació del complex proteïna-DNA com un retard en la migració de les bandes al gel (v. 3.10.1). Podem observar com amb aquest domini les bandes del complex són més difuses i veiem llapissades en més casos que en el cas del Box AS.

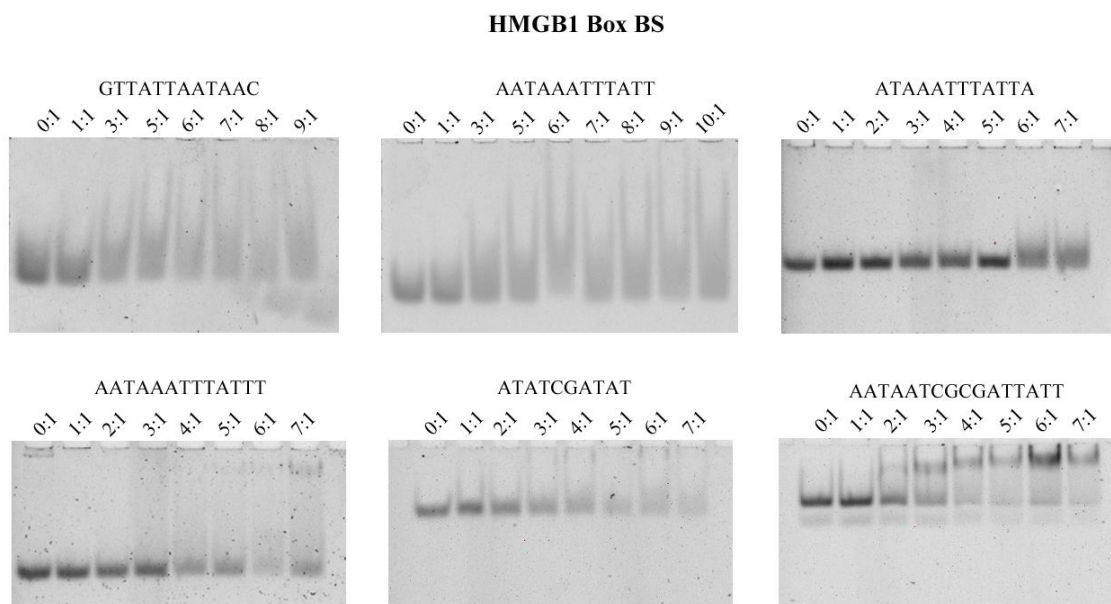


Fig. 4.58. Assaigs EMSA per a la formació de complexos entre la Box BS i diferents oligonucleòtids.

En els carrils s'indica la relació molar proteïna:DNA corresponent a cada incubació.

Així, si comparem aquests gels amb els corresponents gels per a la Box AS (v. fig. 4.45) sembla que la Box A forma complexos més estables i a menor relació molar que la Box BS. Aquest resultat ja és coherent ja que tot i presentar tots dos dominis constants d'unió de l'ordre d' $1 \mu\text{M}$ (K_B), la Box A s'uneix més fortament al DNA lineal que el Box B (el domini Box B presenta una menor K_B degut a una major entalpia d'unió desfavorable) [36].

Resum de les condicions de cristal·lització

A la taula 4.5 es detallen els assajos cristal·logràfics portats a terme amb el fragment Box BS i els diferents oligonucleòtids. En primer lloc es van fer unes proves sense sal (tant amb PEG 3350 com amb MPD), a continuació es van realitzar assajos de cristal·lització amb la matriu comercial JCSG i finalment amb els tampons Natrix HR2-116 (v. annex D).

Taula 4.5. Resum dels assajos cristal·logràfics per a la proteïna HMGB1 Box BS.

Proteïna	Oligonucleòtid	Ràtio P:D (mM)	Condicions			nº gotes	Cristalls	Difracció
			Sal	pH	Precip.			
HMGB1 Box BS Colorado	GTTATTAATAAC	0,8;0,4	-	6; 6,5; 7; 7,5; 8	PEG3350 MPD	84	-	-
	AATAAATTTATT	0,8;0,4	Matriu JCSG** (totes les 96 condicions)			96	Cristalls petits	-
		0,8;0,4	-	6; 6,5; 7; 7,5; 8	PEG3350 MPD	129	Cristalls petits, agulles	Poques taques
		0,8;0,4* 0,6;0,3*	-	6; 6,5; 7; 7,5; 8	MPD	40 (4 i 18°C)	Cristalls petits; agulles	Poques taques
HMGB1 Box BS Barcelona	AATAAATTTATT	0,8;0,4	Natrix (nº 4, 7, 10, 15, 16, 22, 25, 31, 34, 35, 36, 42, 43, 47)			16	Agulles fines; varetes	No difracten
		0,8;0,4 0,4;0,2	-	6; 6,5; 7; 7,5; 8	PEG 3350	27	Agulles	Poques taques
		0,8;0,4	MgCl ₂	6; 7; 8, 9	MPD	8	Varetes	Baixa resolució
		0,8;0,4	Matriu de sals [MgCl ₂ , MgAc, CaCl ₂ , KCl, (NH ₄)Ac, (NH ₄)Cl, NaCl, KAc, CaAc, NaCit] + Tris-HCl pH=7,5;8+PEG 3350			48	Agulles; Cristalls vareta	Bona difracció fins ~3,5Å
	ATAAATTTATTA (v1)	0,8;0,4	-	6; 6,5; 7; 7,5; 8	PEG3350	5	Precipitat cristal·lí	-
		0,8;0,4	MgCl ₂	6; 7; 8, 9	MPD	8	Precipitat cristal·lí	
	AATAAATTTATTT (v2)	0,8;0,4	-	6; 6,5; 7; 7,5; 8	PEG3350	5	-	-
		0,8;0,4	MgCl ₂	6; 7; 8, 9	MPD	20	Varetes	Baixa resolució
	ATATCGATAT	0,8;0,4	Natrix (nº 7, 10, 15, 16, 42, 43)			6	-	-
		0,8;0,4	-	6; 7; 8	PEG3350 MPD	6	-	-
	AATAATCGCGATTATT	0,8;0,4	Natrix (nº 7, 10, 15, 16, 42, 43)			6	Mini varetes	-
		0,8;0,4	-	6; 7; 8	PEG3350	6	-	-

*Es van realitzar gotes barrejant diferents proporcions del tampó i del complex [(3:2)(2:1)(1:1)(1:2)(2:3)]; a la taula només es mostren les concentracions finals per a les barreges (1:1). **Gotes preparades automàticament amb robot.

Els oligonucleòtids ATAAATTTATTA (v1) i AATAAATTTATTT (v2) són variants de l'oligonucleòtid AATAAATTTATT, presenten igual seqüència nucleotídica a excepció dels extrems (v. text). La ràtio P:D és la ràtio molar proteïna:DNA. Tots els assajos s'han realitzat a 4°C, en cas contrari està indicat a la taula.

A la Universitat de Colorado es va obtenir major nombre de cristalls per al domini Box BS que per a les altres HMG-box, especialment en les condicions sense sals i amb l'oligonucleòtid AATAAATTTATT. Però, com veurem a continuació, aquests no presentaven uns bons diagrames de difracció. Per a optimitzar aquests resultats de cristal·lització, es va decidir canviar l'oligonucleòtid AATAAATTTATT per uns de molt semblants però amb extrems no roms: els oligonucleòtids ATAAATTTATTA i AATAAATTTATTT (v. fig. 4.44). Així, amb aquests dos oligonucleòtids es van realitzar assajos cristal·logràfics sense sals i també amb sals de MgCl₂ a Barcelona.

Una altra forma de millorar aquests resultats obtinguts amb el Box BS i l'oligonucleòtid AATAAATTTATT va ser realitzar un conjunt de gotes amb una varietat àmplia de sals. Es van fer servir les sals de MgCl₂, MgAc, CaCl₂, KCl, (NH₄)Ac, (NH₄)Cl, NaCl, KAc, CaAc i NaCit a una concentració final de 5 mM o 20 mM, amb el tampó Tris-HCl a pH=7,5 i com a precipitant el PEG 3350.

Resultats de cristal·lització

a) HMGB1 Box BS - GTTATTAATAAC

No es va obtenir un bon resultat amb l'oligonucleòtid GTTATTAATAAC i la Box BS i no va créixer cap cristall en les gotes que presentaven aquest oligonucleòtid. Únicament es van obtenir esferulites en algunes de les proves realitzades.

b) HMGB1 Box BS - AATAAATTTATT

Els millors resultats de cristal·lització amb la Box BS s'han obtingut amb aquest oligonucleòtid i principalment amb el precipitant PEG com es veurà a continuació.

Amb la matriu JSCG s'han obtingut cristalls petits en dues de les 96 condicions, els quals no es van arribar a congelar degut a la seva mida i morfologia. A la fig. 4.59 es mostra una d'aquestes gotes en la qual van créixer unes agulles fines i curtes.

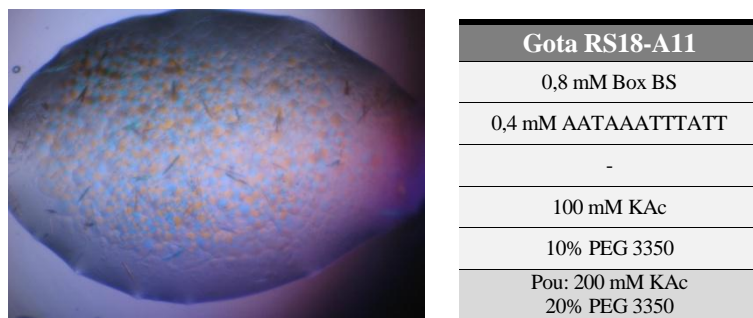


Fig. 4.59. Imatge i condicions de cristal·lització de la gota RS18-A11 corresponent a la proteïna Box BS i a l'oligonucleòtid AATAAATTTATT.

En les gotes amb tampó sense sals i que tenien com a precipitant l'MPD no es va obtenir cap cristall. Aquestes gotes presentaven majoritàriament separació de fases i en alguns casos separació de fases amb unes poques esferulites.

En canvi en les condicions sense sals que contenien PEG 3350 es van obtenir unes agulles fines i llargues en la majoria de les gotes. En moltes ocasions aquestes agulles creixien en un gran nombre formant-se principalment en els límits de la gota com podem observar a la fig. 4.60. Es van difractar un total de 8 agulles d'aquest tipus però els diagrames de difracció corresponents presentaven molt poques taques a baixa resolució (tot i que ben definides). A la dreta de la fig. 4.61 es mostra un diagrama de difracció corresponent a l'agulla assenyalada a l'esquerra.



Fig. 4.60. Imatge i condicions de cristal·lització de la gota RS15-B4 corresponent a la proteïna Box BS i a l'oligonucleòtid AATAAATTTATT.

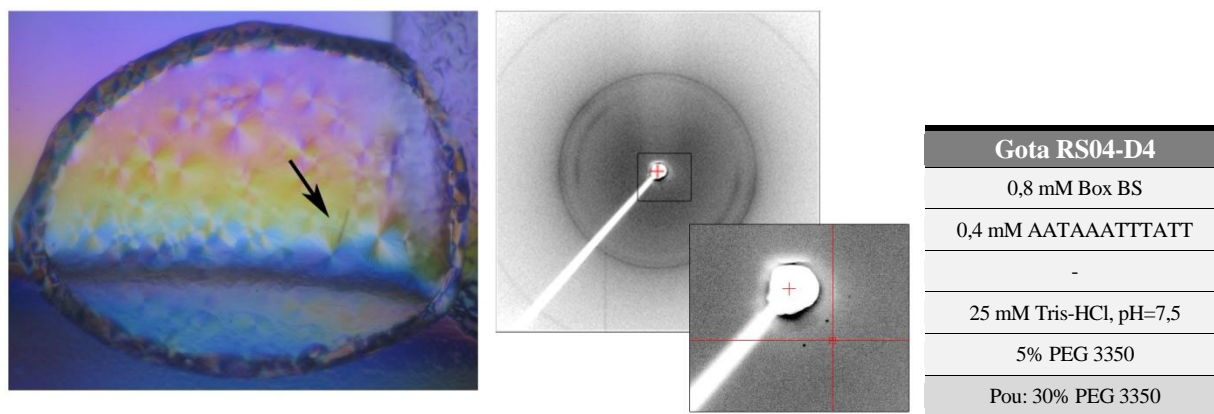


Fig. 4.61. Imatge, condicions de cristal·lització i diagrama de difracció d'un cristall corresponent al domini Box BS i l'oligonucleòtid AATAAATTTATT (gota D4 de la placa RS04, cristall RS04-D4-3).

Cristall difractat a la plataforma cristal·logràfica de la *University of Colorado*: angle d'oscil·lació 0,5°, temps d'exposició 900 s, distància al detector 200 mm i longitud d'ona 1,54 Å.

En aquestes condicions (absència de sals i PEG 3350) es va obtenir també un altre tipus de cristalls. Aquests cristalls presentaven diferents morfologies però tenien en comú que creixien a partir d'un precipitat. És a dir, al realitzar la gota es podia observar un precipitat amorf (no birefringent) en elles que es mantenia després de molts dies. Però després d'un o dos mesos sobre el precipitat van aparèixer uns cristalls els quals creixien a costa del precipitat. A la fig. 4.62 podem observar aquest creixement en una de les gotes on inicialment només hi havia precipitat amorf. La cristal·lització a partir de precipitats es pot produir a partir d'un precipitat amorf (com és el nostre cas) o a partir d'un precipitat cristal·lí (fenomen conegut com a maduració d'Ostwald) [131].

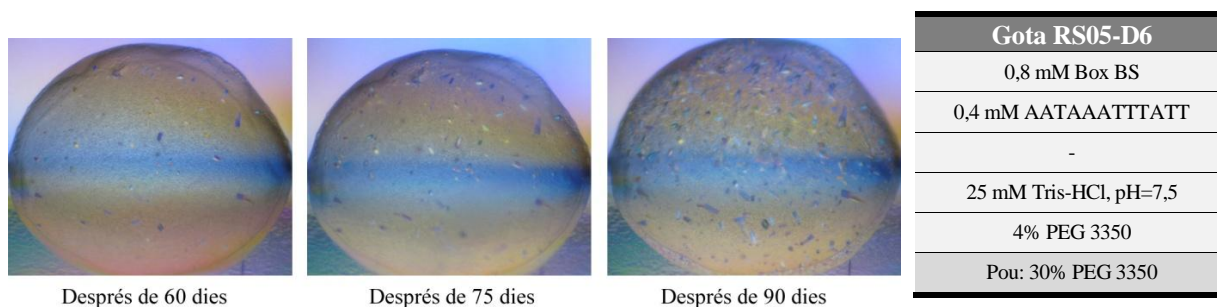


Fig. 4.62. Evolució dels cristalls d'una gota i les seves condicions de cristal·lització.

Imatges de la gota D6 de la placa RS05 corresponent al domini Box BS i l'oligonucleòtid AATAAATTTATT.

La gota amb els millors cristalls d'aquest tipus es mostra a la fig. 4.63. Aquests cristalls van créixer a partir del precipitat amorf després de dos mesos d'haver realitzat la gota. La concentració final del tampó de cristal·lització a la gota era de 25 mM Tris-HCl pH=7,5 i 3% PEG 3350 amb 30% PEG 3350 al pou. És interessant que només varia en un 1% la concentració de PEG a la gota en comparació a la gota RS05-D6 (v. la fig. 4.62). Es van congelar 10 cristalls corresponents a aquesta gota sense i amb crioprotecció. En els seus diagrames de difracció

s'observen taques ben definides però fins a 10 Å en el millor dels casos. En una de les imatges de difracció s'arriben a veure les bandes corresponents a l'*stacking* dels parells de bases del DNA, confirmant així la presència de DNA en aquest tipus de cristalls.

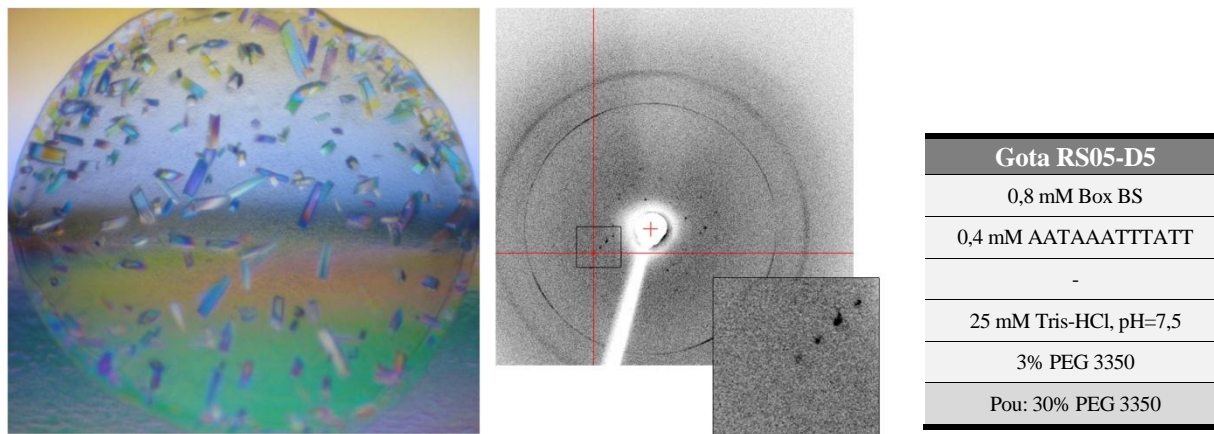


Fig. 4.63. Imatge, condicions de cristal·lització i diagrama de difracció d'un cristall corresponent al domini Box BS i l'oligonucleòtid AATAAATTTATT (gota D5 de la placa RS05, cristall RS05-D5-5).

Cristall difractat a la plataforma cristal·logràfica de la *University of Colorado*: angle d'oscil·lació 0,5°, temps d'exposició 600 s, distància al detector 200 mm i longitud d'ona 1,54 Å. Crioprotecció: 33% PEG 3350 i 10% glicerol.

A la fig. 4.64 i la fig. 4.65 observem dos exemples més de cristalls obtinguts a partir de precipitat. En el primer cas els cristalls presentaven forma de rombe mentre que en el segon forma de cilindre. La morfologia dels cristalls obtinguts sembla estar relacionada amb el pH, obtenint-se varetes a pH 7,5; rombes a 8 i cilindres a 6 o 6,5. En tots dos casos (rombes i cilindres) la difracció dels cristalls, no obstant, presentava molt poques taques.

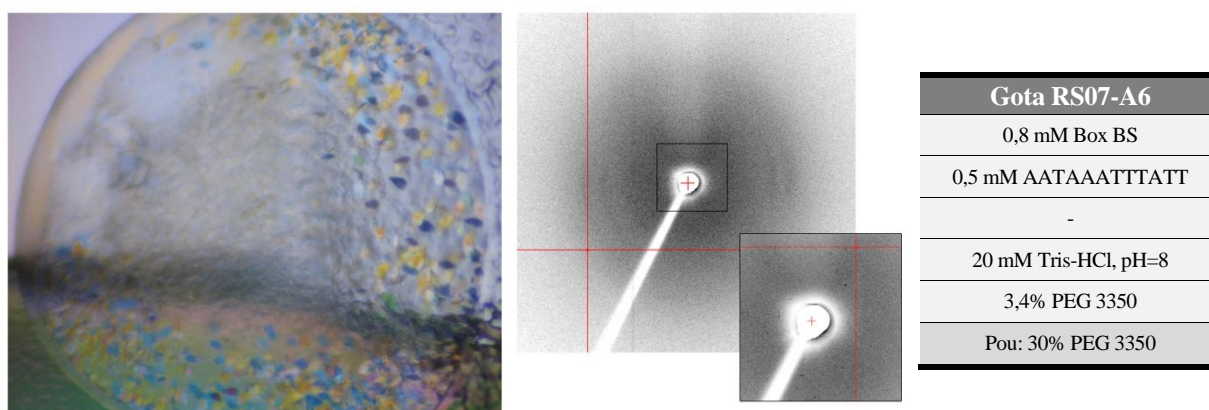


Fig. 4.64. Imatge, condicions de cristal·lització i diagrama de difracció d'un cristall corresponent al domini Box BS i l'oligonucleòtid AATAAATTTATT (gota A6 de la placa RS07, cristall RS07-A6-1).

Cristall difractat a la plataforma cristal·logràfica de la *University of Colorado*: angle d'oscil·lació 0,5°, temps d'exposició 600 s, distància al detector 200 mm i longitud d'ona 1,54 Å.

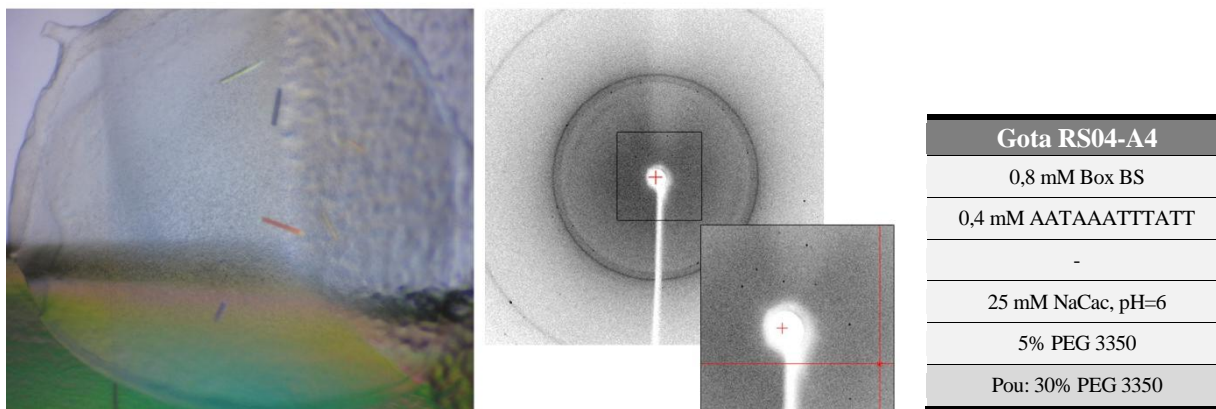


Fig. 4.65. Imatge, condicions de cristal·lització i diagrama de difracció d'un cristall corresponent al domini Box BS i l'oligonucleòtid AATAAATTTATT (gota A4 de la placa RS04, cristall RS04-A4-2).

Cristall difractat a la plataforma cristal·logràfica de la *University of Colorado*: angle d'oscil·lació $0,5^\circ$, temps d'exposició 600 s, distància al detector 200 mm i longitud d'ona $1,54 \text{ \AA}$.

A Barcelona es van voler reproduir les condicions de cristal·lització per tal d'obtenir el mateix tipus de cristalls i veure si s'aconseguia que algun d'ells difractés millor. Es van realitzar gotes sense sals amb Tris-HCl pH=7,5 (condició de pH amb millors resultats) i diferents concentracions de PEG 3350 al pou i a la gota. Es van provar dos concentracions de PEG al pou: al 15% i al 30%. La primera per a obtenir agulles fines i la segona per a aconseguir inicialment un precipitat del que podrien créixer cristalls vareta, com ja havíem vist. En tots els casos es varen obtenir agulles. Aquest resultat no és sorprenent ja que havíem canviat d'altres paràmetres importants per a la cristal·lització (passar de gota assentada a penjada i d'un volum de gota de $1,5 \mu\text{l}$ a $3 \mu\text{l}$). Malgrat no haver obtingut els cristalls a partir del precipitat d'aquestes condicions, les agulles formades (tant amb una concentració inicial de PEG del 15 com del 30%) eren una mica més gruixudes que les obtingudes amb anterioritat i presentaven una gran birefringència. A la fig. 4.66 podem observar aquest tipus d'agulles llargues i fines. El diagrama de difracció d'un cristall d'aquesta gota presentava poques taques ordenades a baixa resolució (v. fig. 4.66) com passava amb les agulles obtingudes a Colorado. De la mateixa manera observem les bandes de l'*stacking* del DNA.

Es van realitzar unes altres proves seguint aquesta línia d'experiments amb una menor concentració de complex ($0,4 \text{ mM}$ proteïna i $0,2 \text{ mM}$ de proteïna) però es van obtenir agulles mal formades. També es va intentar fer precipitar les gotes realitzades per a l'obtenció de cristalls de precipitat augmentant la concentració de PEG, però només es va obtenir precipitat cristal·lí que no va millorar.

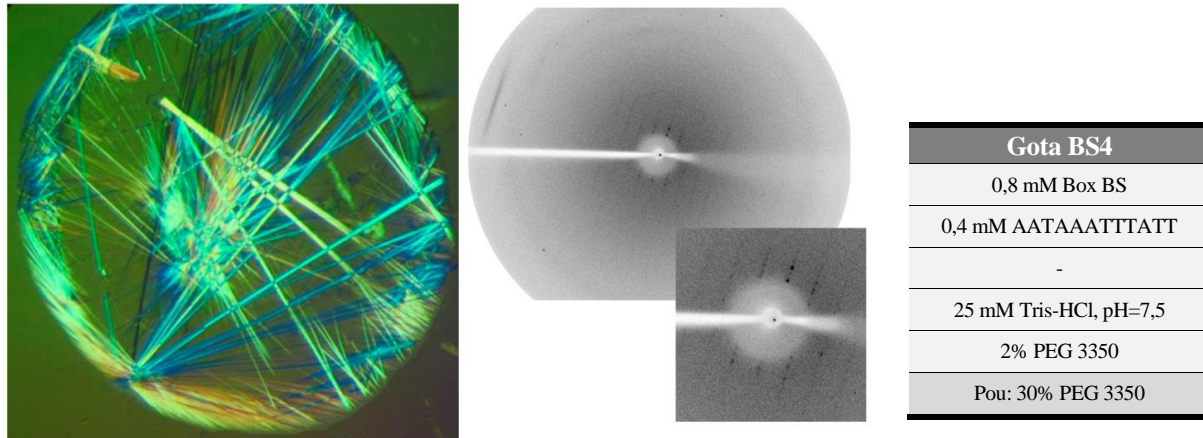


Fig. 4.66. Imatge, condicions de cristal·lització i diagrama de difracció d'un cristall corresponents al domini Box BS i l'oligonucleòtid AATAAATTTATT (gota BS4, cristall BS4-1).

Cristall difractat a la Plataforma Automatitzada de Cristal·lografia del PCB: angle d'oscil·lació 5°, temps d'exposició 600 s, distància al detector 300 mm i longitud d'ona 1,542 Å.

La segona aproximació per a l'optimització de les agulles obtingudes amb PEG i sense sals va ser realitzar uns assajos aplicant una matriu amb diferents sals (les més comunes en la cristal·lització de macromolècules) a una concentració de 5 mM i 20 mM i mantenint constant un 4% de PEG 3350 a la gota i un 30% al pou. Es van provar en aquest cas gotes assegudes i d'1,5 µl per aproximar més a les condicions realitzades a Colorado. En la majoria de les gotes es van obtenir les típiques agulles llargues i fines, en tots els casos massa fines com per a poder ser difractades. En tres de les gotes es van obtenir uns cristalls líquids en forma de varetes curtes amb puntes arrodonides (condicions amb 20 mM de MgCl₂, MgAc i CaCl₂). A més a més, en una de les gotes amb agulles va créixer un cristallet tipus vareta curta molt birefringent a la vora de la gota. A la fig. 4.67 s'assenyala amb una fletxa un cristallet d'aquest tipus que va aparèixer després a la gota. El cristall, que era més gran i presentava millor forma que aquest, va ser congelat en dos fragments. El segon dels fragments va presentar una bona difracció al sincrotró. Els diagrames de difracció tenien una resolució de fins 3,3 Å amb una única direcció de l'*stacking* de les bases del DNA. A l'apartat 4.3 s'analitzen amb més detall els diagrames de difracció obtinguts per a aquest cristall. Aquesta gota contenia la sal d'acetat d'amoni a 20 mM i va ser l'únic cas de gota amb agulles en la qual va aparèixer una vareta a la perifèria.

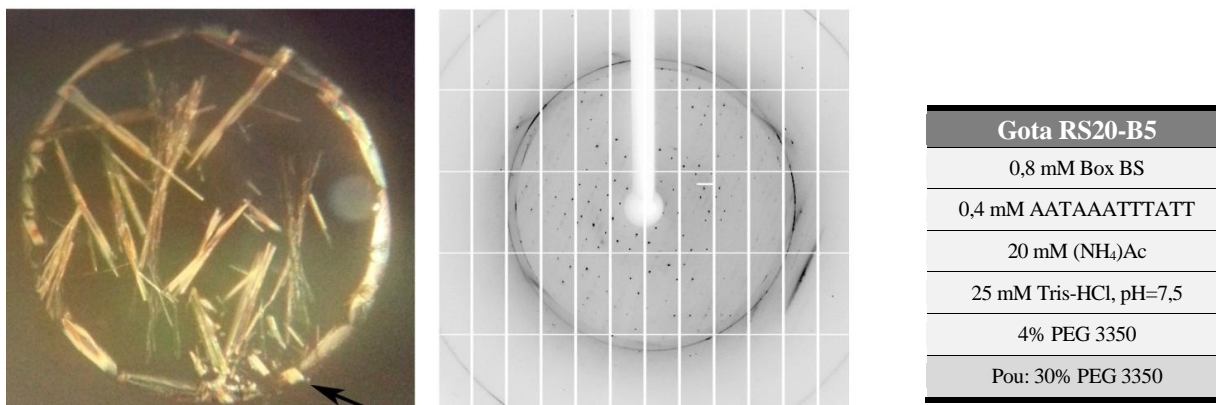


Fig. 4.67. Imatge, condicions de cristal·lització i diagrama de difracció d'un cristall corresponent al domini Box BS i l'oligonucleòtid AATAAATTTATT (gota B5 de la placa RS20, cristall RS20-B5-2).

Cristall difractat al sincrotró ALBA (línia Xaloc): angle d'oscil·lació 2°, temps d'exposició 2 s, distància al detector 550 mm i longitud d'ona 0,9795 Å.

Paral·lelament es van realitzar uns assajos de cristal·lització amb aquest oligonucleòtid (AATAAATTTATT) i alguns dels tampons Natrix HR2. Es van obtenir unes agulles molt fines amb el número 10 (2,5 mM MgSO₄/pH=6/PEG 4000) i unes varetes que presentaven un bon aspecte amb el 15 (20 mM MgCl₂/pH=6/MPD). Van ser congelades dos d'aquestes varetes, corresponents a la càpsula BS41, però les seves difraccions no van ser bones (només es veia l'*stacking* i cap taca).

Per a millorar aquestes varetes es van dur a terme unes altres proves amb aquesta combinació MgCl₂/MPD a diferents valors de pH. Es van fer assajos a 20 i 100 mM de MgCl₂, però només van créixer varetes prou bones en les gotes que contenien 100 mM MgCl₂. Aquestes varetes eren força grans (~500 x 50 µm) i birefringents però presentaven uns diagrames típics de només DNA (v. fig. 4.68). Dos cristalls més d'aquesta i una altra gota confirmaven el mateix resultat.

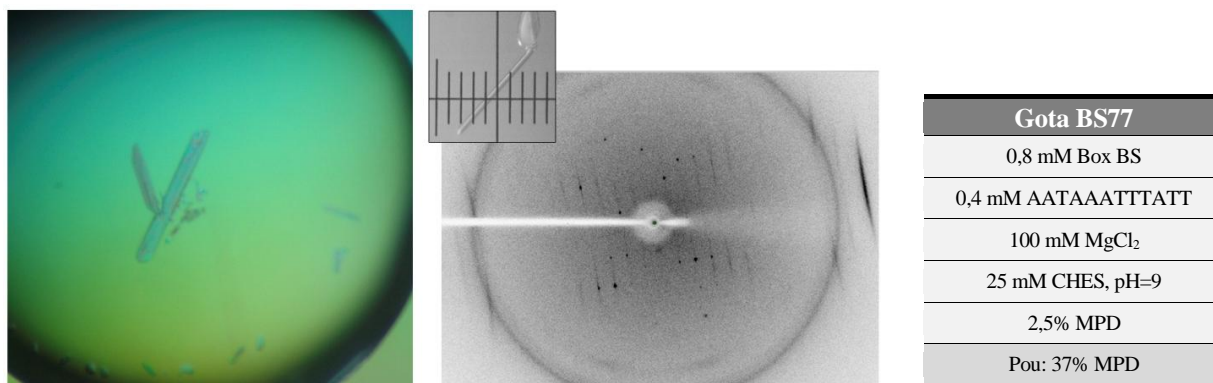


Fig. 4.68. Imatge, condicions de cristal·lització i diagrama de difracció d'un cristall corresponent al domini Box BS i l'oligonucleòtid AATAAATTTATT (gota BS77, cristall BS77-1).

La imatge de la gota es va realitzar un cop pescada una vareta gran. Cristall difractat a la Plataforma Automatitzada de Cristal·lografia del PCB: angle d'oscil·lació 5°, temps d'exposició 240 s, distància al detector 200 mm i longitud d'ona 1,542 Å.

c) *HMGB1 Box BS - variants AATAAATTTATT*

Les proves de cristal·lització de les variants de l'AATAAATTTATT amb el Box BS en les condicions sense sal i PEG 3350 van donar precipitat cristal·lí amb l'ATAAATTTATTA (v1) i separació de fases amb l'AATAAATTTATTT (v2).

Es van realitzar les mateixes proves de cristal·lització amb la combinació MgCl₂/MPD que les realitzades per a l'oligonucleòtid original amb totes dues variants. En el cas de l'oligonucleòtid ATAAATTTATTA (v1) es va obtenir precipitat cristal·lí a 20 mM MgCl₂ i cristalls mal formats a 100 mM. En canvi per a l'AATAAATTTATTT (v2) es va formar separació de fases en les gotes a 20 mM i varetes a 100 mM (similars a la de la fig. 4.69). Es va congelar i difractar una única vareta molt gran d'aquest tipus com observem a la imatge del cristall dins del *loop* (fig. 4.69). El diagrama de difracció sembla també un diagrama típic d'únicament DNA.

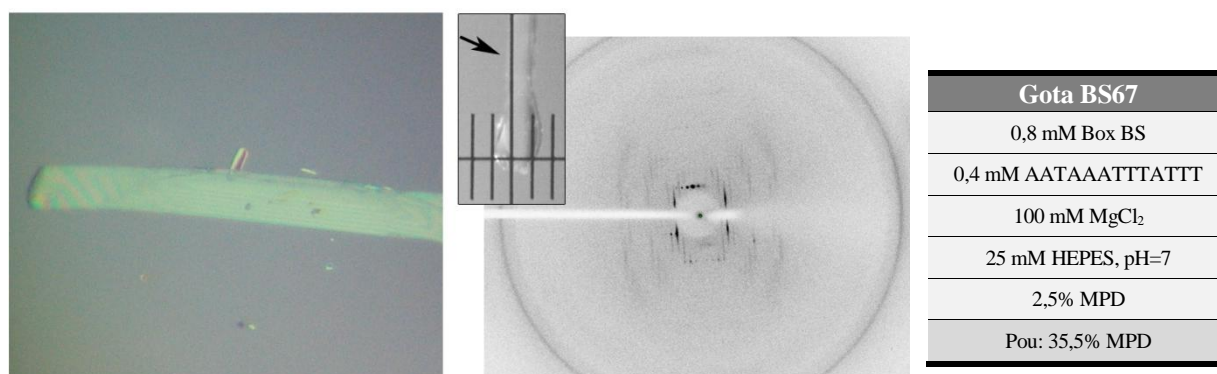


Fig. 4.69. Imatge, condicions de cristal·lització i diagrama de difracció d'un cristall corresponent al domini Box AS i l'oligonucleòtid AATAAATTTATTT (gota BS67, cristall BS67-1).

Es va congelar un cristall similar al de la imatge però que presentava millor forma. Cristall difractat a la Plataforma Automatitzada de Cristal·lografia del PCB: angle d'oscil·lació 5°, temps d'exposició 240 s, distància al detector 200 mm i longitud d'ona 1,542 Å.

d) *HMGB1 Box BS - ATATCGATAT*

Tot i haver obtingut el millor resultat amb aquest oligonucleòtid i la Box AS (v. 4.2.2) no s'ha obtingut ningun resultat positiu en la cristal·lització amb la proteïna Box BS. S'han realitzat gotes sense sal (amb PEG 3350 i MPD) i s'han provat diferents condicions dels tampons Matrix però només s'ha observat separació de fases.

e) *HMGB1 Box BS - AATAATCGCGATTATT*

En aquest cas s'han obtingut només dos resultats positius de les 12 gotes realitzades amb condicions de cristal·lització força diferents. S'han observat varetes petites en les gotes realitzades amb el Natrix 15 (v. fig. 4.70) i el 16. Degut a la seva reduïda mida no es va congelar cap d'aquests cristalls.

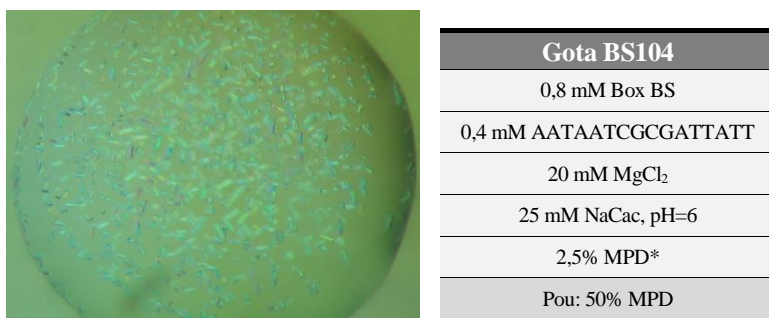


Fig. 4.70. Imatge i condicions de cristal·lització de la gota BS104 corresponent a la proteïna Box BS i l'oligonucleòtid AATAATCGCGATTATT.

*Corresponent al tampó Natrix n°15 (v. annex D).

Electroforesi de cristalls

S'han analitzat electroforèticament els cristalls més representatius obtinguts de la proteïna Box BS per a comprovar si hi ha presència de proteïna. La visualització d'una banda corresponent a la proteïna ens dóna la idea de que hi havia proteïna al cristall però no sobre si es troba de forma ordenada i/o unida al DNA.

A la fig. 4.71 es pot observar el gel de poliacrilamida-SDS corresponent a diferents cristalls de la Box BS. En tots els cristalls s'observa una banda intensa corresponent al fragment Box BS (cristalls BS1, BS34, BS41-2, BS67-1 i RS20-B5-1) excepte en els cristalls BS74 i BS77-1 que presenten una banda més dèbil. En aquests dos cristalls que contenen l'oligonucleòtid AATAAATTTATT i van créixer en unes condicions de 100 mM MgCl₂ ja es sospitava que només contienien DNA degut al diagrama de difracció corresponent al cristall BS77-1 (v. fig. 4.68).

Un resultat inesperat és el trobat per al cristall BS67-1 que presenta una banda molt intensa en aquest gel i també té un diagrama típic de només DNA i ha crescut a 100 mM de sal de magnesi (v. fig. 4.69). En aquest cas l'oligonucleòtid és la variant v2 (AATAAATTTATTT) i creiem que potser la proteïna es troba de forma desordenada i per això observem aquest tipus de diagrama de difracció.

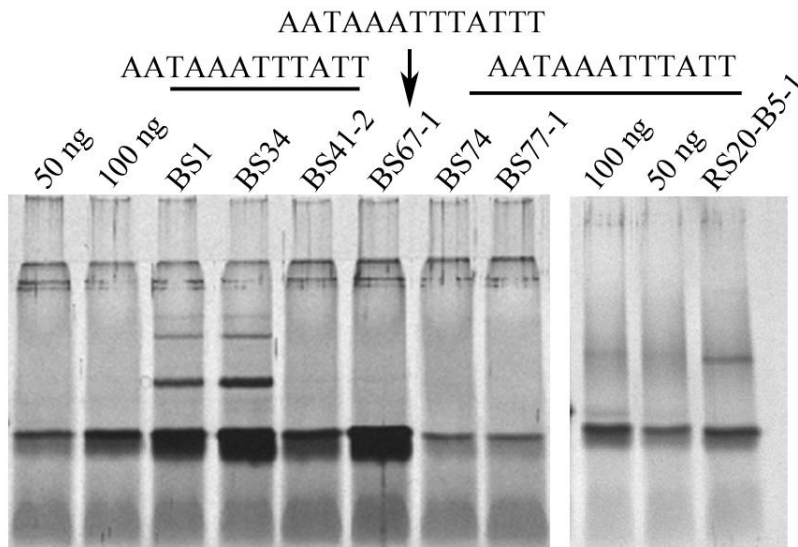


Fig. 4.71. Gel poliacrilamida-SDS d'alguns cristalls de la proteïna Box BS.

S'han carregat al gel controls de proteïna a 50 i 100 ng.

Cristalls amb l'oligonucleòtid AATAAATTTATT: BS1, agulla (semblant a fig. 4.66); BS34, agulla (semblant a fig. 4.66); BS41-2, vareta; BS74; fragment de vareta gran (semblant a fig. 4.68); BS77-1, vareta llarga i fina (v. fig. 4.68) i RS20-B5-1, cristall vareta curta (v. fig. 4.67).

Cristalls amb l'oligonucleòtid AATAAATTTATTT: BS67-1, vareta (v. fig. 4.69).

Conclusions

La cristal·lització del Box BS amb l'oligonucleòtid AATAAATTTATT va presentar uns cristalls molt prometedors amb les condicions amb PEG. Per una banda es van obtenir unes agulles fines, llargues i molt brillants mentre que per una altra es van obtenir uns cristalls amb molt bona forma que van créixer a partir de precipitat en les gotes. Cap dels dos tipus de cristalls va presentar una difracció suficientment bona com per a resoldre la seva estructura, en tots dos casos s'observaven taques ordenades a baixa resolució. En una de les gotes, corresponent a les optimitzacions amb diferents sals, a més a més de les agulles van aparèixer unes varetes curtes. Un d'aquests cristalls va mostrar una difracció fins a 3,3 Å, sent així la millor i única difracció recollida per al Box BS d'aquest treball.

4.2.4. Complexos HMGB1 Box ABS-DNA

Amb l'objectiu de cristal·litzar el fragment Box ABS, que conté els dos HMG-box A i B i el *linker* bàsic que els uneix (v. fig. 4.12), s'han utilitzat uns oligonucleòtids amb unes seqüències més llargues. Concretament hem fet servir l'oligonucleòtid GAATTAACGTACGTTAATTCG (amb 21 pb) i el CGCGTTTTTCGCGAAAAACGCGG (amb 23 pb), anomenats també en aquest treball sub21 i sub23 per simplificar. No sabem amb exactitud quina longitud de seqüència necessita la forma amb els dos dominis per a unir-se al DNA i poder cristal·litzar. Se sap, però, que quan un domini HMG-box s'uneix al DNA produeix un doblegament d'aquest que implica uns sis parells de bases. I a més a més hem de tenir en compte que en les dues estructures que s'han resolt contenint dos dominis HMG-box, la quimera SRY-BoxB i l'HMGB del mitocondri (TFAM), s'han utilitzat unes seqüències de 16 i 22 pb respectivament. Per tant s'ha considerat que aquestes seqüències de 21 i 23 pb presentaven unes longituds adequades per a la unió del fragment Box ABS.

Cal comentar que la cristal·lització de fragments tan llargs de DNA juntament amb una proteïna de 170 aminoàcids és un repte força ambiciós i presenta moltes dificultats ja només tenint en compte la complexitat de l'obtenció d'una estructura ordenada d'aquestes macromolècules.

Prèviament a la realització dels assajos cristal·logràfics es van preparar els complexos entre els oligonucleòtids sub21 i sub23 amb el fragment Box ABS per a carregar en un gel EMSA. Observem clarament la formació dels complexos proteïna-DNA a la fig. 4.72. Podem veure la formació del complex a menor relació molar que en els casos dels HMG-boxes A i B per separat ja que el fragment amb els dos dominis (Box A+ Box B) presenta major afinitat pel DNA [36]. A més a més en aquests gels podem veure la formació d'un segon complex a major concentració de proteïna, segurament format per dues molècules de proteïna i una d'oligonucleòtid.

HMGB1 Box ABS

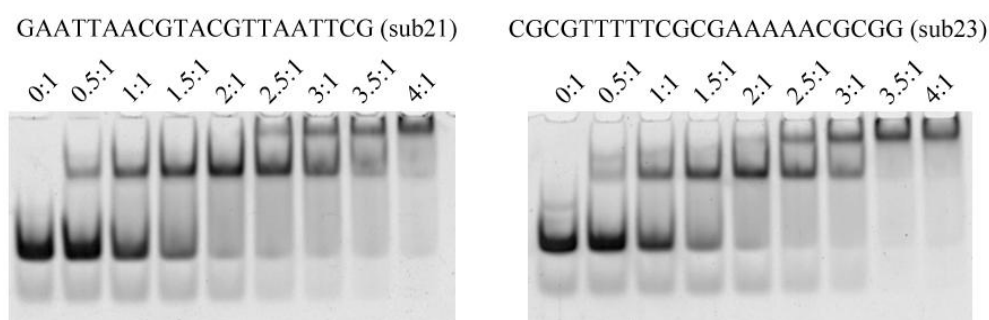


Fig. 4.72. Assaigs EMSA per a la formació de complexos entre la Box ABS i els oligonucleòtids sub21 i sub23. En els carrils s'indica la relació molar proteïna:DNA corresponent a cada incubació.

Resum de les condicions de cristal·lització

A la taula 4.6 es mostra un resum dels assajos cristal·logràfics realitzats amb la proteïna Box ABS. En primer lloc s'han fet diferents proves sense addició de sals amb l'oligonucleòtid sub23 a la Universitat de Colorado. Ja que aquestes proves no van presentar cap resultat positiu es van fer unes quantes proves en condicions més diferents, en aquest cas a Barcelona, amb tots dos oligonucleòtids. Es van provar 8 tampons Natrix HR2 diferents, i tampons sense sal amb PEG i MPD. També es va dur a terme la variació de la relació tampó:complex (indicat amb un asterisc a la taula) per assajar així altres concentracions de complex i precipitant a les gotes.

En cap dels casos, però, es va obtenir un cristall apte per a ser congelat i difractat, com s'explica més detalladament a continuació.

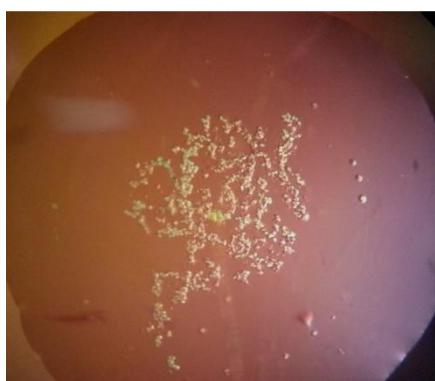
Taula 4.6. Resum dels assajos cristal·logràfics per a la proteïna HMGB1 Box ABS.

Proteïna	Oligonucleòtid	Ràtio P:D (mM)	Condicions			nº gotes	Cristalls	Difracció
			Sal	pH	Precip.			
HMGB1 Box ABS Colorado	CGCGTTTTTCG-CGAAAAACGCGG (sub23)	0,8;0,4	-	6; 6,5; 7; 7,5; 8	PEG 3350; MPD	60	-	-
HMGB1 Box ABS Barcelona	GAATTAACGT-ACGTAAATTCG (sub21)	0,4;0,2	Natrix (nº 7, 10, 15, 16, 25, 34, 42, 43)			8	Precipitat cristal·lí	-
		0,4;0,2	-	6; 6,5; 7; 7,5; 8	PEG 3350	5	-	-
		0,4;0,2*	-	7,5	MPD; PEG 3350	6	-	-
		0,4;0,2*	Natrix (nº 15)			3	Agregats cristal·lins	-
	CGCGTTTTTCG-CGAAAAACGCGG (sub23)	0,4;0,2	Natrix (nº 7, 10, 15, 16, 25, 34, 42, 43)			8	-	-
		0,4;0,2	-	6; 6,5; 7; 7,5; 8	PEG 3350	5	-	-
		0,4;0,2*	-	7,5	MPD; PEG 3350	6	-	-
		0,4;0,2*	Natrix (nº 15)			3	-	-

*Es van realitzar gotes barrejant diferents proporcions del tampó i del complex [(2:1)(1:1)(1:2)]; a la taula només es mostren les concentracions finals per a les barreges (1:1). La ràtio P:D és la ràtio molar proteïna:DNA. Tots els assajos s'han realitzat a 4°C.

*Resultats de cristal·lització**a) HMGB1 Box ABS - GAATTAACGTACGTAAATTCG (sub21)*

Les proves de cristal·lització amb el fragment Box ABS i l'oligonucleòtid sub21 van donar principalment separació de fases en la majoria de les gotes realitzades. No obstant això, vam observar l'aparició d'un precipitat cristal·lí homogeni en les condicions dels Natrix n° 15 i 16 (v. annex D). En la repetició de la gota contenint el Natrix n°15 i en la variació 2:1 (tampó:complex) amb aquest mateix tampó es van obtenir uns agregats cristal·lins com el que s'observa a la fig. 4.73. Caldria optimitzar aquest resultat per tal d'obtenir uns cristalls ben formats en lloc d'aquests agregats o precipitats cristal·lins.



Gota ABS34
0,4 mM Box ABS
0,2 mM sub21 GAATTAACGTACGTAAATTCG
20 mM MgCl ₂
25 mM NaCac, pH=6
2,5% MPD*
Pou: 50% MPD

Fig. 4.73. Imatge i condicions de cristal·lització de la gota ABS34 corresponent a la proteïna Box ABS i a l'oligonucleòtid sub21.

*Corresponent al tampó Natrix n°15 (v. annex D).

b) HMGB1 Box ABS - CGCGTTTTTCGCGAAAAACGCGG (sub23)

No es va obtenir cap resultat positiu en els assajos realitzats amb l'oligonucleòtid sub23. Ni tan sols es va observar cap tipus de precipitat, només separació de fases en totes les gotes preparades.

Conclusions

La cristal·lització del fragment Box ABS, contenint els dos HMG-box de l'HMGB1 i el *linker* entre ells, no va presentar cap condició de cristal·lització òptima en els assajos cristal·logràfics realitzats. Només es van obtenir precipitats o agregats cristal·lins.

4.2.5. Complexos NHP6A-DNA

Es va abordar la cristal·lització de la proteïna HMGB del llevat *Saccharomyces cerevisiae*, l'NHP6A, de la mateixa manera que per als dominis HMG-box per separat (Box AS i Box BS). Així, hem fet servir per a la cristal·lització de l'NHP6A, la qual conté un únic HMG-box, els mateixos 6 oligonucleòtids que per als dominis Box AS i BS: GTTATTAATAAC, AATAAATTTATT, ATATCGATAT AATAATCGCGATTATT i les dues variants de l'AATAAATTTATT (ATAAATTTATTA i AATAAATTTATTT). Degut a les dificultats per a cristal·litzar aquesta proteïna amb aquests oligonucleòtids es van provar dos oligonucleòtids més: ATATATATATATAT i CATATATATG.

Com en els casos anteriors, la formació dels complexos proteïna-DNA s'ha confirmat mitjançant electroforesis en gels EMSA. A la fig. 4.74 podem observar els assajos de canvi en la mobilitat electroforètica per als diferents oligonucleòtids utilitzats en cristal·lografia i concentracions creixents de la proteïna NHP6A. La constant de dissociació de l'NHP6A és menor que en el cas de l'HMG-box per separat, sent de l'ordre de nanomolar per a la primera ($K_D=10$ nM, [44]) i de micromolar per a les segones [36]. Per aquest motiu podem observar com en aquest cas la formació de complex amb proteïna NHP6A es produeix abans (menor relació proteïna:DNA) que en el cas de les HMG-box de l'HMGB1 (Box AS i BS) (v. també fig. 4.45 i fig. 4.58). A més a més podem apreciar la formació d'un segon complex a l'augmentar la concentració de proteïna, tal i com passava amb el fragment amb els dos dominis (Box ABS) (v. fig. 4.72). Aquest segon complex està format per dues molècules de proteïna per una de DNA com es va comprovar per anàlisi quantitatiu de les bandes de proteïna obtingudes en la tinció amb plata del gel (no mostrat).

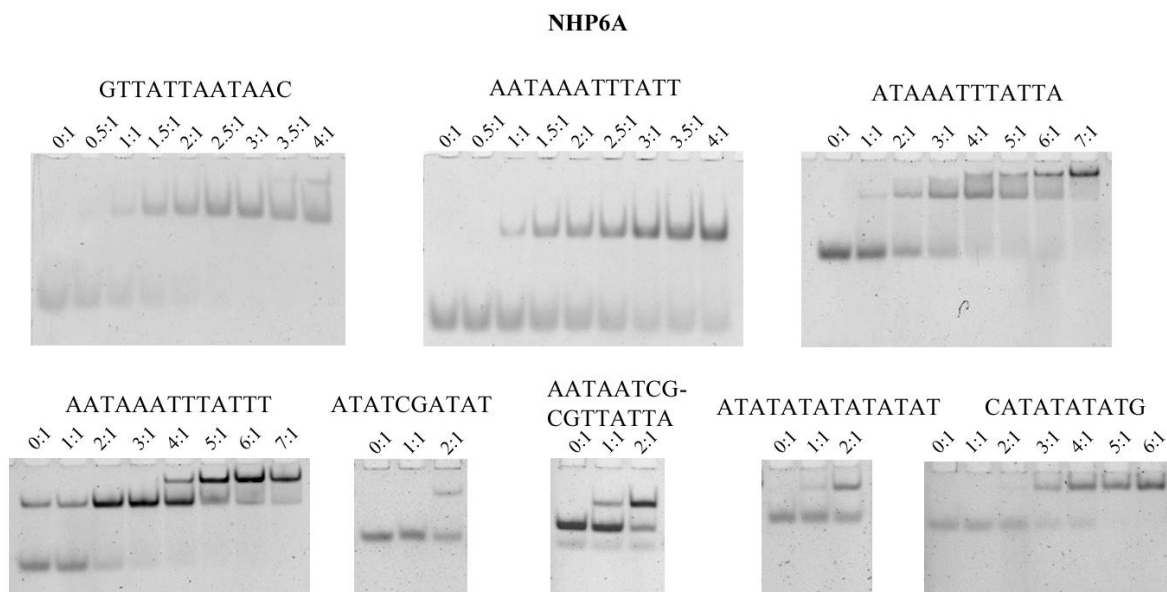


Fig. 4.74. Assaigs EMSA per a la formació de complexos entre l'NHP6A i els diferents oligonucleòtids.

En els carrils s'indica la relació molar proteïna:DNA corresponent a cada incubació.

Resum de les condicions de cristal·lització

A la taula 4.7 es resumeixen els assajos cristal·logràfics realitzats per a la proteïna NHP6A.

Taula 4.7. Resum dels assajos cristal·logràfics per a la proteïna NHP6A.

Proteïna	Oligonucleòtid	Ràtio P:D (mM)	Condicions			nº gotes	Cristalls	Difracció
			Sal	pH	Precip.			
NHP6A Colorado	GTTATTAATAAC	0,6;0,2	-	6; 6,5; 7; 7,5; 8	PEG3350 MPD	60	-	-
	AATAAATTTATT	0,6;0,2	-	6; 6,5; 7; 7,5; 8	MPD	30	-	-
NHP6A Barcelona	AATAAATTTATT	0,8;0,4	Matriu PAC9mod i PAC10*			384	Cristalls petits	No difracten
		0,8;0,2	-	6; 6,5; 7; 7,5; 8	PEG 3350	5	-	-
		0,8;0,2	MgCl ₂	9	PEG 8000	49	Agregats cristal·lins; Mini-varetes	No difracten
	ATAAATTTATTA (v1)	0,8;0,2	-	6; 6,5; 7; 7,5; 8	PEG3350	5	-	-
	AATAAATTTATTT (v2)	0,8;0,2	-	6; 6,5; 7; 7,5; 8	PEG3350	5	-	-
	ATATCGATAT	0,8;0,2	Natrix (nº 15)			1	-	-
		0,8;0,2	-; MgCl ₂	9	PEG 8000	2	Agregats cristal·lins	-
		0,8;0,2	-	7	PEG 3350	1	-	-
	AATAATCGCGATTATT	0,8;0,2	Natrix (nº 15)			1	-	-
		0,8;0,2	-; MgCl ₂	9	PEG 8000	2	Agregats cristal·lins	-
		0,8;0,2	-	7	PEG 3350	1	-	-
	ATATATATATATAT	0,8;0,2	Natrix (nº 15)			1	-	-
		0,8;0,2	-; MgCl ₂	9	PEG 8000	2	Làmines no homogènies	-
		0,8;0,2	-	7	PEG 3350	1	-	-
	CATATATATG	0,8;0,2	Natrix (nº 15)			1	-	-
0,8;0,2		-; MgCl ₂	9	PEG 8000	2	Cristalls rombe petits	-	
0,8;0,2		-	7	PEG 3350	1	-	-	

*Gotes preparades automàticament amb robot. Els oligonucleòtids ATAAATTTATTA (v1) i AATAAATTTATTT (v2) són variants de l'oligonucleòtid AATAAATTTATT, presenten igual seqüència nucleotídica a excepció dels extrems (v. fig. 4.44). La ràtio P:D és la ràtio molar proteïna:DNA. Tots els assajos s'han realitzat a 4°C.

Les proves cristal·logràfiques de la proteïna NHP6A amb els oligonucleòtids GTTATTAATAAC i AATAAATTTATT (i les seves variants) en les condicions sense sals no va donar cap resultat positiu, només es va observar separació de fases en el millor dels casos. Per aquest motiu es va decidir portar a terme la cristal·lització del complex amb les matrius PAC9mod i PAC10 (disponibles a la Plataforma Automatitzada de Cristal·lografia del Parc Científic de Barcelona), i així veure el resultat en 192 condicions de cristal·lització diferents. Es van provar dues relacions molars de proteïna:DNA (2:1 i 4:1) amb l'oligonucleòtid AATAAATTTATT. A més a més prèviament a la realització d'aquestes proves es va realitzar un test de precristal·lització (Hampton Research) que va confirmar que la concentració de treball del complex era l'adequada (obtenció de precipitats en aquest test).

De totes les possibles condicions de cristal·lització d'aquestes dues matrius (v. annex D), únicament van créixer cristalls en tres de les gotes, la resta estaven clares i algunes presentaven separació de fases. Aquestes tres condicions tenien en comú el tampó (CHES pH=9) i la sal ($MgCl_2$) i dos d'elles el precipitant (PEG 8000). Per aquesta raó es van realitzar diferents assajos amb variacions d'aquestes condicions per a l'optimització d'aquests cristalls (v. taula 4.7).

Finalment es va provar la cristal·lització d'aquesta proteïna amb dos altres oligonucleòtids (ATATATATATATAT i CATATATATG) a més a més dels oligonucleòtids ATATCGATAT i AATAATCGCGATTATT. Amb aquests quatre oligonucleòtids es van fer només unes poques proves de cristal·lització amb les condicions amb més èxit amb aquesta i les altres proteïnes HMG-box (Natrix n°15; PEG 3350 sense sal; $MgCl_2$ i CHES pH=9).

A continuació s'explica amb més detall aquests resultats de cristal·lització.

Resultats de cristal·lització

a) NHP6A - GTTATTAATAAC

No es va obtenir cap cristall en la cristal·lització de l'oligonucleòtid GTTATTAATAAC amb aquesta proteïna. Es van provar diferents concentracions de PEG 3350 i MPD sense cap sal afegida al tampó. Només es va observar separació de fases en algunes de les gotes.

b) NHP6A - AATAAATTTATT

En els assajos sense sals realitzats a Colorado amb el precipitant PEG 3350 i a Barcelona amb MPD no es va obtenir cap resultat positiu, únicament separació de fases.

Per a explorar un espai de condicions més ampli es van realitzar unes plaques d'alt rendiment amb les 192 condicions de les matrius PAC9m i PAC10 (v. annex D) amb el robot de Cartesian Technologies a la Plataforma Automatitzada de Cristal·lografia del Parc Científic de Barcelona. De totes les proves realitzades només es van obtenir cristalls en tres gotes. Dos d'elles corresponien a la mateixa condició (MgCl₂, CHES pH=9 i PEG 8000) a les dues relacions molars (2:1) i (4:1). Els cristalls formats eren petits i en forma de varetes. Es va difractar un d'aquests cristalls amb crioprotecció, corresponent a la gota RS23-H12-d1 (v. fig. 4.75). El seu diagrama de difracció, però, no presentava cap taca possiblement per la seva mida (s'ha de tenir en compte que aquestes gotes són molt petites, de 200 nl). En l'altra condició, RS23-H08-d1 (MgCl₂, CHES pH=9 i PEG MME 500), van créixer uns cristallets en forma de cucs que no van ser pescats (v. fig. 4.76).

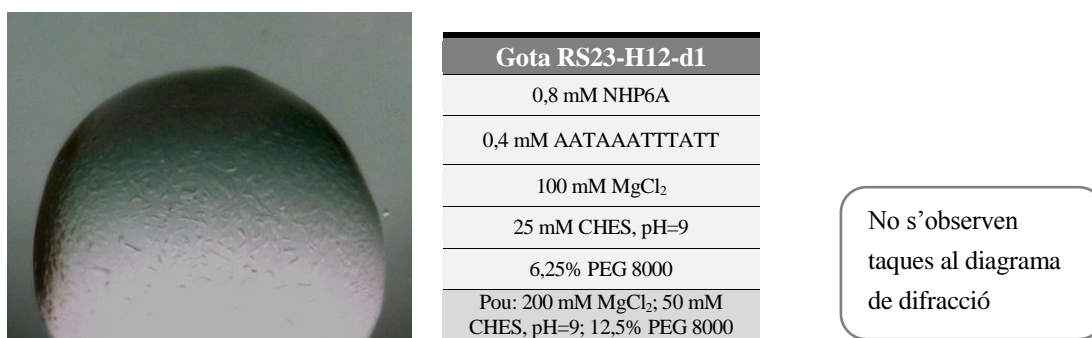


Fig. 4.75. Imatge i condicions de cristal·lització de la gota RS23-H12-d1 corresponent a la proteïna NHP6A i l'oligonucleòtid AATAAATTTATT.

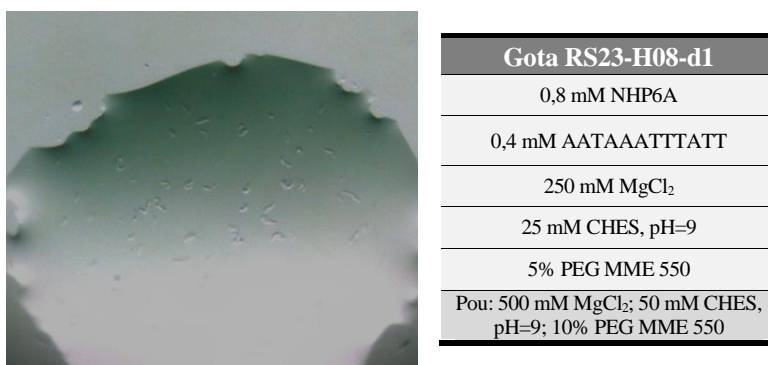


Fig. 4.76. Imatge i condicions de cristal·lització de la gota RS23-H08-d1 corresponent a la proteïna NHP6A i l'oligonucleòtid AATAAATTTATT.

Es va intentar optimitzar aquests cristalls vareta petits, variant lleugerament les condicions de cristal·lització, per una banda en plaques amb gotes assentades (de 0,8 µl de volum) i també en càpsules amb gotes més grans penjades (3 µl de volum). La relació molar va ser de 4:1 ja que s'obtenien menor nombre de cristalls que en la 2:1.

En la placa es van obtenir varetes petites en algunes de les gotes (v. fig. 4.77) que seguien sent de mida molt reduïda. A les càpsules es va observar un precipitat cristal·lí en forma de boles en la majoria de les gotes (v. fig. 4.78) i unes varetes petites en un cas (v. fig. 4.79). En algunes de les proves, però, s'observaven les varetes i les boletes a la vegada.

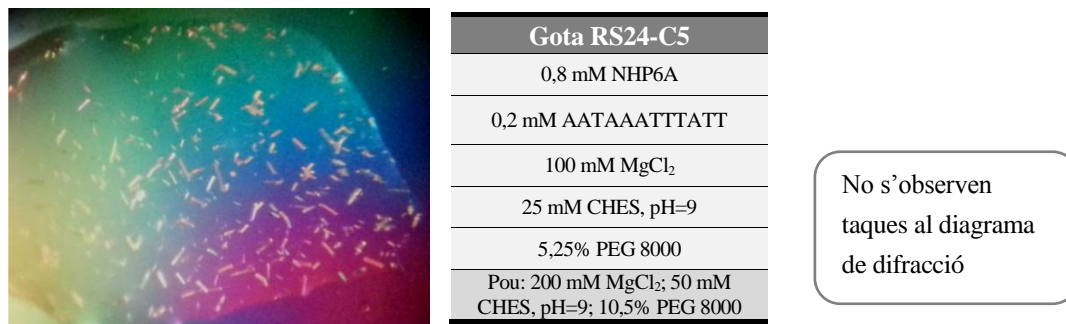


Fig. 4.77. Imatge i condicions de cristal·lització de la gota RS24-C5 corresponent a la proteïna NHP6A i a l'oligonucleòtid AATAAATTTATT.

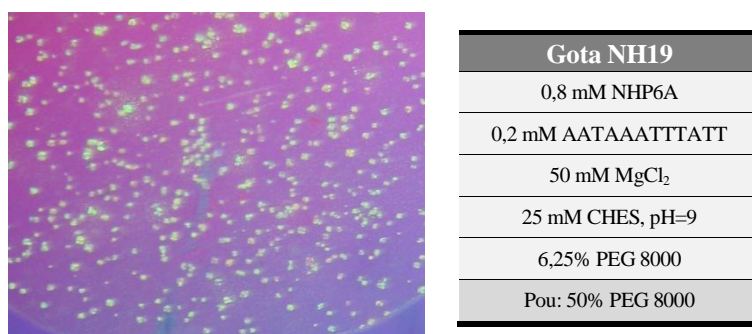


Fig. 4.78. Imatge i condicions de cristal·lització de la gota NH19 corresponent a la proteïna NHP6A i a l'oligonucleòtid AATAAATTTATT.

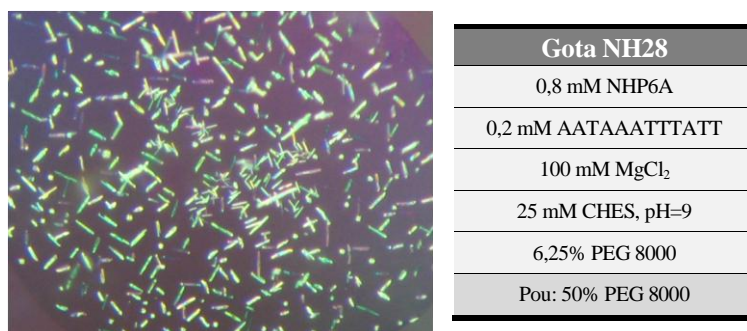


Fig. 4.79. Imatge i condicions de cristal·lització de la gota NH28 corresponent a la proteïna NHP6A i a l'oligonucleòtid AATAAATTTATT.

c) *NHP6A - variants AATAAATTTATT*

En les condicions sense sals provades amb PEG amb les variants de l'oligonucleòtid AATAAATTTATT només es va obtenir separació de fases.

d) *NHP6A – ATATCGATAT, AATAATCGCGATTATT, ATATATATATATAT, CATATATATG*

Amb aquests quatre oligonucleòtids addicionals només es va observar l'obtenció de cristalls en les gotes que presentaven MgCl₂, CHES pH=9 i el precipitant PEG 8000. En el cas de l'oligonucleòtid ATATCGATAT va créixer el precipitat cristal·lí en forma de boletes així com amb l'oligonucleòtid CATATATATG. Per a l'AATAATCGCGATTATT es van obtenir unes làmines no homogènies (v. fig. 4.80) i per a l'ATATATATATATATAT uns cristallets en forma de rombe plans (v. fig. 4.81), cap d'ells era òptim per a congelar i ser difractat.

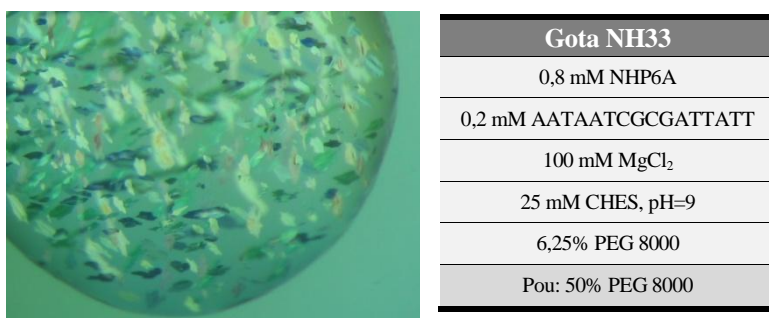


Fig. 4.80. Imatge i condicions de cristal·lització de la gota NH33 corresponent a la proteïna NHP6A i a l'oligonucleòtid AATAATCGCGATTATT.

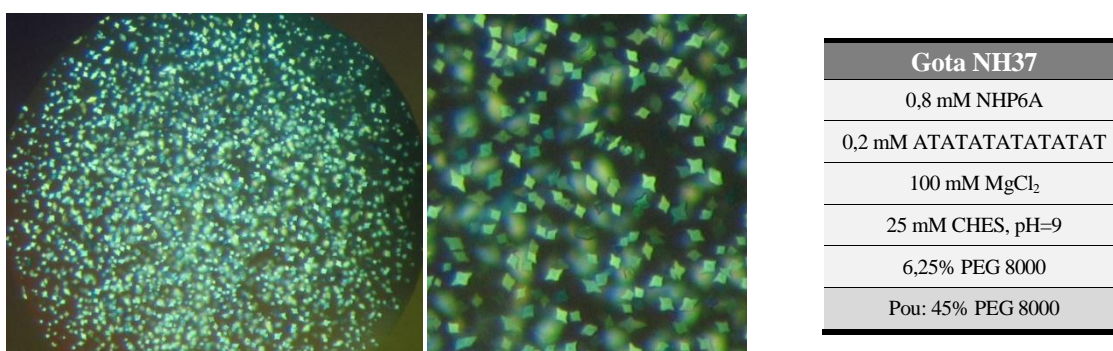


Fig. 4.81. Imatge i condicions de cristal·lització de la gota NH37 corresponent a la proteïna NHP6A i a l'oligonucleòtid ATATATATATATATAT.

Conclusions

Tot i realitzar molts assajos cristal·logràfics amb gran varietat de condicions de cristal·lització no es van obtenir resultats positius amb la proteïna NHP6A. Només en els casos amb CHES pH=9 i alta concentració de $MgCl_2$ s'han obtingut cristalls, que possiblement només contindran el DNA. L'oxidació de la metionina de la proteïna NHP6A, la qual interacciona amb el DNA segons dades d'RMN [44], podria ser un impediment per a l'obtenció de cristalls del complex. Inicialment, al realitzar els assajos cristal·logràfics, es va creure que treballant amb un estoc de proteïna amb l'agent reductor DTT era suficientment per a impedir l'oxidació de la metionina. Però pot ser que el seu efecte es vegi disminuït amb el temps. Per tant una possible optimització de les gotes realitzades amb l'NHP6A en aquest treball podria ser l'aplicació de condicions més reductores a les gotes com per exemple substituint el DTT pel TCEP.

4.3. Anàlisi dels diagrames de difracció

A partir dels diagrames de difracció d'un cristall d'una macromolècula podem obtenir el seu mapa de densitat electrònica. Aquest mapa s'interpreta en termes de les seves posicions atòmiques, així que la precisió i fins i tot la fiabilitat de l'anàlisi de l'estructura del cristall depèn del límit de resolució d'aquest. De fet, com ja s'ha donat a entendre, la capacitat d'obtenció de cristalls que difractin amb resolució suficientment alta és un factor limitant en la determinació de l'estructura d'un cristall. Així a nivells de resolució de 3,5 Å no es pot distingir un parell de bases dels seus veïns en una molècula de DNA. A 2 Å de resolució els àtoms individuals no es poden diferenciar però sí endevinar-se la seva forma molecular. Mentre que a resolucions d'1,5 Å (propera a les distàncies d'enllaç) els àtoms es resolten de forma parcial i a majors resolucions els àtoms es veuen amb claredat.

A la taula 4.8 es troben els cristalls amb millors diagrames de difracció, i per tant amb els límits de resolució més alts, de les proves cristal·logràfiques realitzades en aquest treball. Aquests corresponen a 6 cristalls del fragment Box AS i un del Box BS (v. apartat 4.2.2 i 4.2.3 per més detalls). També s'ha inclòs el cristall amb el pèptid *AT-hook* de l'HMGA1 obtingut durant el PFC i analitzat també en el present estudi. S'ha assignat una lletra de la A a la F a les diferents cel·les unitàries calculades com a primera aproximació per als diferents cristalls. Aquestes han estat calculades amb els programes *Denzo/HKL2000* [98] per al cristall 45-2 i amb l'*iMosflm* [100] per a la resta (v. apartat 3.9.4).

En aquest apartat s'analitzaran els diagrames de difracció corresponents als cristalls de la taula 4.8, s'estudiaran els paràmetres de cel·la obtinguts juntament amb el contingut de la unitat asimètrica i es proposaran possibles models d'empaquetament dels complexos. En el cas de la cel·la C, corresponent a un cristall del complex Box AS-ATATCGATAT, es va anar un pas més enllà. Al presentar una major resolució i unes taques millor definides en els diagrames de difracció vam poder arribar a resoldre la seva estructura a nivell atòmic com s'explicarà al següent apartat (v. 4.4).

Taula 4.8. Resum dels cristalls amb millors diagrames de difracció de les proteïnes HMG.

Cristall	Proteïna o pèptid	Oligonucleòtid	Condicions				Resol. màx. (Å)	Cel·la
			Sal	pH	Precip.	Precip. pou		
45-2 (v. fig. 4.42)	PRGRP*	AATAATTATT (10 pb)	5 mM MgCl ₂ 0,25 mM NiCl ₂	25 mM Tris pH=7,5	5% isop.	42% MPD 5% isop.	3,1	A
AS20-1 (v. fig. 4.50)	HMGB1 Box AS	AATAAATTTATT (12 pb)	20 mM MgAc	25 mM NaCac pH=6	15% MPD	40% MPD	2,8	B
AS32-2 (v. fig. 4.53)		ATATCGATAT (10 pb)	20 mM MgCl ₂	25 mM NaCac pH=6	2,5% MPD	40% MPD	1,8	C
AS32-3 (v. fig. 4.53)			20 mM MgCl ₂	25 mM NaCac pH=6	2,5% MPD	40% MPD	1,8	
AS35-2 (v. fig. 4.55)			5 mM MgCl ₂ 25 mM NH ₄ Ac	25 mM Tris pH=7,5	5% MPD	40% MPD	1,8	
AS36-1 (v. fig. 4.56)			-	25 mM NaCac pH=6	5% MPD	70% MPD	2,4	D
AS67-1 (v. fig. 4.54)			20 mM MgCl ₂	25 mM NaCac pH=6	2,5% MPD	35% MPD	2,1	E
RS20-B5-2 (v. fig. 4.67)	HMGB1 Box BS	AATAAATTTATT (12 pb)	20 mM NH ₄ Ac	25 mM Tris pH=7,5	4% PEG 3350	30% PEG 3350	3,3	F

Precip., precipitant. isop., isopropanol. A la taula s'indica la figura corresponent on es mostra la imatge, les condicions de la gota i el diagrama de difracció per a cada cristall. Els diferents tipus de cel·la unitària calculats s'han anomenat A-F (v. text). Per a més informació veure apartat anterior.

*Pèptid corresponent a la seqüència comuna als *AT-hooks* II i III de l'HMG.

4.3.1. Oligonucleòtid AATAATTATT i pèptid PRGRP

En primer lloc, hem indexat una de les millors imatges dels diagrames de difracció amb el programa *Denzo* [98]. Hem fet una selecció de les taques més importants i hem descartat les menys representatives. El programa ens mostra els paràmetres de cel·la que correspondrien a les nostres dades per a cada xarxa de Bravais i un error percentual respecte a la cel·la de menor simetria (la primitiva triclínic). En aquest cas la xarxa que ens dóna un menor error percentual és la xarxa centrada monoclínic. Per tant, seleccionem el grup espacial C2 per a realitzar el processat de la resta de diagrames de difracció. Degut a la falta de reflexions de les dades de partida (tenim molt pocs diagrames de difracció) els paràmetres estadístics, després de realitzar l'escalat amb el programa *Scalepack* [98], no són gaire òptims. De totes formes, amb l'escalat, obtenim un conjunt de reflexions amb els seus corresponents índexs de Miller (hkl) i les seves intensitats que ens ajudaran a la construcció d'un model aproximat de l'estructura. Dos dels diagrames de difracció obtinguts per a aquest cristall són els que podem veure a la fig. 4.82 juntament amb els paràmetres de cel·la d'aquest.

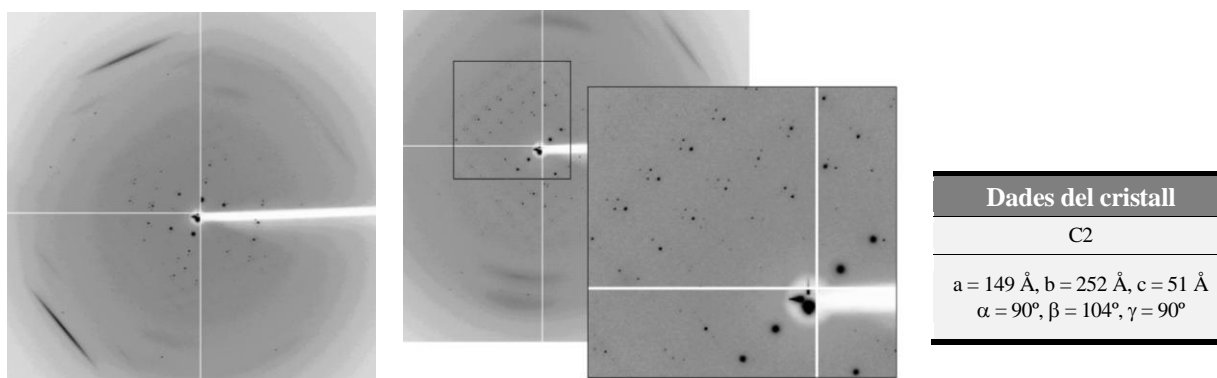


Fig. 4.82. Diagrames de difracció i dades del cristall 45-2.

La imatge de l'esquerra correspon a la foto número 13 i la de la dreta a la número 9 amb un detall de les reflexions. A la taula es recullen el grup espacial i els paràmetres de cel·la del cristall.

Si tenim en compte el volum d'aquesta cel·la, el grup espacial i el volum per parell de bases típic de seqüències únicament ATs ($\sim 1900 \text{ \AA}^3/\text{pb}$) [70] podem determinar que tenim 24 dúplexs dins de la unitat asimètrica. Com es tracta del grup C2, que té multiplicitat 4, obtindríem un total de 96 dúplexs de DNA dins la cel·la unitària. Tenint en compte els antecedents en empaquetament d'oligonucleòtids únicament ATs, la geometria de la cel·la i la dels oligonucleòtids, s'ha cregut que els oligonucleòtids estaven distribuïts a la unitat asimètrica de la següent forma: 8 cadenes paral·leles formades per 3 oligonucleòtids cadascuna al llarg de l'eix b i amb una inclinació de 40° respecte l'eix a. Per la simetria del grup espacial, obtenim la distribució dels oligonucleòtids a la cel·la segons el posicionament de la unitat asimètrica en els punts equivalents (x, y, z), (-x, y, -z), ($\frac{1}{2}+x, \frac{1}{2}+y, z$) i ($\frac{1}{2}-x, \frac{1}{2}+y, -z$), d'aquesta manera les cadenes de DNA queden creuades en

l'espai. Amb la col·laboració del Dr. Joan A. Subirana s'ha proposat un model d'organització preliminar d'aquesta estructura construït amb el programa *Cerius*² [107] (fig. 4.83).

En un primer moment es va creure que aquestes cadenes es trobaven sobre un pla perpendicular a l'eix *c* i que formaven dos pisos sobre *c*. Però aquesta organització ens proporcionaria una intensitat molt forta del pla (0 0 2); si observem les intensitats (*hkl*) podem veure que no tenim aquest pla amb tanta intensitat. Per tant, es va creure que les cadenes no es distribuïen perpendiculars a *c* sinó al llarg de la diagonal de la cel·la. D'aquesta manera els oligonucleòtids s'encadenarien d'una cel·la a l'altra veïna.

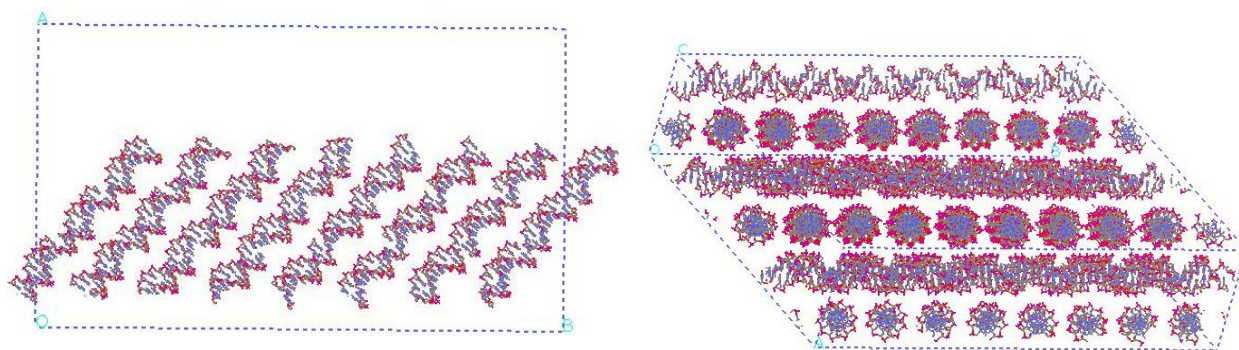


Fig. 4.83. Model de l'estructura corresponent al cristall 45-2 que conté l'oligonucleòtid AATAATTATT.

La imatge de la esquerra correspon a la vista frontal del pla *ab* d'una cel·la on s'han representat les 8 cadenes d'oligonucleòtids de la unitat asimètrica. A la imatge de la dreta observem tota la cel·la en una projecció sobre una diagonal de la cel·la.

Si observem detingudament els diagrames de difracció podem veure com aquests presenten reflexions més fortes envoltades de reflexions “satèl·lit” menys intenses, formant com un hexàgon. Es veu clarament al detall de la imatge de difracció de la fig. 4.82. Les reflexions satèl·lit s'obtenen pel gir de les cadenes de 3 oligonucleòtids sobre si mateixes respecte les veïnes (gir d'uns 45°). Si aquest gir no existís només obtindríem les reflexions més intenses i treballaríem amb una subcel·la 8 cops més petita ($a=74,4 \text{ \AA}$, $b=63 \text{ \AA}$, $c=51 \text{ \AA}$, $\beta=104^\circ$) que només constaria de dos cadenes de tres oligonucleòtids a la seva unitat asimètrica i amb un grup espacial P2 (v. fig. 4.84).

Per últim cal comentar que les molècules d'aquest cristall prenen una organització HASO (*Helical Arrangement of Stacked Oligonucleotides*). Aquest terme defineix un tipus d'organització en forma de columnes independents dins les cel·les amb aparellament Watson-Crick. Les bases terminals dels diferents dúplex interaccionen, definint una pseudo-hèlix continua (les bases defineixen un esglaió virtual on els fosfats són absents). Les estructures HASO sempre s'han trobat en seqüències riques en ATs [70]. Una altra característica és que cada dúplex gira un angle Ω respecte el seu veí. En el nostre cas el valor de Ω és de 330°, un angle més gran que el que s'havia obtingut en estudis anteriors amb l'oligonucleòtid AATAATTATT.

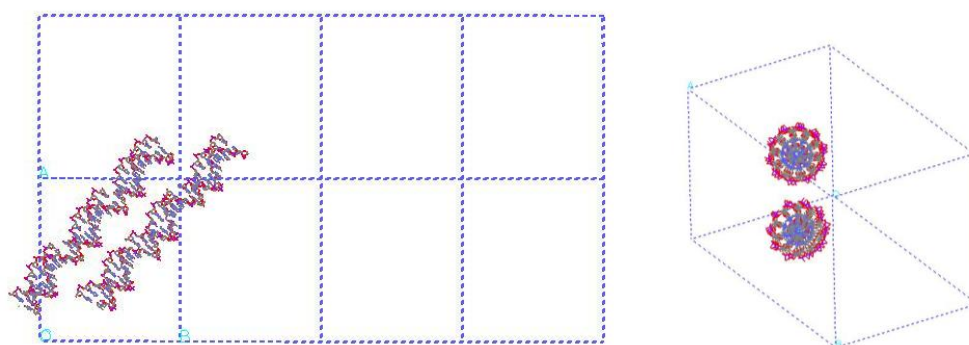


Fig. 4.84. Model de la subcel·la de l'estructura corresponent al cristall 45-2 que conté l'oligonucleòtid AATAATTATT.

A l'esquerra, vista frontal del pla ab de la cel·la on s'han representat les 2 cadenes d'oligonucleòtids de la unitat asimètrica de la subcel·la (rectangles més petits). A la dreta, projecció sobre la diagonal d'una subcel·la amb les dues cadenes d'oligonucleòtids (3 dúplex d'oligonucleòtids a cadascuna de les cadenes).

Per a comprovar que el model proposat s'assembla a l'estructura del cristall, obtindrem amb el programa *Cerius*² [107] un diagrama de difracció corresponent al model construït que podem comparar amb les dades de l'escalat del cristall 45-2 o bé amb els diagrames de difracció. Hem pogut constatar que la majoria dels hkl's corresponents al diagrama de difracció del model es troben també a l'escalat i amb una intensitat semblant. Un exemple, serien els plans $(-4\ 0\ 2)$ i $(4\ 0\ -2)$ que presenten una intensitat elevada tant al model com a les dades del cristall (v. fig. 4.85).

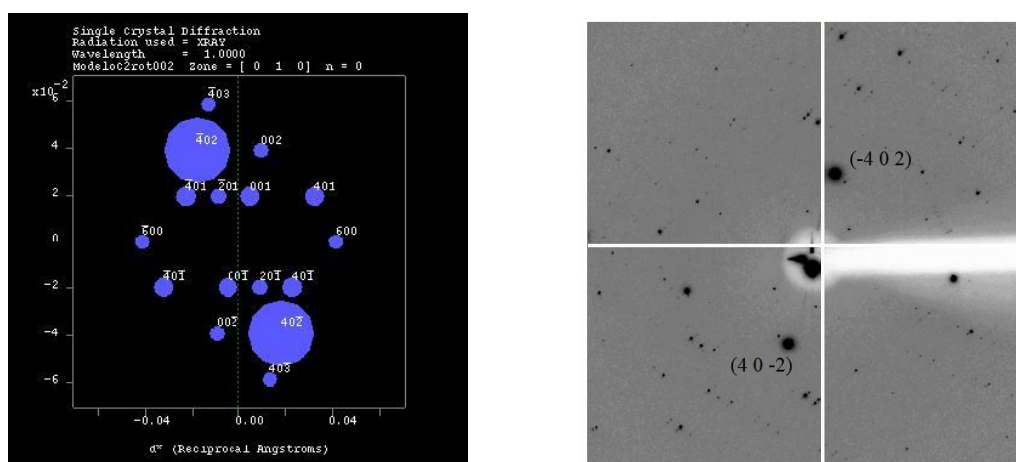


Fig. 4.85. Comparació del model construït amb les dades del cristall 45-2.

A l'esquerra, diagrama de difracció del model obtingut amb el programa *Cerius*² [107]. A la dreta, un dels diagrames de difracció del nostre cristall. En les dues imatges observem dos reflexions fortes corresponents als plans $(-4\ 0\ 2)$ i $(4\ 0\ -2)$.

4.3.2. Proteïnes HMGB1 Box AS i Box BS amb DNA

Hem obtingut uns bons diagrames de difracció corresponents al fragment Box AS de l'HMGB1 amb el dodecàmer AATAAATTTATT i el decàmer ATATCGATAT i per al domini Box BS amb el dodecàmer AATAAATTTATT. A la taula 4.9 es recullen els grups espacials i els paràmetres de cel·la corresponents a aquests cristalls estudiats amb el programa *iMosflm* [100]. En aquestes anàlisis preliminars de l'empaquetament s'ha triat el grup espacial amb major simetria i menor penalització per a cada cristall, però caldria processar els diagrames de difracció amb més profunditat per a confirmar el grup espacial. Observem que per a la Box AS i l'oligonucleòtid ATATCGATAT hem obtingut cristalls amb tres empaquetaments diferents (cel·les C, D i E).

El contingut dins de la unitat asimètrica de les cel·les trobades es pot estimar tenint en compte el seu volum. Conèixer el contingut macromolecular és la primera aproximació per a saber com estan les molècules posicionades a la cel·la, és a dir, l'empaquetament. Hem realitzat dues aproximacions per a calcular el nombre de molècules dins de la unitat asimètrica en aquest treball. Per una banda si el contingut a la cel·la és únicament DNA, tindríem en compte el volum típic en seqüències ATs per a parells de bases ($\sim 1900 \text{ \AA}^3/\text{pb}$) [70]), com hem fet en l'estructura anterior (v. 4.3.1). D'altra banda, si el cristall presenta tant proteïna com DNA estimaríem el contingut a la unitat asimètrica amb les probabilitats de *Matthews*. Això vol dir estimar el contingut en funció de l'opció més probable, la que presenta una major freqüència basant-se en les estructures de complexos ja resoltes [104]. En un principi els nostres cristalls han presentat contingut en proteïna en gels d'SDS, per tant caldria fer ús de la segona aproximació.

Un dels problemes en el càlcul del contingut a la unitat asimètrica resideix en saber la relació molar entre la proteïna i l'oligonucleòtid al complex. Idealment aquesta relació seria 1:1, basant-nos en altres estructures d'HMG-boxes, i per tant és aquesta la situació que hem agafat de referència per a calcular el contingut a la unitat asimètrica. Aquesta situació podria no reflectir la realitat, podent-se trobar major o menor nombre de molècules de proteïna per oligonucleòtid o bé presentar-se desordenada, però ens serveix com a primera aproximació. Amb el dodecàmer s'ha calculat també el contingut en només DNA per a tenir també una referència (v. taula 4.9).

A més a més l'observació dels perfils de taques a la zona de l'*stacking* del DNA (a $\sim 3,3 \text{ \AA}$) també ens aporta una informació molt valuosa. En les cel·les B (Box AS + dodecàmer), C i E (Box AS+ decàmer) s'observen en els diagrames de difracció dos direccions d'*stacking* indicant que les cadenes de DNA es troben creuades a l'espai. L'obtenció d'aquests dos *stackings* pot ser deguda a l'empaquetament dels oligonucleòtids en dues direccions creuades com observàvem per exemple en el model anterior (v. 4.3.1) o bé pel doblegament que produeixen les HMG-box a l'interaccionar amb el DNA com es veurà a l'estructura resolta en aquest treball (cel·la C, v. 4.4).

En el cas del dodecàmer AATAAATTTTATT observem que el perfil de taques en l'*stacking* és més definit apreciand-se una reflexió meridional (ratlla en el diagrama de difracció a $\sim 3,3$ Å) a diferència de l'oligonucleòtid ATATCGATAT on s'observen taques puntuals més intenses (comparar fig. 4.50 i fig. 4.53). Aquestes reflexions meridionals són molt típiques d'oligonucleòtids només ATs que normalment difracten a una resolució més baixa degut a l'alta hidratació pròpia d'aquestes seqüències [70].

Amb totes aquestes consideracions es pot tenir una idea de l'empaquetament de les estructures que no hem arribat a resoldre (cel·les B, D, E i F), però és evident que el model d'organització a la cel·la no és directament deduïble de l'anàlisi dels diagrames de difracció i caldria intentar resoldre aquestes estructures més complexes per a confirmar-ne el seu contingut. A continuació es discutiran els possibles empaquetaments.

A la taula 4.9 es detallen, juntament amb els grups espacials i els paràmetres de cel·la, les característiques més destacades en els diagrames de difracció dels cristalls dels fragments de l'HMGB1 corresponents a les cel·les B, C, D E i F .

Taula 4.9. Empaquetament dels diferents cristalls de les proteïnes HMGB1.

Cristall	Proteïna	Oligo-nucleòtid	Resol. màx.	Nom cel·la	Grup espacial i paràmetres de cel·la	Volum de cel·la	Stacking a la difracció	Contingut unitat asimètrica
AS20-1	HMGB1 Box AS	AATAAA- TTTATT (12 pb)	2,8	B	Si P1 (m=1) a=39 Å, b=39 Å, c=54 Å $\alpha=81^\circ$, $\beta=86^\circ$, $\gamma=74^\circ$	81400	Dos <i>stackings</i> creuats i ben definits	4 dúplexs (o bé 2 complexos)
AS32-2			1,8	C	Si P222 (m=4) a=43 Å, b=84 Å, c=94 Å $\alpha=90^\circ$, $\beta=90^\circ$, $\gamma=90^\circ$	338000	Dos <i>stackings</i> creuats i difusos	2 complexos
AS32-3			1,8					
AS35-2		1,8	ATATC- GATAT (10 pb)	D	Si C222 (m=8) a=94 Å, b=94 Å, c=100 Å $\alpha=90^\circ$, $\beta=90^\circ$, $\gamma=90^\circ$	880000	No es veuen els <i>stackings</i>	3 complexos
AS36-1		2,4		E	Si C222 (m=8) a=105 Å, b=163 Å, c=89 Å $\alpha=90^\circ$, $\beta=90^\circ$, $\gamma=90^\circ$	1530000	Dos <i>stackings</i> creuats i difusos	5 complexos
AS67-1		2,1	AATAAA- TTTATT (12 pb)	F	Si P222 (m=4) a=39 Å, b=68 Å, c=91 Å $\alpha=90^\circ$, $\beta=90^\circ$, $\gamma=90^\circ$	238000	Un <i>stacking</i> ben definit	3 dúplexs (o bé 2 proteïnes i 1 oligo o bé 1 proteïna i 2 oligos)
RS20-B5-2	3,3							

Resol. màx., límit de resolució, és a dir, màxima resolució observada d'una reflexió per als diagrames de difracció d'un cristall. S'ha indicat el grup espacial més probable (escollint el grup espacial amb major simetria i menor penalització), la multiplicitat d'aquest (m) i els paràmetres de cel·la per a aquest grup calculats amb el programa *iMosflm* [100]. El contingut en la unitat asimètrica ha estat estimat amb el coeficient de *Matthews*, fent servir unes masses moleculars per als oligonucleòtids i per a les proteïnes de: 7280 Da (AATAAATTTTATT), 6050 Da (ATATCGATAT), 8930 Da (Box AS) i 8430 Da (Box BS).

a) Proteïna Box AS amb AATAAATTTATT (cristall AS20, cel·la B)

En aquest cas observem reflexions meridionals dels *stackings* de les bases del DNA creuades uns 60° i teòricament el contingut de la unitat asimètrica (i també de la cel·la ja que és P1) és de 2 complexos o bé, si només hi estigués present el DNA, 4 dúplexs. Com hem comentat, aquestes dues direccions de les cadenes del DNA poden indicar que els oligonucleòtids estan creuats entre ells a l'espai o bé que els oligonucleòtids es troben doblegats degut a la proteïna.

S'ha proposat un model preliminar d'empaquetament d'aquest cristall a partir del processat dels diagrames de difracció. S'ha realitzat un integrat i un escalat dels diagrames de difracció i a continuació unes proves de reemplaçament molecular amb un model de l'oligonucleòtid AATAAATTTATT ideal (amb un *rise* de $3,2 \text{ \AA}$) amb els programes *iMosflm* [100] i *MolRep* [108], respectivament (v. apartat 3.9.5). En aquesta anàlisi preliminar de l'estructura s'han pogut posicionar només dos oligonucleòtids com s'observa a la fig. 4.86.

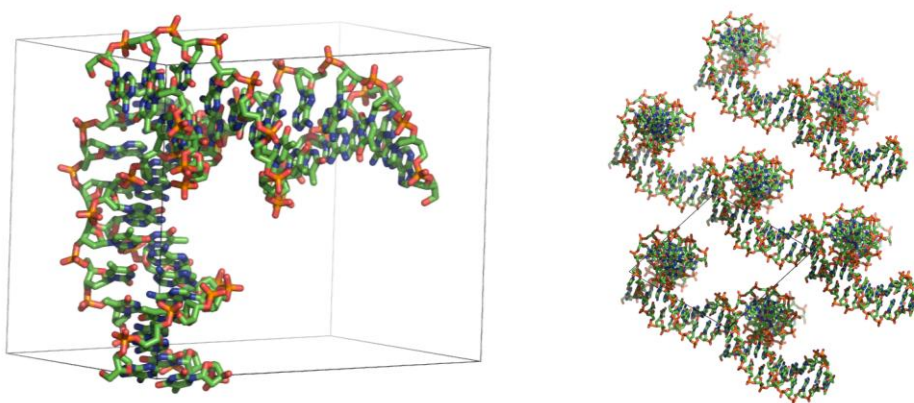
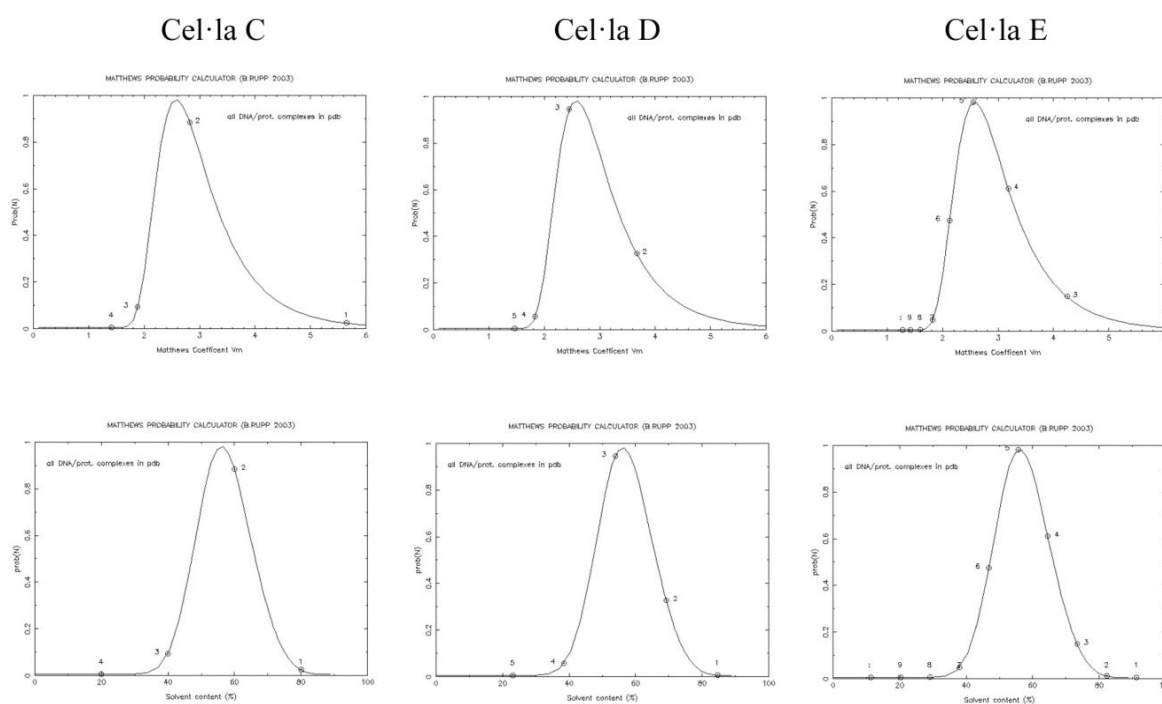


Fig. 4.86. Model de l'estructura corresponent al cristall AS20 que conté l'oligonucleòtid AATAAATTTATT. La imatge de la esquerra correspon a la vista d'una cel·la unitària contenint dos dúplexs de DNA creuats uns 60° . A la dreta podem observar l'empaquetament de les molècules.

Caldria veure si es pot localitzar la proteïna dins de la cel·la i si aquesta es troba de forma ordenada en els llocs buits que queden en l'empaquetament en forma de columnes creuades del DNA. O si, per una altra banda, l'oligonucleòtid es troba doblegat per la proteïna i per tant caldria canviar aquest model de columnes creuades per un model on l'oligonucleòtid estigués doblegat i interaccionant amb la proteïna.

b) Proteïna Box AS amb ATATCGATAT (cel·les C, D i E)

Les cel·les C (cristalls AS32 i AS35-2) i E (cristall AS67-1) presenten perfils de taques corresponents a l'*stacking* del DNA similars, mostrant dues direccions creuades en el posicionament de les dobles hèlixs del DNA als cristalls. La cel·la C ($a=43 \text{ \AA}$, $b=84 \text{ \AA}$, $c=94 \text{ \AA}$), corresponent al grup espacial P222, presenta un menor volum que la cel·la E ($a=105 \text{ \AA}$, $b=163 \text{ \AA}$, $c=89 \text{ \AA}$) amb un grup espacial C222 (v. taula 4.9). Per aquesta raó el contingut estimat a la unitat asimètrica de la primera cel·la és de 2 complexos mentre que en el de la segona és de 5. S'obtidrien així, tenint en compte la multiplicitat de cada grup espacial, 8 complexos en tota la cel·la C i 40 complexos en el cas de la cel·la E. A la fig. 4.87 podem observar les distribucions de probabilitats per al coeficient de *Matthews* i per al percentatge de solvent per aquestes cel·les on podem veure que els màxims dels pics ens indiquen el nombre de complexos estimat a la unitat asimètrica.



Cel·la	Nº complexos	V_m ($\text{\AA}^3/\text{Da}$)	Contingut en solvent (%)
C	2	2,82	59,98
D	3	2,45	53,83
E	5	2,55	55,68

Fig. 4.87. Probabilitats de *Matthews* per al coeficient de *Matthews* (V_m) i per al contingut en solvent.

S'ha estimat el nombre de complexos a cada cel·la amb les probabilitats de *Matthews* [104], suposant que els complexos s'han format en una relació molar 1:1 (una molècula de proteïna per un dúplex). A les gràfiques poden observar el nombre de complexos més probable (màxim del pic) a la unitat asimètrica en relació a la freqüència del coeficient de *Matthews* (fila superior) i del contingut en solvent (fila inferior). A la taula es detallen els valors del coeficient de *Matthews* i del contingut en solvent per al nombre estimat de complexos a la unitat asimètrica.

Es va decidir prioritzar la resolució de l'estructura de la cel·la C degut a la millor forma de les taques en els diagrames de difracció, al millor límit de resolució i al menor nombre de molècules a la unitat asimètrica en relació amb la cel·la E. A l'apartat 4.4 s'explicarà amb més detall el processat dels diagrames de difracció i la resolució per reemplaçament molecular de l'estructura.

D'altra banda, els diagrames de difracció corresponents a la cel·la D són completament diferents als observats fins ara. Observem una resolució força bona però no s'aprecien unes taques més intenses a les resolucions de l'*stacking* de les bases del DNA. Creiem que podria ser perquè la proteïna està unida de tal manera al DNA que el doblega completament i aquest no presenta unes direccions preferents de la seva doble hèlix. L'altre opció és que el DNA del complex es trobi de forma desordenada entre les posicions que ocupa la proteïna.

c) Proteïna Box BS amb AATAAATTTATT (cristall RS20-B5-2, cel·la F)

En el cas de la proteïna Box BS amb l'oligonucleòtid AATAAATTTATT en el cristall RS20-B5-2, observem una única direcció d'*stacking* al llarg de l'eix x en els diagrames de difracció. Aquest fet, indicaria que els oligonucleòtids del complex estan posicionats en columnes paral·leles al llarg del costat *a* de la cel·la seguint una única direcció. Així aquest dodecàmer d'aproximadament 38,4 Å d'alçada encaixaria al llarg del costat *a* ($a=39$ Å) (v. fig. 4.88).

En unes proves preliminars de reemplaçament molecular realitzades amb el programa *MolRep* [108] (v. apartat 3.9.5) es van arribar a posicionar dos dúplexs de DNA. El tercer oligonucleòtid, que estaria contingut en aquesta cel·la i de forma paral·lela als altres dos, no es va aconseguir posicionar indicant que potser en aquest espai es trobaria la proteïna. Podria ser que la unitat asimètrica contingui dos dúplexs de DNA per una molècula de proteïna als llocs que queden buits o bé que només hi hagués un DNA amb dues molècules de proteïna. En tots dos casos semblaria que la interacció d'aquesta proteïna no provocaria un doblegament important de l'oligonucleòtid.

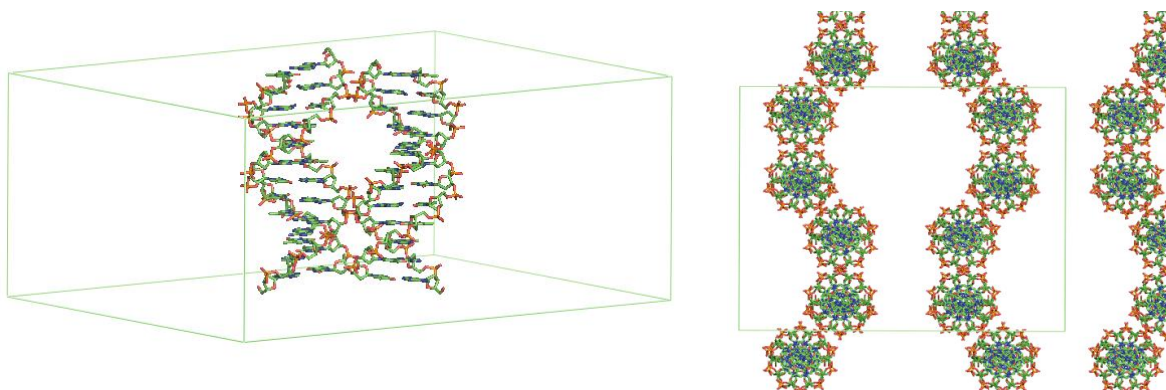


Fig. 4.88. Model de l'estructura corresponent al cristall RS20-B5-2 que conté l'oligonucleòtid AATAAATTATT.

La imatge de l'esquerra correspon a la vista d'una cel·la unitària contenint dos dúplex de DNA a la unitat asimètrica. A la dreta podem observar l'empaquetament de les molècules en la cel·la. Els oligonucleòtids, també en aquest cas, formarien columnes paral·leles apilades com la majoria dels oligonucleòtids només ATs.

4.4. Resolució de l'estructura del complex HMGB1 Box AS-DNA

En aquest treball ens hem centrat en la resolució de l'estructura corresponent al cristall AS32-3 ja que era el que presentava millor resolució d'entre els cristalls obtinguts. Aquest cristall correspon al complex Box AS amb l'oligonucleòtid ATATCGATAT i va difractar fins a una resolució d'1,8 Å (veure apartat anterior per a més detalls). A continuació s'explicarà com s'ha arribat a trobar l'estructura atòmica d'aquest cristall i a l'apartat 4.5 s'analitzaran els trets més importants en la interacció del domini Box AS amb el DNA observats en aquesta estructura.

4.4.1. Anàlisi dels diagrames de difracció

El primer pas en la resolució de l'estructura d'un cristall és l'anàlisi dels diagrames de difracció per a l'obtenció final d'un fitxer que contingui les intensitats corresponents a cada reflexió (hkl) (v. apartat 3.9.4). Es va partir d'una difracció a la línia Xaloc del sincrotró ALBA del cristall AS32-3 que consistia en una recollida de 270 imatges de difracció cada 1° (temps d'exposició 4 s, distància al detector 425 mm i longitud d'ona 0,9795 Å). Per a processar aquests diagrames es va fer servir el programa *HKL2000* [99].

En primer lloc es va triar un diagrama de difracció, es van definir els límits de resolució de l'indexat i la integració i es va fer una cerca de les reflexions o taques d'aquest. A continuació el programa realitza l'indexat de les reflexions i mostra els paràmetres de cel·la per a les diferents xarxes de Bravais indicant un error percentual respecte a la cel·la de menor simetria (primitiva triclinica). En el nostre cas es va obtenir l'error més baix per a la primitiva monoclinica P2 (0,04%) i per a la primitiva ortoròmbica P222 (0,06%) (v. annex E). Es va triar una xarxa de Bravais P222 (menor error amb major simetria) per a continuar fent el refinament de l'indexat. Prèviament a l'indexat o en aquest punt es poden variar alguns paràmetres com per exemple la mida dels radis de les taques i del *background*, i la mida de les caixes d'integració o la mosaïcitat per tal d'ajustar el perfil de reflexions estimat al nostre experiment. En aquest pas també es poden afinar automàticament d'altres paràmetres de la integració (triats per l'usuari) com la distància al detector, les coordenades x i y del raig o els paràmetres de cel·la. En tot moment aquest programa ens permet veure la predicció de taques i el paràmetre estadístic χ^2 per a optimitzar aquests paràmetres el màxim possible per a les nostres dades de difracció.

Tot seguit es va dur a terme la integració de les reflexions de tots els diagrames de difracció per tal d'obtenir un llistat de les reflexions i les seves intensitats per a cada diagrama de difracció (fitxers *.x* amb *HKL2000*). A continuació vam fer l'escalat per a posar a la mateixa escala les intensitats en un únic arxiu.

Durant l'escalat el programa ens permet millorar la qualitat dels resultats variant diferents paràmetres de l'escalat o exclouent diagrames de difracció que presentin major error respecte la resta. També es pot definir el grup espacial dins de la xarxa de Bravais triada a l'indexat i fixar el límit de resolució. La visualització gràfica dels paràmetres estadístics d'avaluació de l'escalat permet realitzar una optimització de forma iterativa d'aquests paràmetres més fàcilment (v. fig. 4.89). Concretament en el nostre cas va caldre excloure alguns dels últims diagrames de difracció de la recollida ja que presentaven un B-factor elevat degut a la incidència prolongada de raigs X sobre el cristall (i per tant també una valors alts de χ^2 i R_{merge}) (v. fig. 4.89 a i b). Es va definir una resolució límit de 2.00 Å per a evitar la introducció d'error (normalment una bona manera per a valorar el límit de resolució de la difracció és la ràtio entre la intensitat i l'error de la intensitat que hauria de ser $I/\sigma > 2$) (v. fig. 4.89 c). A continuació es van ajustar els valors de l'*error model* per tal de portar χ^2 a valors propers a 1. Com a grup espacial es va triar el $P2_12_12_1$.

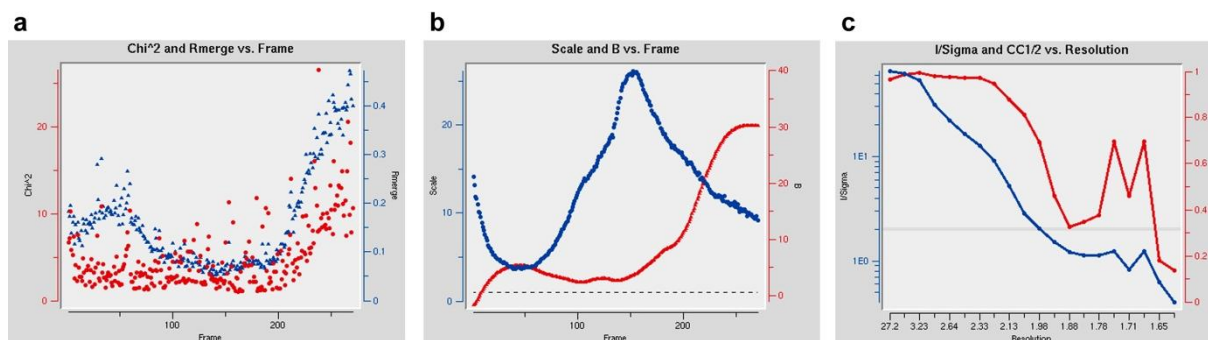


Fig. 4.89. Algunes gràfiques dels paràmetres d'avaluació obtingudes al principi de l'escalat.

- a) Representació de χ^2 (bonança de l'adequació) (vermell) i R_{merge} (concordança entre reflexions simètriques) (blau) en funció del diagrama de difracció (*Frame*). b) Representació del B-factor (vermell) en funció del diagrama de difracció (*Frame*). c) Representació de I/σ (blau) en funció de la resolució.

Finalment, els valors estadístics que es van obtenir per a aquesta difracció són els que es troben a la fig. 4.90. Observem com a mesura que augmenta la resolució empitjoren relativament els paràmetres estadístics (R_{merge} o relació entre I i error) com ja és d'esperar.

Shell limit	Lower Angstrom	Upper Angstrom	Average I	Average error	Average stat.	Norm. Chi**2	Linear R-fac	Square R-fac	Rmeas	Rpim	CC1/2	CC*
50.00	4.31	4.31	1365.1	103.8	17.9	1.157	0.089	0.160	0.096	0.036	0.976	0.994
	4.31	3.42	1105.3	46.9	12.1	1.752	0.078	0.094	0.084	0.030	0.994	0.999
	3.42	2.99	652.4	24.1	9.3	1.991	0.096	0.120	0.103	0.037	0.995	0.999
	2.99	2.71	202.9	11.3	6.2	1.200	0.141	0.158	0.151	0.054	0.984	0.996
	2.71	2.52	114.9	7.0	5.2	1.134	0.203	0.235	0.217	0.076	0.986	0.996
	2.52	2.37	84.9	6.5	5.2	1.365	0.282	0.569	0.302	0.106	0.966	0.991
	2.37	2.25	63.4	6.2	5.3	1.338	0.373	0.000	0.400	0.143	0.917	0.978
	2.25	2.15	46.5	6.2	5.4	0.988	0.489	0.677	0.524	0.185	0.950	0.987
	2.15	2.07	32.2	6.9	6.4	0.846	0.553	0.788	0.600	0.227	0.891	0.971
	2.07	2.00	23.4	12.3	9.0	0.460	0.658	0.899	0.727	0.301	0.783	0.937
All reflections			384.6	24.0	8.3	1.268	0.113	0.140	0.122	0.045		

Fig. 4.90. Sortida de l'escalat del programa *HKL2000* mostrant els valors estadístics obtinguts.

Valors obtinguts de l'anàlisi dels diagrames de difracció del cristall AS32-3 amb el programa *HKL2000* [99].

Es mostren per a deu intervals de resolució els diferents valors obtinguts per a paràmetres com la intensitat (I), l'error (error), χ^2 (Chi**2) i R_{merge} (R-fac lineal).

L'elecció del grup espacial $P2_12_12_1$ es va confirmar amb el programa Pointless [101]. Aquest programa puntua els diferents grups de Laue emparellant reflexions simètricament equivalents i comprova absències sistemàtiques per indicar possibles grups espacials.

A la taula 4.10 es recullen les principals dades obtingudes de l'anàlisi dels diagrames de difracció d'aquest cristall. L'últim pas va consistir en transformar les intensitats per a cada reflexió obtingudes a factors d'estructura amb el programa *Truncate* [102] (del paquet CCP4 [103]). És a dir passar d'un fitxer *.sca* obtingut amb el programa *HKL2000* a un *.mtz* per a fer servir en el reemplaçament molecular.

Taula 4.10. Dades de l'anàlisi dels diagrames de difracció i refinament.

Dades de l'anàlisi dels diagrames de difracció del cristall AS32-3	
Grup espacial	$P2_12_12_1$
Paràmetres de cel·la	
a, b, c (Å)	42,8 84,2 94,2
α, β, γ (°)	90,0 90,0 90,0
Resolució (Å)	2,00 (2,07-2,00)
R_{merge}	11,3 (65,8)
I/σ	16,0 (1,90)
Completitud (%)	98,9 (93,5)
Redundància	7,1 (5,1)

Entre parèntesis es recullen els valors corresponents a l'últim interval de resolucions.

4.4.2. Consideracions prèvies al reemplaçament molecular

Com hem vist, el contingut en la unitat asimètrica estimat amb el coeficient de *Matthews* [104] és de dos proteïnes (és a dir, dos dominis Box AS) i dos oligonucleòtids (dos dúplexs ATATCGATAT) (v. 4.3.2.). Per a aquest contingut el coeficient de *Matthews* té un valor de 2,82 Å³/Da i el contingut en solvent és del 58,98%.

Abans d'iniciar el reemplaçament molecular per a resoldre l'estructura vam obtenir un mapa de Patterson (v. fig. 4.91) i es va fer la funció d'autorotació (*self-rotation function*) (v. fig. 4.92) amb les eines del programa CCP4 [103]. Amb aquestes dues anàlisis preliminars es pretén obtenir alguna informació addicional sobre l'empaquetament de les molècules dins del cristall abans d'iniciar el reemplaçament molecular. A l'annex F es desenvolupen una mica més els conceptes de mapa de Patterson i la funció d'autorotació.

Al mapa de Patterson (v. fig. 4.91) s'observa un perfil de densitats en les diagonals internes de la cel·la corresponents a unes distàncies repetitives entre àtoms de l'estructura. Normalment s'obté un perfil d'aquest tipus en estructures de DNA on les distàncies entre bases (*stacking*) dels oligonucleòtids apilats es troben repetides al llarg de la cel·la. L'obtenció d'aquestes dues orientacions preferents en les dues diagonals internes de la cel·la i creuades a l'espai suggeria que el nostre oligonucleòtid podia estar orientat d'aquesta manera al cristall. Aquesta observació ja és coherent amb els *stackings* creuats que presentaven els diagrames de difracció.

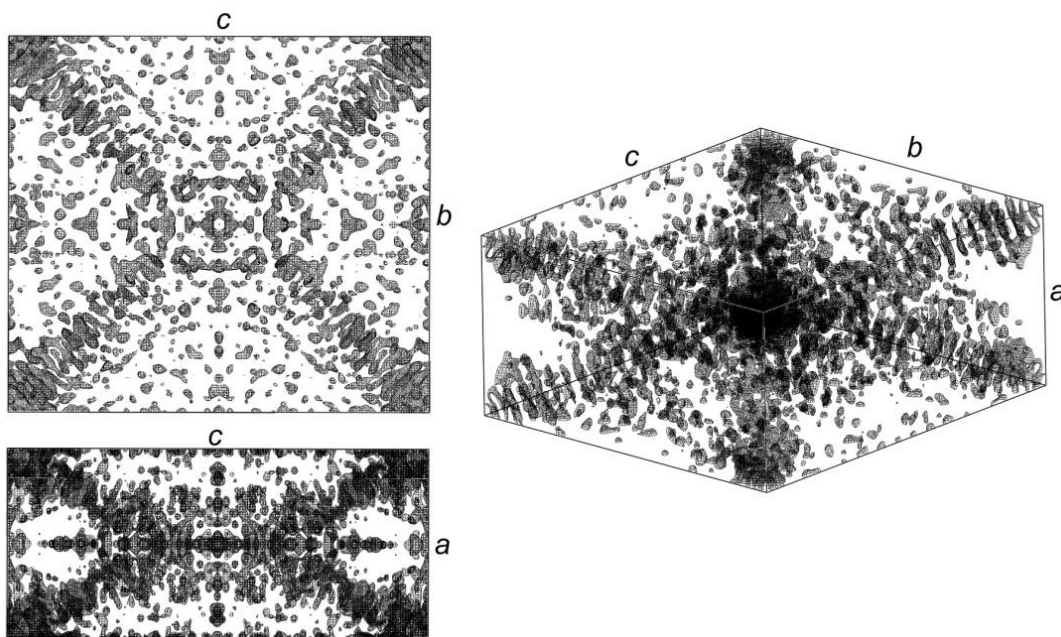


Fig. 4.91. Mapa de Patterson per al cristall AS32-3.

A l'esquerra s'observen dues projeccions sobre les cares *cb* i *ca* de la cel·la del cristall amb el mapa de Patterson obtingut i a la dreta una projecció sobre una de les diagonals internes.

En la funció de rotació observem en la secció $\kappa=180^\circ$ (indicada com a Chi a la fig. 4.92) la presència de quatre eixos binaris. Tres d'aquests eixos corresponen als tres eixos helicoïdals propis del grup espacial $P2_12_12_1$ del cristall (indicats amb * a la figura) mentre que el quart representaria un eix de simetria addicional (indicat amb + a la figura). A aquests eixos addicionals no relacionats amb la simetria del grup espacial se'ls anomena eixos de simetria no cristal·logràfica (NCS, de les seves sigles en anglès). El nostre eix NCS implica que dins de la unitat asimètrica hi ha dues molècules (o subunitats) que presenten una simetria entre elles. En l'afinament de l'estructura es va fer servir aquesta simetria no cristal·logràfica com un paràmetre de restricció aplicat pel programa *Refmac5* (v. 4.4.4).

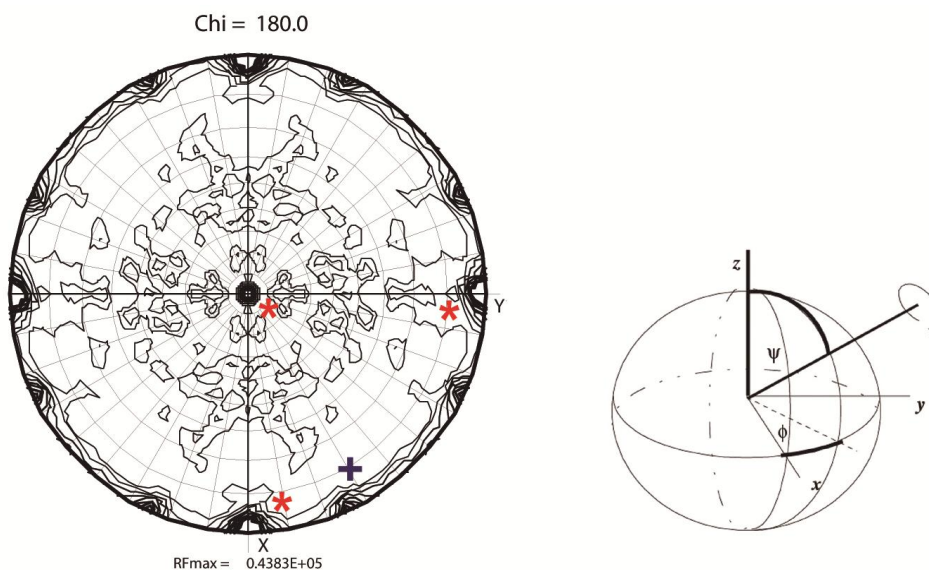


Fig. 4.92. *Self-rotation function* per al cristall AS32-3.

Observem tres eixos binaris corresponents al grup espacial del cristall ($\phi=0^\circ, \psi=0^\circ$; $\phi=0^\circ, \psi=90^\circ$; $\phi=90^\circ, \psi=90^\circ$) (indicats per *) i un eix binari de simetria no cristal·logràfica ($\phi=25^\circ, \psi=90^\circ$) (indicat per +).

A la dreta es mostra un esquema de la representació de les coordenades polars ϕ i ψ .

4.4.3. Reemplaçament molecular

Per a realitzar el reemplaçament molecular és necessari partir d'uns models corresponents al domini Box A i al dúplex ATATCGATAT. Com a model del dúplex ATATCGATAT es va generar primerament un dúplex ideal corresponent a aquesta seqüència amb uns valors de 36° pel *twist* i $3,32 \text{ \AA}$ pel *rise* amb el programa *Turbo-Frodo* [105] (v. fig. 4.93 a).

D'altra banda, com a model de la proteïna, s'ha fet servir l'estructura del Box A cristal·litzat amb un DNA modificat amb cisplatí [43] corresponent a l'entrada del *Protein Data Bank* (PDB) 1CKT. Es va modificar aquest fitxer per a tenir un model (fitxer *pdb*) contenint un únic domini Box A (v. fig. 4.93 b). S'haurà de tenir en compte durant l'afinament que aquest fragment de la proteïna és una mica més curt i caldrà modelar els aminoàcids restants (la seqüència aminoacídica comprèn del residu Pro8 al Tyr77 mentre que el nostre fragment va del residu Lys7 al Pro80 i a més a més conté una Gly i una Ser N-terminals producte del clonatge).

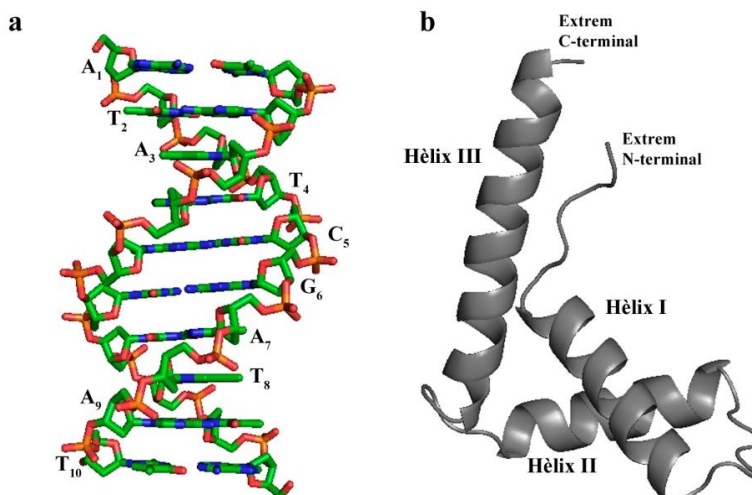


Fig. 4.93. Models del dúplex de DNA i de la proteïna fets servir per al reemplaçament molecular.

a) Model ideal del dúplex ATATCGATAT preparat amb el programa *Turbo-Frodo* [105]. b) Model de la proteïna obtingut del *Protein Data Bank* per a l'entrada 1CKT. Imatge preparada amb el programa *Pymol* [105].

La fase del reemplaçament molecular va ser la que ens va requerir més temps durant la resolució de l'estructura degut a les dificultats que presentava. En primer lloc desconèixiem si els oligonucleòtids es trobaven creuats a l'espai posicionats en les diagonals de la cel·la com indicava el mapa de Patterson o bé si les dues orientacions del mapa podrien ser degudes a un doblegament del DNA provocat per la proteïna. A més a més, tot i que pels gels realitzats s'esperava la presència de proteïna al cristall, no podíem saber si aquesta s'hauria ordenat o no (i si, per tant, podríem localitzar-la). El coeficient de *Matthews* estimava que podria haver-hi dos complexos a la unitat asimètrica però, ja que el dúplex i la proteïna presenten masses moleculars semblants, podria ser també que trobéssim tres dúplexs i una proteïna per exemple o qualsevol altra combinació. Per tant, les proves a realitzar durant el reemplaçament molecular eren molt nombroses. A més a més també es van fer servir diferents programes pel reemplaçament molecular: *MolRep* [108] (del paquet CCP4), *AMoRe* [109] i *Phaser* [110] (del paquet CCP4) ja que a l'utilitzar mètodes diferents per a realitzar les funcions de rotació i translació poden permetre l'obtenció de diferents solucions.

Amb totes aquestes consideracions es va començar realitzant els primers reemplaçaments moleculars amb diferents models del dúplex del DNA ideal de la fig. 4.93 a (i amb d'altres valors pel *twist* i pel *rise*). Es van realitzar reemplaçaments posicionant quatre oligonucleòtids a la unitat asimètrica per tal que quedessin a les diagonals però les solucions no eren les correctes (R molt alt i C molt baix). També es va provar de posicionar dos dúplexs a les diagonals i una o dues proteïnes, però tampoc donaven bons valors de R i C.

A continuació es van realitzar reemplaçaments amb models de DNA doblegats, tant d'altres estructures de complexos d'HMG-box (amb seqüències nucleotídiques diferents) o amb models de la nostra seqüència doblegats de diferents formes (generats amb el programa *Namot* [132]

amb diferents valors pel *roll* i *twist*). Es van provar també diferents combinacions i ordres de cerca (primer proteïna i després DNA i a l'inversa). Cap d'elles ens va donar una bona solució.

Finalment es va pensar que potser a la cel·la es trobaven dúplexs rectes posicionats en les dues direccions de les diagonals internes de la cel·la connectats entre ells per dúplexs doblegats possiblement per un domini Box AS. Amb aquest plantejament es va aconseguir trobar la solució bona. Es va fer servir un model del fragment del dúplex ATAT (generat amb *Turbo-Frodo* [105]) com a model de cerca del reemplaçament molecular amb el programa *Phaser* [110]. Un cop localitzats dos fragments d'aquesta seqüència a la unitat asimètrica, es van posicionar dos dominis Box AS amb *Molrep* [108] (en aquest punt es va veure que una segona molècula de proteïna encaixava bé). Els dos parells de bases CG restants van ser construïts amb *Coot* [112]. Finalment el dúplex recte es va posicionar amb *Molrep* [108].

Per a conèixer l'adequació de les solucions es tenen en compte els valors de R i C que donen els programes de reemplaçament molecular. En el cas de *Molrep* el paràmetre del coeficient de correlació (C) és substituït per un *score* (producte del coeficient de correlació i el màxim valor de la funció d'empaquetament, la qual penalitza les solucions amb solapament de molècules). Comparant les solucions no adequades amb la solució escollida es va passar d'uns valors per a R de ~ 65 i un *score* de ~ 22 a un R de 45 i un *score* de 74, demostrant que aquesta solució és la que més s'adequa al nostre resultat.

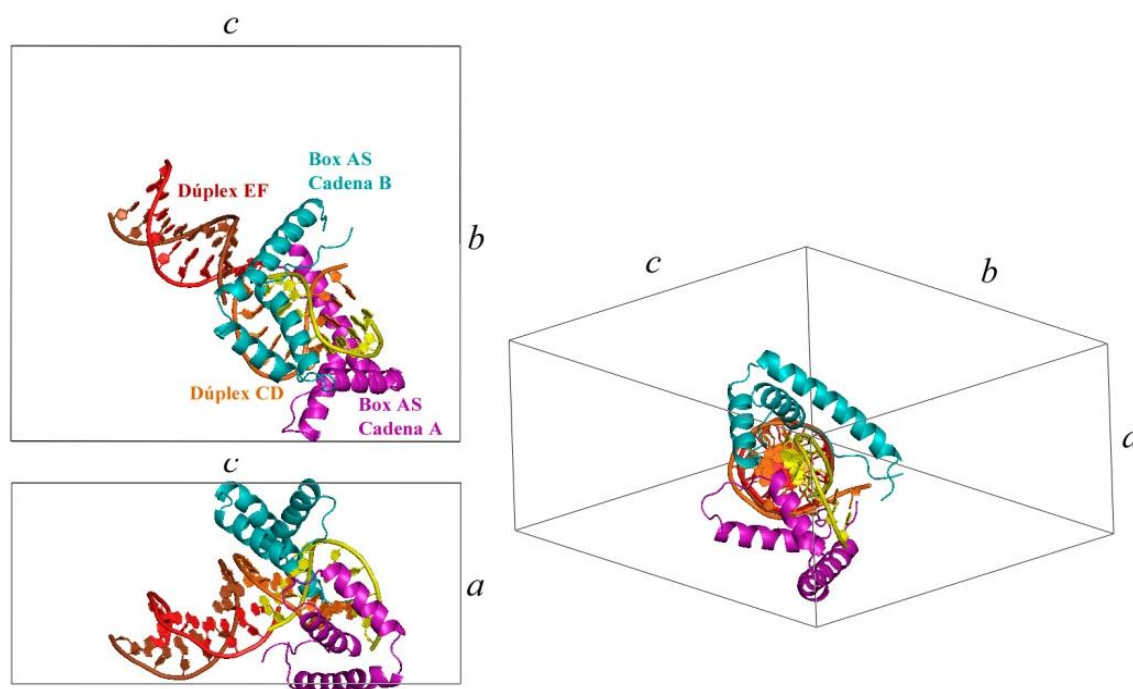


Fig. 4.94. Contingut de la unitat asimètrica.

Es mostra el contingut localitzat en els diferents passos del reemplaçament molecular a la unitat asimètrica: dos dominis Box AS (cian i magenta) i dos dúplexs de DNA (vermell/marró i taronja/groc).

4.4.4. Afinament i validació

L'afinament permet ajustar millor les posicions dels àtoms i afegir molècules, ions o residus que faltessin. En el nostre cas vam realitzar uns primers cicles d'afinament per probabilitat màxima (*maximum likelihood refinement*) amb el programa *Refmac5* i a continuació l'observació de zones verdes al mapa de diferències ens va permetre afegir les molècules d'aigües i un ió magnesi hexahidratat al model. També es van construir els residus que mancaven a l'extrem N-terminal (Gly5-Ser6-Lys7) i a l'extrem C-terminal (Ile78-Pro79), excepte l'última prolina que no presentava mapa (Pro80). Per a la resta de residus es va buscar una a una la conformació del rotàmer adient que millor s'adequés al mapa de densitat. La visualització dels mapes de densitat i de diferències amb el model i aquestes modificacions es van realitzar amb el programa *Coot* [112].

Arribats a aquest punt cal realitzar petites variacions del model per una banda perquè encaixi amb el mapa i per una altra per a assegurar-nos que no hi hagin xocs entre cadenes i que els angles i les distàncies d'enllaç i els paràmetres tèrmics siguin raonables. Per ajudar-nos a analitzar aquests paràmetres estereoquímics vam fer servir el programa *Molprobit* [114]. El programa ens va proporcionar els paràmetres que tenien valors fora de la norma. A continuació es van rectificar les posicions dels àtoms per a complir els valors acceptables per al programa (els valors acceptables corresponen a la mitjana obtinguda per aquest valor més o menys quatre cops la desviació estàndard). En el nostre cas, per exemple, va ser necessari modificar algunes conformacions de les cadenes laterals d'alguns residus de la proteïna per tal que presentessin una conformació del rotàmer preferent i també es van ajustar algunes distàncies i alguns angles d'enllaç aplicant uns paràmetres de restricció durant els cicles de l'afinament. En concret va ser necessari imposar algunes distàncies corresponents als ponts d'hidrogen entre bases de la doble hèlix i alguns angles de la cadena sucre-fosfat del dúplex (especialment del dúplex recte).

També va ser necessari aplicar una segona conformació a la Cys44(A), a l'Asp32(B) i a la Ser45(B) ja que el mapa de densitat ho demandava. Per tal d'aplicar el percentatge d'ocupació de cada conformació es van fer cicles de refinament de forma iterativa.

Finalment cal comentar que en tots aquests passos d'afinament per probabilitat màxima es van aplicar també unes restriccions de simetria no cristal·logràfica que relacionen els dos dominis Box AS de la unitat asimètrica (la cadena A amb la B). L'eix de simetria no cristal·logràfica que s'havia vist amb la funció d'autorotació corresponia justament a la relació entre les dues proteïnes. En el nostre cas no se li va aplicar molt de pes a la restricció NCS ja que empitjorava els valors de R_{work} i R_{free} i per tant observàvem que una proteïna respecte l'altra presentava certes diferències.

Taula 4.11. Dades obtingudes durant l'afinament del cristall AS32-3.

Dades del refinament cristall AS32-3	
Resolució (Å)	2,00
Nombre de reflexions	22219
$R_{\text{work}}/R_{\text{free}}$	19,9/23,4
Nombre d'àtoms	
Proteïna	1254
DNA	808
Mg ⁺²	1
Aigües	115
Total	2178
B-factors	
Proteïna	44,4
DNA	37,6
Mg ⁺²	21,1
Mitjana	42,4
Desviacions R.m.s.d.	
Distàncies d'enllaç (Å)	0,016
Angles d'enllaç (°)	1,62

Un cop els paràmetres estereoquímics presenten valors acceptables (v. fig. 4.95 i fig. 4.96) i hem modificat el model tot el possible perquè s'adapti al mapa, es dona el model per bo. A la taula 4.11 trobem els paràmetres estadístics obtinguts per al model final de l'estructura. En aquest cas els paràmetres que es fan servir per a controlar l'adequació de l'afinament són els valors de R_{work} i R_{free} que han de disminuir en cada cicle per a veure que el model s'està afinant correctament. Així amb aquest afinament es va passar d'uns $R_{\text{work}} \sim 26$ $R_{\text{free}} \sim 30$ pels primers cicles inicials abans d'afegir les aigües a uns $R_{\text{work}} = 19,9$ $R_{\text{free}} = 23,4$ pel model final.

All-Atom Contacts	Clashscore, all atoms:	5.31	97 th percentile* (N=715, 2.00Å ± 0.25Å)
	Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.		
Protein Geometry	Poor rotamers	0	0.00% Goal: <1%
	Ramachandran outliers	0	0.00% Goal: <0.05%
	Ramachandran favored	149	100.00% Goal: >98%
	MolProbity score [^]	1.28	99 th percentile* (N=12522, 2.00Å ± 0.25Å)
	Cβ deviations >0.25Å	0	0.00% Goal: 0
	Bad backbone bonds:	0 / 1286	0.00% Goal: 0%
	Bad backbone angles:	0 / 1704	0.00% Goal: <0.1%
Nucleic Acid Geometry	Bad bonds:	0 / 904	0.00% Goal: 0%
	Bad angles:	1 / 1388	0.07% Goal: <0.1%

In the two column results, the left column gives the raw count, right column gives the percentage.

* 100th percentile is the best among structures of comparable resolution; 0th percentile is the worst. For clashscore the comparative set of structures was selected in 2004, for MolProbity score in 2006.

[^] MolProbity score combines the clashscore, rotamer, and Ramachandran evaluations into a single score, normalized to be on the same scale as X-ray resolution.

Fig. 4.95. Sortida del programa Molprobity [114] obtinguda per al model final del cristall AS32-3.

En la taula es mostren resumits els paràmetres analitzats pel programa, com ara la puntuació per col·lisions entre àtoms, els rotàmers no preferents, l'adequació dels angles i distàncies d'enllaç o la distribució dels angles de torsió dins del diagrama de Ramachandran. Observem que tots els paràmetres són acceptables pel programa.

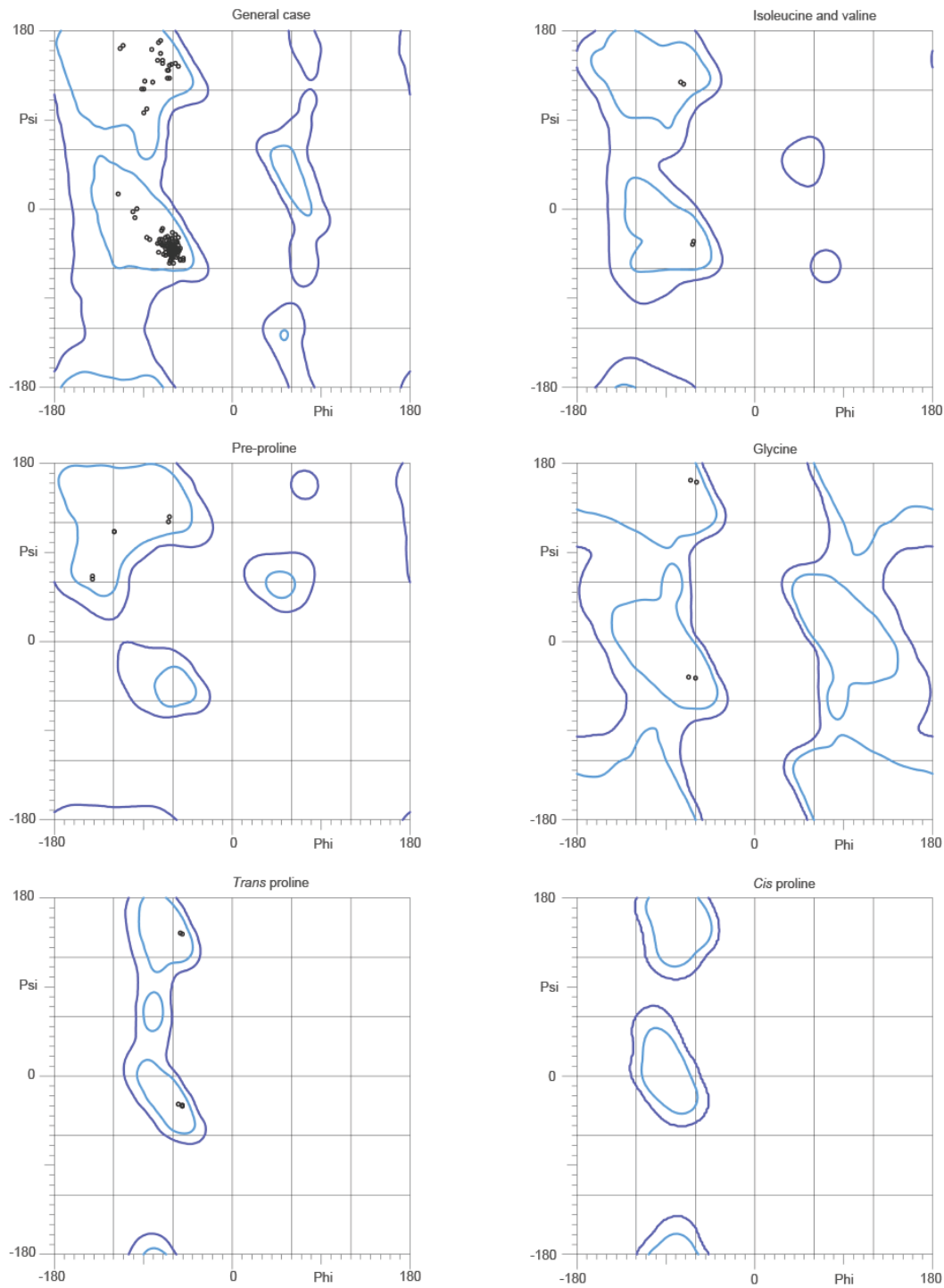


Fig. 4.96. Ramachandran per al model final.

Aquests diagrames de *Ramachandran* s'han obtingut amb el programa *Molprobability* [114][133]. Observem que els 153 residus de l'estructura (150 residus més les tres conformacions addicionals) es troben en regions preferents.

Per últim, el model final es va dipositar al *Protein Data Bank* amb el codi d'entrada 4QR9.

4.5. Estructura del complex HMGB1 Box AS-DNA

4.5.1. Estructura general del cristall Box AS-ATATCGATAT

La unitat asimètrica del cristall AS32-3 conté dos dúplex de l'oligonucleòtid ATATCGATAT i dues molècules corresponents al domini Box AS de la proteïna HMGB1 (fig. 4.97 a). Tots dos dominis Box AS s'uneixen a un dels dúplexs (cadena CD), provocant un important doblegament d'aquest d'uns 85° mentre que l'altre dúplex (cadena EF) no fa a penes contactes amb la proteïna i roman recte (v. fig. 4.97 b). Els dúplexs doblegats i els rectes formen una hèlix continua al llarg del cristall dibuixant un patró en forma de ziga-zaga (v. fig. 4.97 a). L'*helical twist* entre els parells de bases d'una hèlix i de la següent és de -25° .

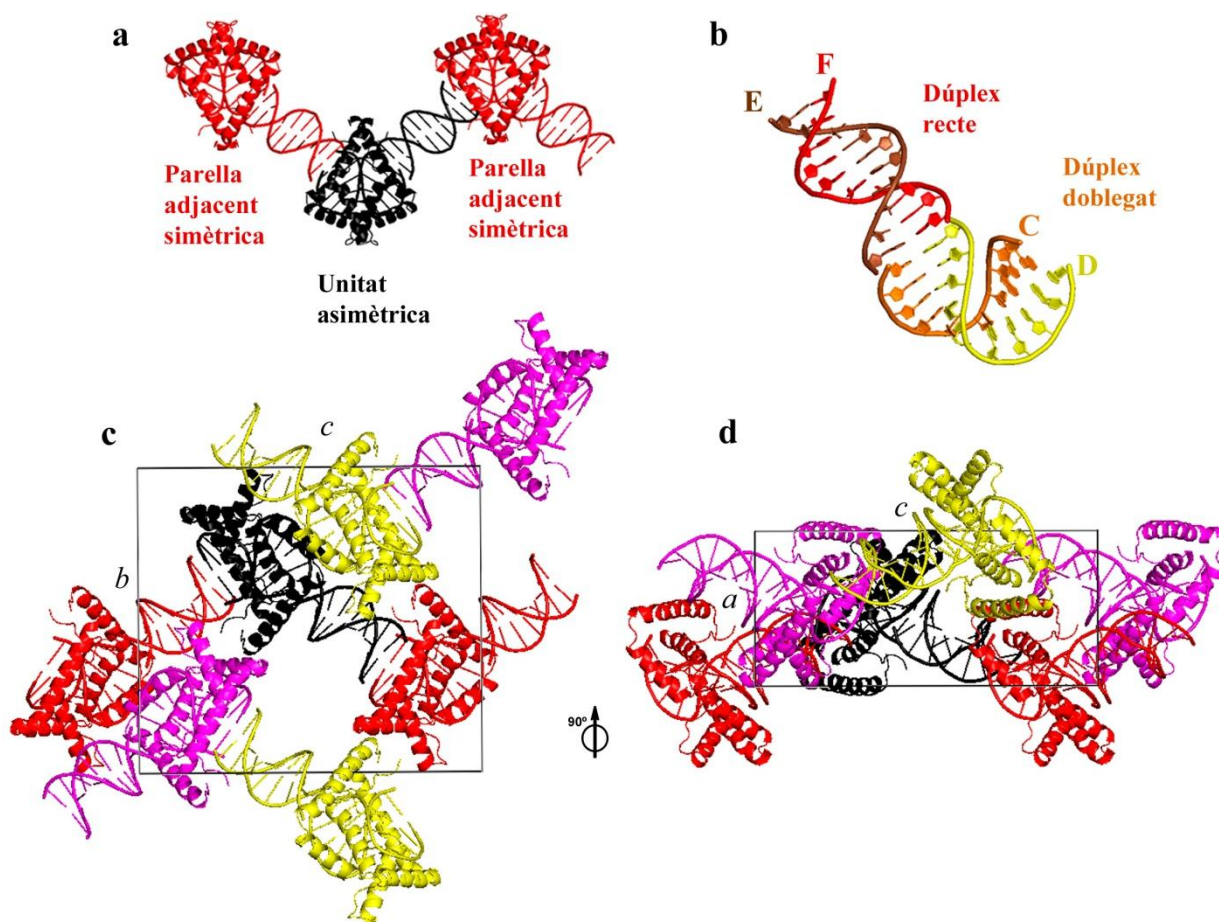


Fig. 4.97. Organització dins de la cel·la del complex Box AS-DNA.

- Unitat asimètrica (negre) formada per dos dúplexs i dos dominis Box AS i dues veïnes seves adjacents (vermell).
- Els dos dúplexs de la unitat asimètrica: un de recte (cadena EF, marró i vermell) i un de doblegat (cadena CD, taronja i groc).
- Vista superior (vista del pla *cb*) mostrant el contingut de la cel·la: una unitat asimètrica en negre amb les seves unitats asimètriques veïnes relacionades amb ella pels tres eixos binaris helicoidals del grup espacial (en vermell, groc i lila). Per a tenir una visió més clara de l'empaquetament s'han representat un altre conjunt vermell, un altre groc i un altre lila de les cel·les adjacents.
- Vista lateral (vista del pla *ac*) de c).

A partir de dos dúplexs i dos Box AS de la unitat asimètrica (en negre a la fig. 4.97 c i d) es generen tres cops més aquest contingut a la cel·la unitària degut a les operacions de simetria dels tres eixos helicoidals del grup espacial ($P2_12_12_1$) (en groc, vermell i lila). Així en la cel·la obtenim 8 dúplexs del decàmer ATATCGATAT i 8 dominis Box AS. Els oligonucleòtids rectes queden posicionats a les diagonals internes de la cel·la. Aquestes dues direccions són les corresponents als dos *stackings* creuats que observàvem als diagrames de difracció.

Els dos Box AS s'uneixen a l'oligonucleòtid palindròmic ATATCGATAT d'una forma aproximadament simètrica (v. fig. 4.98). Un dels dominis Box AS (cadena A) contacta amb una meitat del dúplex mentre que l'altre (cadena B) amb l'altra meitat. Les proteïnes interaccionen amb el DNA provocant un desenrotllament i un considerable doblegament. Les fenilalanines 37 de tots dos dominis s'intercalen entre els parells de bases CG centrals, quedant localitzades en el sot que es crea en el solc menor del DNA (v. fig. 4.98 d). Podem observar com els dos dominis interaccionen d'una forma cua-cua com era el cas de la proteïna TFAM (veure apartat 1.1.2), però en cap altra estructura d'HMG-box en complex amb DNA s'havia observat una intercalació de dos residus corresponents al segon lloc d'intercalació en el mateix pas de parell de bases.

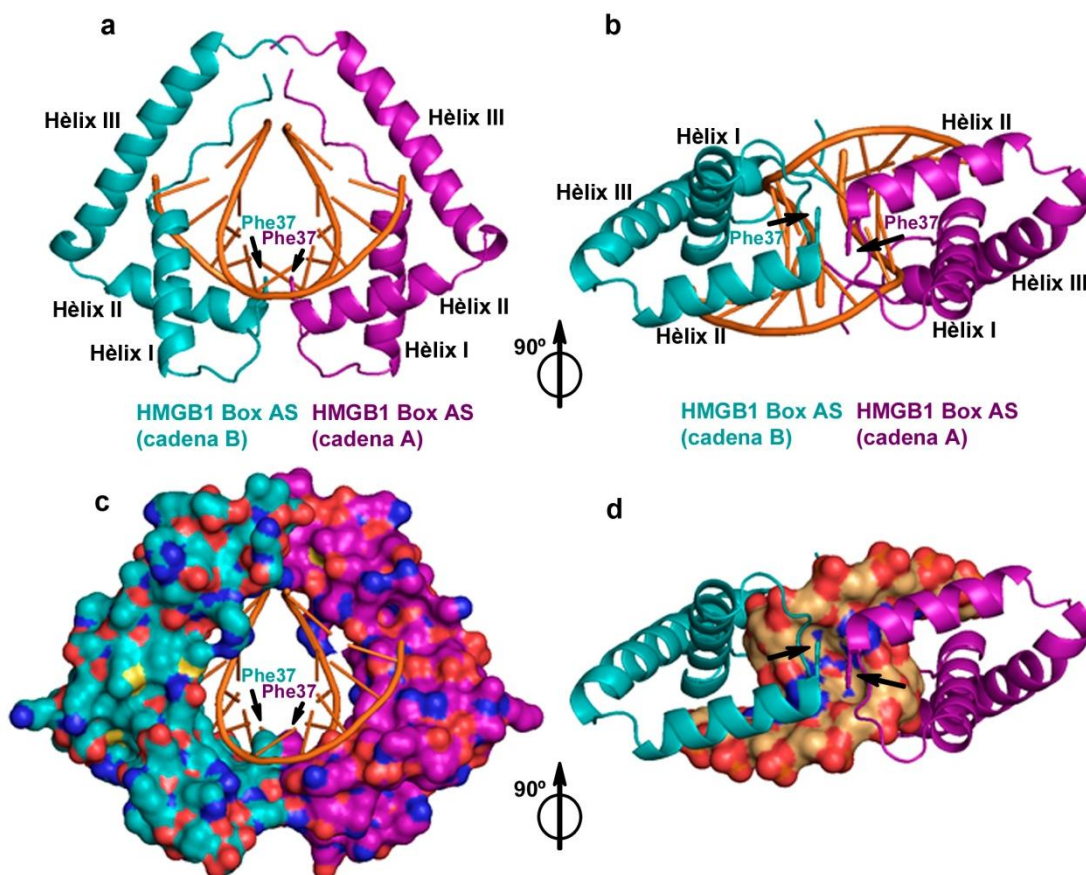


Fig. 4.98. Interacció de dos dominis Box AS per a doblegar el DNA.

a) Vista lateral dels dos dominis, cadena A (lila) i cadena B (cian), amb l'oligonucleòtid en taronja. b) Vista inferior del complex. c) Representació en superfície de les proteïnes. d) Representació en superfície del DNA on s'observa el sot que ocupen les dues Phe37. En tots els casos s'indiquen les dues fenilalanines amb fletxes.

A la fig. 4.99 es mostra el mapa de densitat electrònica obtingut per aquest model. Podem veure que amb aquesta resolució de 2 Å s'observen molt clarament els perfils de les bases del DNA i les molècules d'aigua per exemple. A la figura superior es mostra l'oligonucleòtid doblegat amb les dues fenilalanines entre les bases C₁₅ i G₁₆ i a la inferior un magnesi hexahidratat que s'ha localitzat a la part còncava del DNA doblegat.

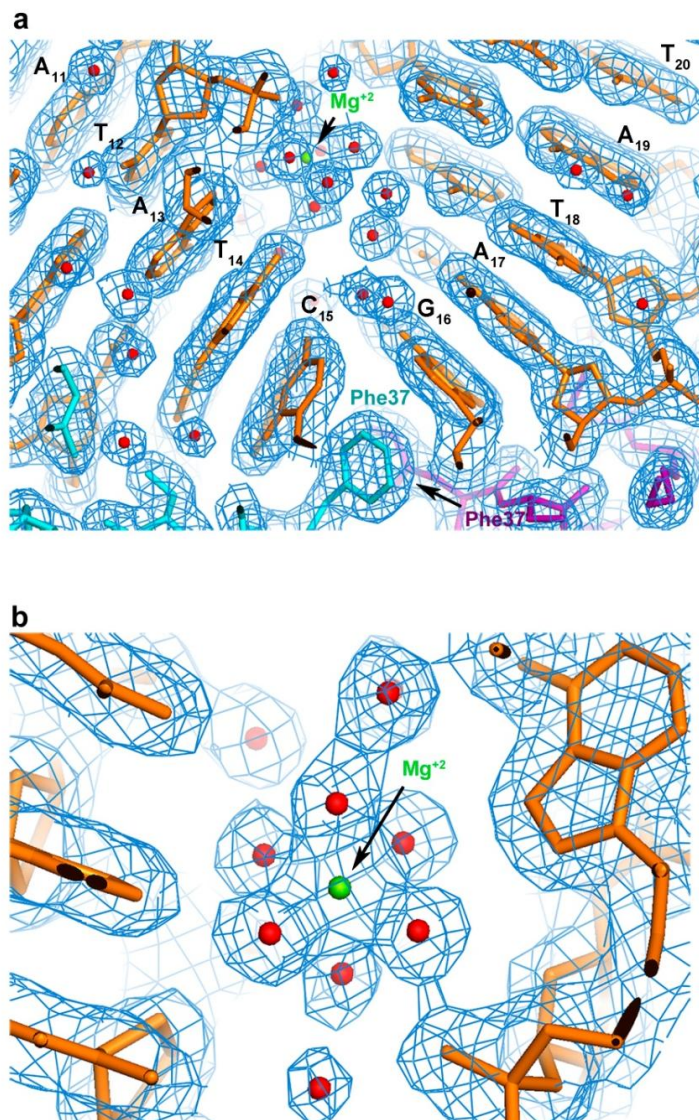


Fig. 4.99. Mapes de densitat electrònica de l'estructura.

a) Vista del doblegament del DNA amb les Phe intercalades. b) Detall de l'ió magnesi hexahidratat.

A més a més s'ha trobat una cadena d'aigües que connecten amb ponts d'hidrogen els *loops* entre l'hèlix I i l'hèlix II dels dos Box AS (v. fig. 4.100).

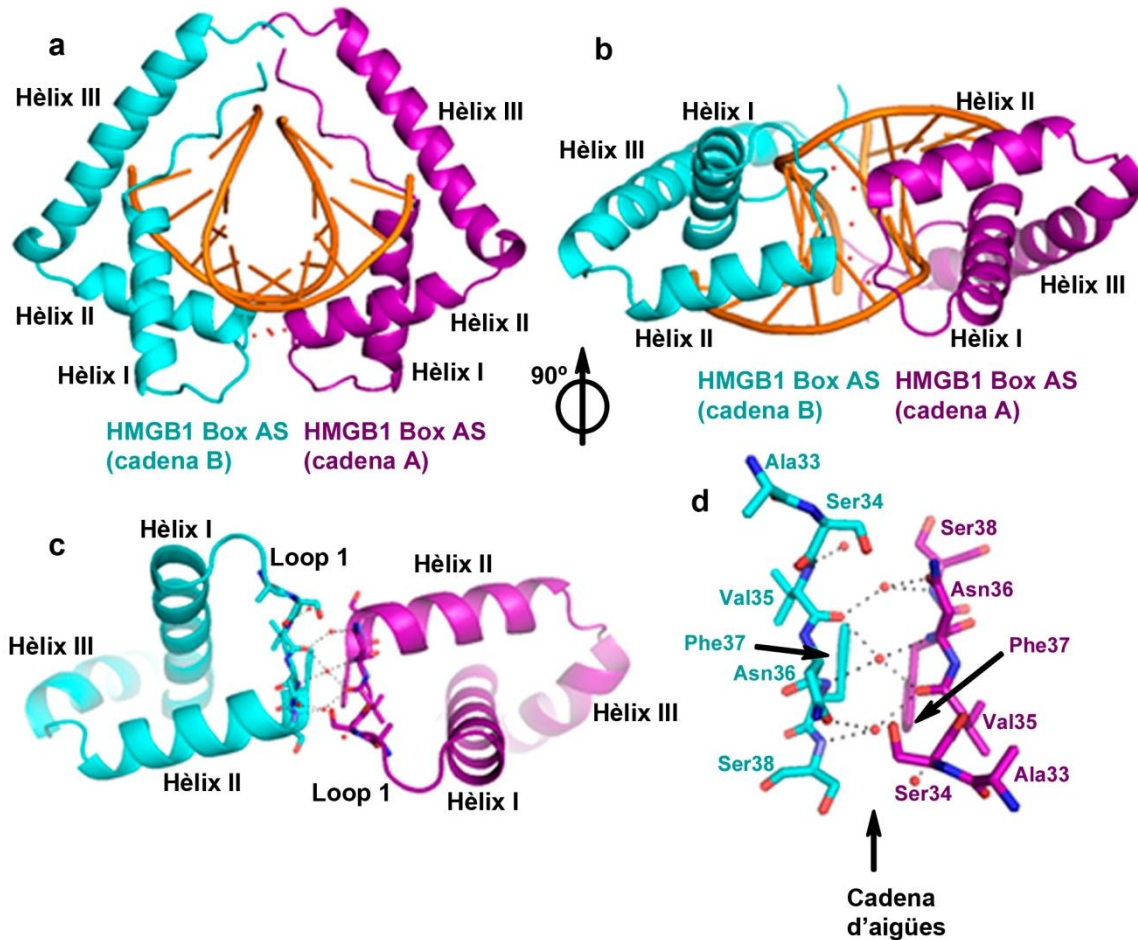


Fig. 4.100. Interaccions entre els dos dominis Box AS a través d'una cadena d'aigües.

- a) Vista lateral dels dos dominis Box AS (lila i cian) amb l'oligonucleòtid (taronja) i les aigües connectant els dos dominis representades com esferes vermelles. b) Vista inferior del complex. c) Representació (com línies discontinues) dels punts d'hidrogen entre les aigües. d) Detall de c) indicant els residus que intervenen en la superfície de dimerització entre els dos dominis Box AS.

4.5.2. Interaccions proteïna-DNA

La unió proteïna-DNA està establitzada per punts d'hidrogen directes, punts d'hidrogen mediat per aigües i forces de van der Waals (v. fig. 4.101 a). Les interaccions de les dues Box AS amb el DNA d'aquesta estructura comparteixen algunes característiques amb d'altres complexos HMG-box-DNA. El braç curt del Box AS juntament amb la cadena N-terminal estesa es posicionen en el solc menor del DNA fent contactes amb aquest. Concretament, la regió N-terminal de l'hèlix I i tota l'hèlix II d'un domini Box AS fan contactes amb una de les cadenes del DNA mentre que la part central de l'hèlix I i la cadena N-terminal estesa ho fan amb l'altra cadena. L'hèlix III, però, no interacciona amb el DNA.

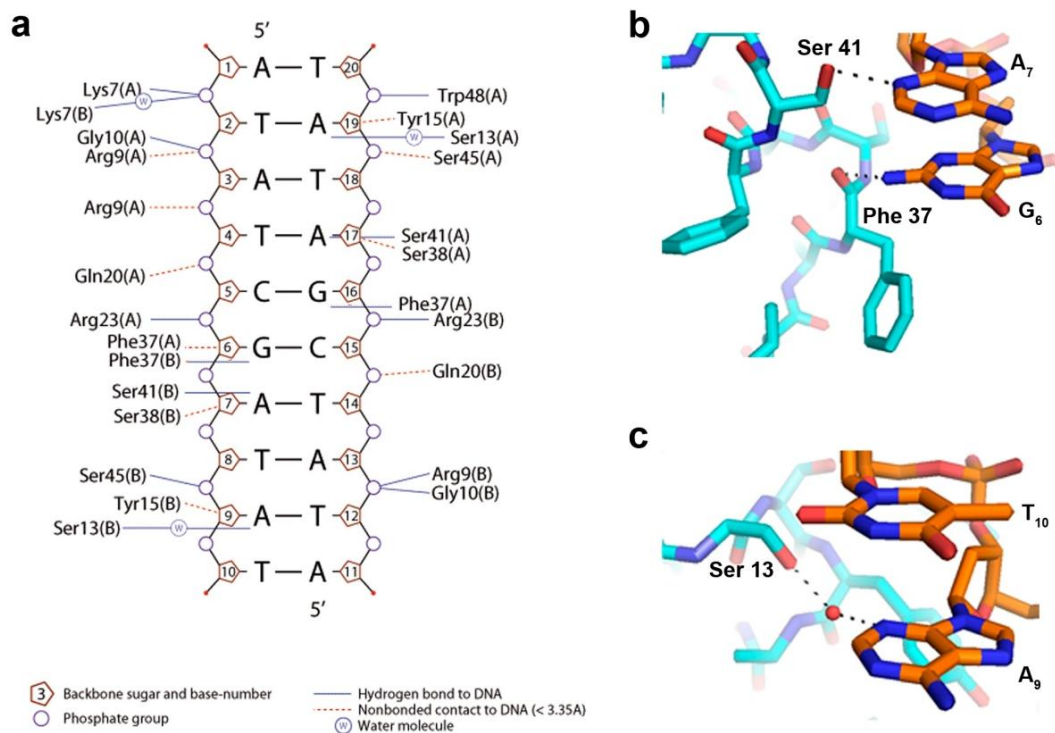


Fig. 4.101. Interaccions entre la proteïna i el DNA.

a) Contactes proteïna-DNA representats amb el programa *Nucplot* [134]. S'han representat únicament els contactes entre residus de la proteïna ((A), cadena A i (B), cadena B) i el DNA (les línies contínues equivalen a ponts d'hidrogen mentre que les discontinues equivalen a interaccions de van der Waals). b-c) Vistes en detall dels ponts d'hidrogen amb les bases del DNA que estableixen els residus Ser41, Phe37 (b) i Ser13 (c).

En el complex Box AS-DNA trobem tres ponts d'hidrogen (dos d'ells directes i un mediat amb una molècula d'aigua) entre el domini Box AS i les bases del DNA mentre que el nombre de ponts d'hidrogen amb l'esquelet sucre-fosfat és més abundant (v. fig. 4.101 a). L'hidrogen del grup hidroxil de la cadena lateral de la serina 41 (de la cadena B) forma un pont d'hidrogen amb el N3 de l'adenina 7, i l'oxigen del grup amida de la cadena principal de la fenilalanina 37 (de la cadena B) el fa amb l'hidrogen del N2 de la guanina 6 (v. fig. 4.101 b).

Els residus equivalents per a les proteïnes amb especificitat de seqüència SRY (Ser38/Ser41) [41] i LEF-1 (Ser37/Asn41) [42] es creu que contribueixen a aquesta especificitat formant diferents ponts d'hidrogen directes amb les bases del DNA. També es pensa que el residu 13 està implicat en l'especificitat de seqüència, sent asparagina en les proteïnes amb especificitat i serina en les de especificitat limitada. En el nostre cas la serina 13 (de la cadena B) fa un pont d'hidrogen mediat amb una molècula d'aigua amb el N3 de l'adenina 9 (v. fig. 4.101 c). Aquesta mateixa interacció no específica també s'ha observat per als residus equivalents de l'HMGD [45] i del Box B [44], però no en el Box A de l'estructura amb el DNA modificat amb cisplatí, on hi ha una citosina en aquesta posició. A més a més, justament aquesta Ser13 juntament amb la Ser41 i la Tyr15 participen en una xarxa de ponts d'hidrogen entre aigües que interconnecten les bases A₃, T₄ i C₅.

En el cas dels ponts d'hidrogen amb l'esquelet sucre-fosfat del DNA són destacables els enllaços que formen l'Arg23 i el Trp48 amb dos fosfats de l'oligonucleòtid. Aquestes interaccions estan molt conservades al llarg de la família de les HMG-box trobant-se també en les estructures resoltes en complex amb DNA de la quimera SRY-Box B [44], el Box A del cisplatí-DNA [43], l'HMGD [45], l'NHP6A [44], el SRY [41] i el LEF-1 [42]. De fet, juntament amb l'alanina 63, aquests dos residus (arginina i triptòfan) són els únics conservats en aquestes proteïnes (v. fig. 1.12 a l'apartat 1.1.2). Totes aquestes interaccions es troben de forma simètrica també a la cadena A de la proteïna Box AS amb les corresponents bases simètriques.

4.5.3. Estructura del DNA

El dúplex on interaccionen els dos dominis Box AS es troba doblegat uns 85° fent un colze en les bases centrals CG observant-se aquí un solc estret més ample i un desenrotllament de la doble hèlix (v. taula 4.12). Concretament en el pas dels parells de bases C₅G₆/C₁₅G₁₆ observem un valor per al *roll* de 74,84°, per al *twist* de 4,86° i per al *rise* de 6,65 Å. Els valors promitjos del *roll*, *twist* i *rise* són de 0,6°, 36,0° i 3,32 Å per al B-DNA estàndard i de 8,0°, 31,1° i 3,32 Å per a l'A-DNA estàndard en estructures cristal·litzades a alta resolució [63].

La unió de la proteïna provoca una desviació de la forma B del DNA i observem en aquest oligonucleòtid tan distorsionat trets característics de la forma A. Per exemple, el solc menor es troba eixamplat a expenses del solc major (v. taula 4.12). També observem un menor angle de torsió delta relacionat amb uns plegaments del sucre del tipus C3' *endo* en les bases A₃, T₄, C₅ i simètricament A₁₃, T₁₄ i C₁₅ (v. taula 4.13). En contrast, l'oligonucleòtid recte on la proteïna no es troba unida presenta unes característiques pròpies de B-DNA (valors no mostrats).

Taula 4.12. Paràmetres geomètrics dels parells de bases i amplada dels solcs del dúplex doblegat.

Pas	Amplada del solc menor (Å)	Amplada del solc major (Å)	Rise (Å)	Roll (°)	Twist (°)
A ₁ T ₂ /A ₁₉ T ₂₀	-	-	3,15	0,47	29,97
T ₂ A ₃ /T ₁₈ A ₁₉	-	-	3,46	3,76	35,92
A ₃ T ₄ /A ₁₇ T ₁₈	17,7	20,9	3,26	8,43	27,48
T ₄ C ₅ /G ₁₆ A ₁₇	20,6	17,0	3,07	7,89	24,81
C ₅ G ₆ /C ₁₅ G ₁₆	22,4	14,9	6,65	74,84	4,86
G ₆ A ₇ /T ₁₄ C ₁₅	20,6	17,1	3,06	7,97	24,84
A ₇ T ₈ /A ₁₃ T ₁₄	17,8	21,3	3,17	8,36	27,28
T ₈ A ₉ /T ₁₂ A ₁₃	-	-	3,43	2,53	35,20
A ₉ T ₁₀ /A ₁₁ T ₁₂	-	-	3,27	0,57	27,00

Valors del *rise*, el *roll* i el *twist* i amplada del solc menor i major (equivalent a la distància entre fosfats) per al dúplex doblegat (cadena CD) obtinguts amb el programa 3DNA [62].

Taula 4.13. Paràmetres geomètrics de l'esquelet sucre-fosfat del dúplex doblegat.

Pas	<i>Delta</i>	<i>Èpsilon</i>	<i>Zeta</i>	Conformació del fosfat	Plegament sucre
A₁/T₂₀	149/94	-174/-	-83/-	BI/BI	C3' <i>exo</i> /C4' <i>exo</i>
T₂/A₁₉	113/124	-176/-174	-106/-100	BI/BI	C1' <i>exo</i> /C1' <i>exo</i>
A₃/T₁₈	78/90	-168/-176	-77/-91	BI/BI	C3' <i>endo</i>/O4' <i>endo</i>
T₄/A₁₇	91/120	-144/178	-82/-92	BI/BI	C3' <i>endo</i>/C1' <i>exo</i>
C₅/G₁₆	83/138	-123/174	-70/-96	BI/BI	C3' <i>endo</i>/C2' <i>endo</i>
G₆/C₁₅	139/89	179/-122	-99/-77	BI/BI	C2' <i>endo</i>/C3' <i>endo</i>
A₇/T₁₄	115/88	176/-148	-88/-77	BI/BI	C1' <i>exo</i>/C3' <i>endo</i>
T₈/A₁₃	91/73	-177/-169	-87/-72	BI/BI	O4' <i>endo</i>/C3' <i>endo</i>
A₉/T₁₂	103/98	-175/-179	-87/-96	BI/BI	O4' <i>endo</i>/O4' <i>endo</i>
T₁₀/A₁₁	95/158	-/-174	-/-86	BI/BI	O4' <i>endo</i>/C3' <i>exo</i>

Valors dels angles de torsió *delta*, *èpsilon* i *zeta* de l'esquelet. També s'indica la conformació del fosfat (calculada a partir de la diferència entre *èpsilon* i *zeta*) i el plegament del sucre. Es mostren els valors per a cada parell de bases separats per una /. Els plegaments del sucre del tipus C3' *endo*, propis de la forma A del DNA, es troben ressaltats en negreta. Els paràmetres de la taula s'han calculat amb el programa 3DNA [62].

La distorsió en el DNA de la nostra estructura és lleugerament superior a l'observada en el DNA distorsionat amb el *crosslink* produït a les bases GG amb cisplatí [43]. En aquesta última estructura la fenilalanina 37 es posiciona en el buit creat en el solc menor degut a la lesió produïda pel cisplatí unit a les bases GG distorsionant lleugerament el DNA prèviament doblegat, mentre que en la nostra estructura la intercalació de dos fenilalanines 37 de tots dos dominis en el pas CG d'un DNA lineal força el doblegament tan pronunciat del dúplex focalitzat justament entre aquestes dues bases. A la fig. 4.102a es mostra la superposició d'ambdós oligonucleòtids.

De fet, els valors d'r.m.s.d. (*root-mean-squared deviation*) per al dúplex de la nostra estructura i el DNA modificat amb cisplatí és de només 3,23 Å [43] (comparant tots els àtoms i fragments de la mateixa longitud) i de 2,59 Å comparant aquest últim amb un oligonucleòtid similar que conté l'adducte cisplatí però sense la proteïna [135]. Així, la unió de dos dominis Box AS crea una distorsió similar a la del DNA prèviament doblegat pel *crosslink* produït pel cisplatí.

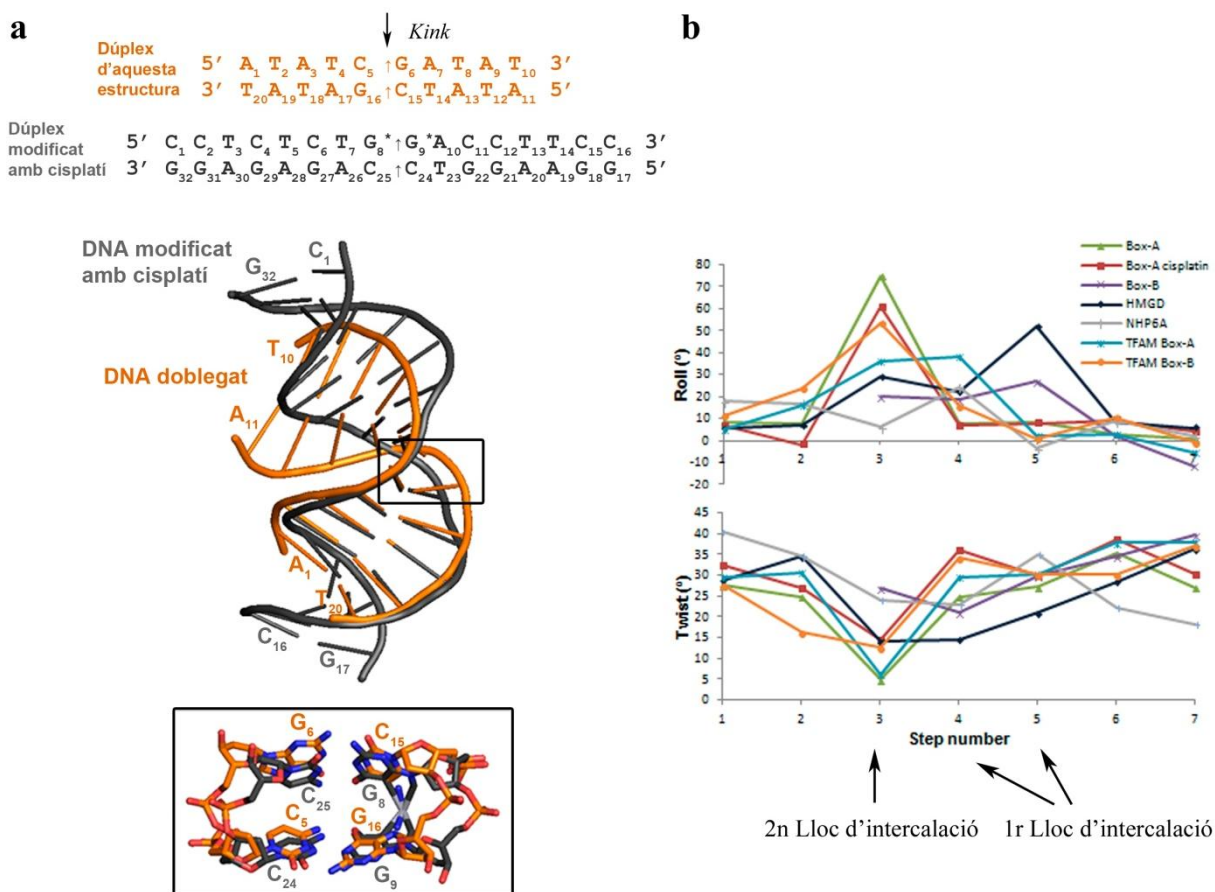


Fig. 4.102. Comparació amb altres oligonucleòtids interaccionant amb HMG-boxes.

a) Superposició del DNA de la nostra estructura amb el DNA modificat amb cisplatin en complex amb el Box A [43].

La superposició s'ha realitzat alineant les dues estructures amb el programa *Pymol*. Es mostren les seqüències nucleotídiques de tots dos oligonucleòtids i s'ha indicat amb fletxes els llocs d'intercalació. b) Gràfiques per als valors del *roll* i el *twist* en alguns complexos d'HMGBs. S'han representat els diferents valors per a aquests angles alineant les seqüències dels DNA en complex amb diferents HMG-boxes agafant com a referència el segon lloc d'intercalació. Per a la nostra estructura s'han calculat amb el programa *3DNA* mentre que per a les altres s'han usat els valors del *Nucleic Acids Data Bank* per als codis d'accés 1CKT, 2GZK, 1QRV, 1J5N i 3TMM.

Si comparem els paràmetres geomètrics *roll* i *twist* de les diferents estructures de DNA units a proteïnes HMG-box de limitada especificitat de seqüència resoltes fins al moment observem que els valors més allunyats del B-DNA en un pas (major *roll* i menor *twist*) són els de la nostra estructura (v. fig. 4.102 b). Així, la unió d'aquests dos dominis Box AS i la intercalació de les dues fenilalanines provoca el major doblegament i desenrotllament en un pas de parell de bases vist fins ara per a aquestes proteïnes.

És interessant veure que la distorsió es produeix en un únic punt, el corresponent al segon lloc d'intercalació del domini Box AS, mentre que en el primer lloc l'alanina no arriba a intercalar-se entre les bases del DNA. Així la distorsió del DNA es focalitza en únic punt presentant aquest una forma de colze (*kinked DNA*). En canvi, les altres proteïnes que presenten intercalació en tots dos llocs provoquen una distorsió en dos punts del DNA, com per exemple en la proteïna HMGD (v. fig. 4.102 b en blau marí), observant-se un doblegament més progressiu del dúplex de DNA (*bent DNA*) (v. també les seves corresponents estructures a la fig. 1.13 a l'apartat 1.1.2).

4.5.4. Comparació amb altres Box AS

L'estructura del domini Box A presenta apreciables diferències entorn a la fenilalanina 37 si aquest es troba unit a un DNA lineal (la nostra estructura) o a un DNA prèviament doblegat amb cisplatí [43]. L'r.m.s.d. global entre les dues proteïnes és d'1,68 Å i d'1,61 Å si comparem només les tres hèlixs alfa de les dues estructures (v. taula 4.14). Les diferències principals que s'observen entre les dues són que el *loop* entre l'hèlix I i l'hèlix II es troba posicionat de forma diferent mentre que l'hèlix I presenta una forma més estesa en la unió de la Box A al DNA modificat amb cisplatí (v. fig. 4.103 a). A més a més podem apreciar com l'orientació de l'anell de la fenilalanina està girat l'un respecte l'altre.

L'hèlix II d'aquestes dues estructures unides al DNA presenta una forma més recta en comparació amb les estructures dels dominis Box A lliures estudiades per RMN (v. fig. 4.103 b i c). Aquesta geometria específica més estesa del Box A facilitaria la interacció i la intercalació de la Phe37 amb la doble hèlix del DNA. Tot això ens dóna una idea de la possible flexibilitat conformacional d'aquest domini per adaptar-se a diferents entorns.

Comparació del Box A unit a un B-DNA d'aquesta estructura amb:

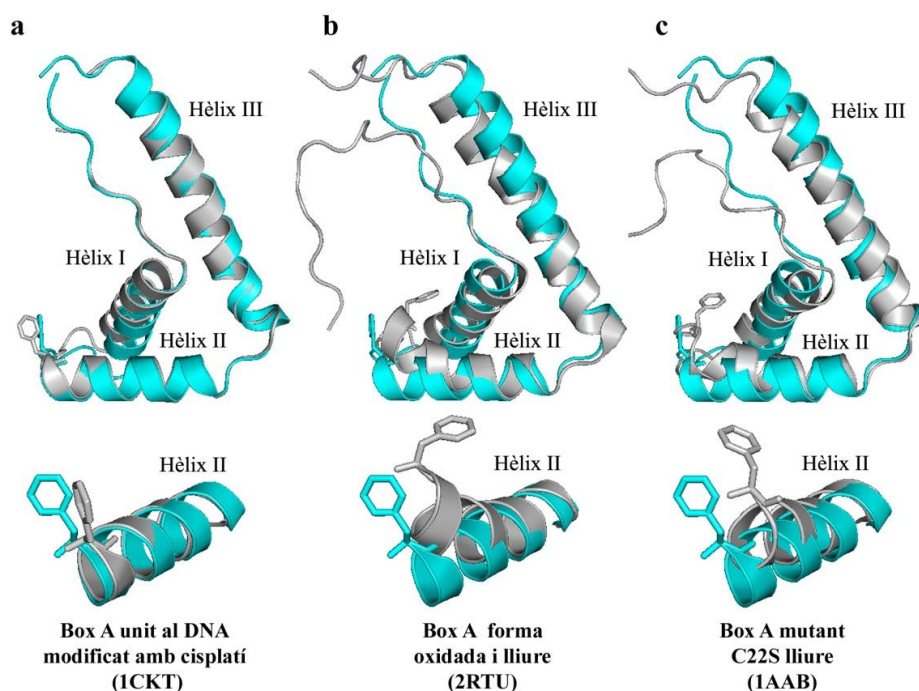


Fig. 4.103. Comparació entre diferents dominis Box A i les conformacions de les fenilalanines 37.

Superposició del Box A de la nostra estructura (cian) amb: a) el Box A corresponent a l'estructura de raigs X amb el DNA modificat amb cisplatí (gris) (codi pdb d'accés 1CKT) [43]. b) el Box A en la seva forma oxidada i lliure (gris) corresponent a l'estructura d'RMN amb codi d'accés al pdb 2RTU [136]. c) el mutant C22S Box A en la seva forma reduïda i lliure (gris) corresponent a l'estructura d'RMN amb codi d'accés al pdb 1AAB [32].

Taula 4.14. Valors dels r.m.s.d. per a les superposicions del Box A d'aquesta estructura amb d'altres.

Superposició de la Box AS (cadena B) amb:	Tots els àtoms ¹ (Å)			Hèlixs alfa ² (Å)		
	C α	Esquelet	Tot	C α	Esquelet	Tot
Box AS (cadena A) d'aquesta estructura	0,46	0,46	0,95	0,34	0,33	1,01
Box A unit al DNA modificat amb cisplatí (1CKT)	1,23	1,20	1,68	1,05	1,06	1,61
Forma lliure i oxidada del Box A (2RTU)	2,58	2,54	3,46	2,01	1,97	2,73
Mutant C22S Box A lliure (1AAB)	2,55	2,63	3,08	1,72	1,75	2,50

Els valors d' r.m.s.d. s'han calculat superposant la cadena B de la nostra estructura amb la cadena A, el box A de l'estructura del cisplatí [43], la forma oxidada i lliure de la Box A [136] i la forma reduïda i lliure [32]. S'ha fet servir el programa *Superpose* [137] per a realitzar les superposicions amb els carbonis alfa dels residus, la cadena principal polipeptídica (esquelet) i tots els àtoms de la molècula.

¹Del residu 7 al 77 segons la fig. 1.12. ²Residus 13-28, 38-50 i 53-75 de la fig. 1.12.

Se sap que les cisteïnes 22 i 44 del domini Box A poden formar un enllaç disulfur reversible en condicions d'oxidació suaus. Aquesta oxidació per una banda redueix l'afinitat d'unió al DNA [138] i a més a més està relacionada amb l'activitat inflamatòria de l'HMGB1 [139]. Concretament, l'estat oxidat de l'HMGB1 promouria l'activitat dels leucòcits per secretar citoquines proinflamatòries.

Recentment ha estat resolta per RMN l'estructura del domini Box A lliure en la seva forma oxidada (amb l'enllaç disulfur format entre cisteïnes) [136]. En aquesta estructura, en comparació al Box A de la nostra, l'hèlix II està doblegada cap a l'hèlix III, posicionant-se així la fenilalanina més lluny d'un possible lloc d'intercalació al DNA (v. fig. 4.103 b i fig. 4.104). També podem observar com la zona N-terminal de l'hèlix I està lleugerament més a prop de la zona C-terminal de l'hèlix II degut a la formació d'aquest enllaç. Així doncs aquesta disposició a l'espai on el residu responsable de la intercalació al DNA està voltejat cap a un altra banda explicaria la disminució en l'afinitat d'unió al DNA de la forma oxidada.

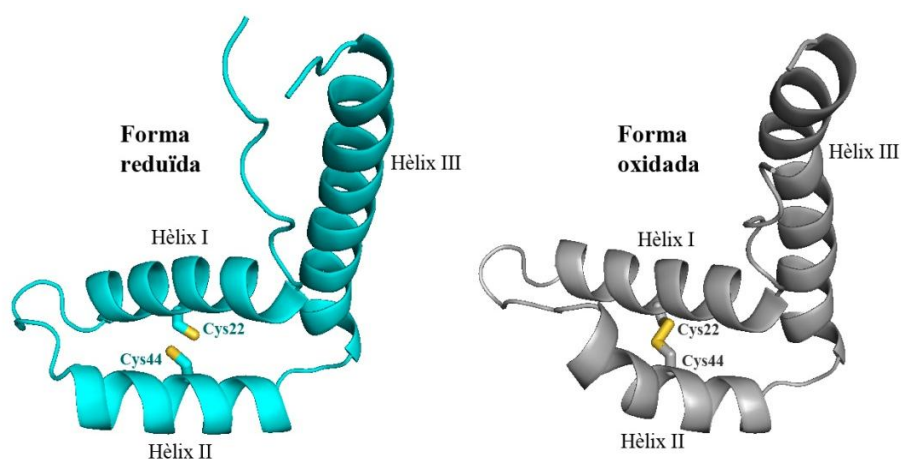


Fig. 4.104. Comparació entre la forma reduïda i l'oxidada del domini Box A.

S'ha representat en forma d'*sticks* les cisteïnes 22 i 44 del domini Box A de la estructura d'aquest treball (cian) i la forma oxidada resolta per RMN recentment (gris) (codi d'accés al pdb 2RTU [136]). Podem observar com la conformació de la cisteïna 22 ha canviat en la formació del pont disulfur Cys22-Cys44.

4.5.5. Conclusions

El domini Box A de la proteïna HMGB1 té una major afinitat per estructures del DNA no canòniques com DNA modificat amb cisplatí o *four-way junctions* que el Box B (revisat a [35]) i per a aquest motiu gran part dels estudis sobre aquest domini s'han centrat en la seva unió amb estructures distorsionades. Malgrat això, la unió del domini Box A a B-DNA és també de gran importància per arribar a entendre els mecanismes d'acció d'aquests factors arquitectònics de la cromatina. Mitjançant la difracció de raigs X d'un cristall del domini Box A amb un oligonucleòtid lineal i no modificat hem pogut estudiar la gran distorsió que provoca en aquest. Concretament l'acció de la proteïna ha provocat un doblegament d'uns 85°, un desenrotllament de la doble hèlix i en general una distorsió del DNA en la que observem trets característics de la forma A del DNA.

La majoria dels dominis HMG-box de les proteïnes HMGB amb especificitat de seqüència limitada o inexistent presenten dos residus no polars que s'intercalen al DNA provocant un doblegament d'entre 60° i 110°. En el cas del domini Box A de l'HMGB1, al tenir una alanina en la posició corresponent al primer lloc d'intercalació, només s'intercala al DNA la fenilalanina del segon lloc d'intercalació. Es creu que justament per a aquest motiu el Box A distorsiona en menor grau el DNA lineal. No obstant això, a l'estructura resolta en aquest treball hem vist que la unió de dos Box A, amb la corresponent intercalació d'ambdós fenilalanines en el mateix pas de parells de bases en el DNA, provoca una distorsió comparable i fins i tot major a les observades en un únic punt fins al moment. Entre les interaccions proteïna-DNA hem observat que el residu Ser41 i la cadena principal de la Phe37 fan ponts d'hidrogen directes amb les bases del DNA. La Ser13 fa un pont mediat amb una aigua amb una base del DNA mentre que la Lys7 i l'Arg9 fan ponts d'hidrogen directes amb l'esquelet sucre-fosfat del DNA; cap d'ells s'havien vist prèviament.

Resta pendent conèixer si aquesta interacció de tots dos dominis es realitza de forma cooperativa o simultània i si d'altres HMG-box són capaces d'interaccionar d'aquesta manera.

4.6. Estudi de la introducció de *supercoiling* negatiu

Una de les característiques de les proteïnes HMGB és l'efecte que poden produir sobre la superhelicitat del DNA [140]. Aquest efecte va ser estudiat ja l'any 1984, quan es va observar la diferent capacitat d'introduir *supercoiling* per part de l'HMGB1 i els seus dominis HMG-box per separat [141].

A partir de l'estructura cristal·lina obtinguda per al Box A hem vist el grau de deformació que produeix aquest al doblegar el DNA i desenrotllar la doble hèlix. Aquest desenrotllament provocat per les proteïnes HMGB està directament relacionat amb la introducció de superenrotllament negatiu en presència de topoisomerasa I.

En aquest treball hem volgut veure l'efecte en la topologia del DNA que produeixen els fragments AS i ABS (corresponents al Box A i al Box A+Box B de l'HMGB1 (v. fig. 4.12)). Per a veure la introducció de superenrotllament degut a l'acció de les proteïnes en DNAs circulars relaxats es porta a terme una reacció en presència de topoisomerasa I i s'observa el canvi en el grau de superenrotllament dels topoisòmers obtinguts en un gel d'agarosa (v. apartat 3.10.5).

A la figura següent es mostra el gel d'agarosa corresponent a les diferents reaccions d'un plasmidi circular de 4,6 kb (pSTAT-CEN2-S2 (v. fig. 4.110)) a concentracions creixents de Box ABS i Box AS en presència de topoisomerasa I.

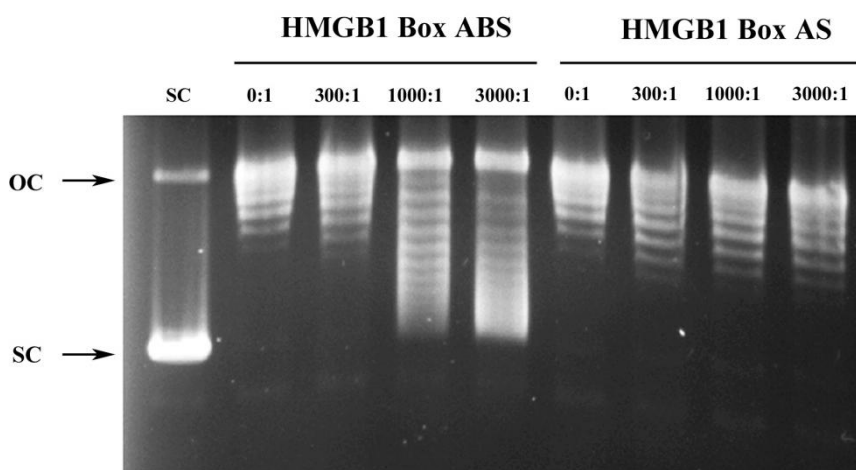


Fig. 4.105. Assaig d'introducció de *supercoiling* a alta força iònica (~140 mM).

SC és la mostra del DNA superenrotllat. A partir d'una ràtio molar de 1000:1 de Box ABS podem veure introducció de superenrotllament negatiu. Indicat amb fletxes es mostren el DNA superenrotllat (SC, de l'anglès *supercoiled*) i el DNA *nicked* (OC, de l'anglès *open circular*).

Podem observar com per a la proteïna HMGB1 Box ABS s'observen topoisòmers amb superenrotllament creixent a mesura que augmenta la concentració d'aquesta. En canvi, per al domini Box AS no s'observa cap canvi en el grau de superenrotllament del DNA circular.

Segons Stros *et al* [142] i Teo *et al* [143] la força iònica d'aquesta reacció és determinant en els assajos de *supercoiling* d'algunes proteïnes o fragments de les HMGB1 i han observat que el grau de superenrotllament és major a menor concentració de sal. Tenint això en compte, es va repetir l'assaig anterior però aquest cop amb una menor concentració de sal, disminuint de 100 a 10 mM la concentració de NaCl del tampó de reacció (v. taula 3.27). Efectivament, en aquestes condicions, el grau de superenrotllament és major però de nou s'observa com l'efecte de superenrotllament és superior en el cas de la proteïna Box ABS (fig. 4.106). Així, tot i que observem clarament una introducció de superenrotllament negatiu creixent per a concentracions creixents dels dos fragments proteics l'efecte del fragment ABS amb els dos dominis HMG-box (Box A i B), és aproximadament tres cops superior al produït pel domini Box A (podem tenir un grau de superenrotllament similar amb aproximadament tres cops menys de proteïna) (fig. 4.106).

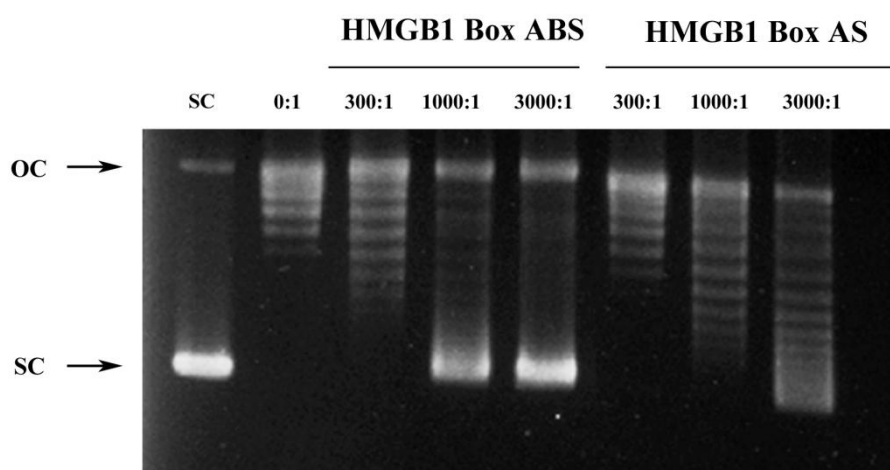


Fig. 4.106. Assaig d'introducció de *supercoiling* a baixa força iònica (~50 mM).

Indicat amb fletxes es mostren el DNA superenrotllat (SC, *supercoiled*) i el DNA *nicked* (OC, *open circular*).

En aquest l'assaig, el desenrotllament (*untwist*) provocat per la proteïna en el DNA circular relaxat ($\Delta Tw < 0$) és compensat amb superenrotllament positiu en una altra part de la molècula ($\Delta Wr > 0$). Aquest superenrotllament positiu és relaxat per la topoisomerasa ($\Delta Lk < 0$, $Wr = 0$). En eliminar la topoisomerasa juntament amb la proteïna de la barreja, s'obtenen topoisòmers amb superenrotllament negatiu ($Lk < 0$, $Wr < 0$, $\Delta Tw = 0$). A més a més pot haver-hi algun altre efecte de superenrotllament (Wr) produït per la proteïna degut al doblegament que provoca en el DNA. Per a confirmar l'obtenció de topoisòmers amb superenrotllament negatiu produïts principalment pel desenrotllament es van reproduir les mateixes reaccions a baixa força iònica però aquest cop es va desenvolupar l'electroforesi en presència de 0,6 $\mu\text{g/ml}$ cloroquina (fig. 4.107).

La cloroquina s'intercala en el DNA provocant en aquest superenrotllament positiu. Per tant si analitzem topoisòmers amb superenrotllament positiu en un gel d'agarosa en presència de cloroquina aquests presentaran encara més voltes de superhelicitat positives. En el cas contrari, si es tracta de topoisòmers amb superenrotllaments negatius, observariem topoisòmers amb menors superenrotllaments negatius i es podrien arribar a formar topoisòmers amb voltes de superhelicitat positives. En el nostre cas es tracta de topoisòmers amb superenrotllament negatiu ja que en el gel amb cloroquina observem com l'efecte del superenrotllament es veu disminuït (fig. 4.107).

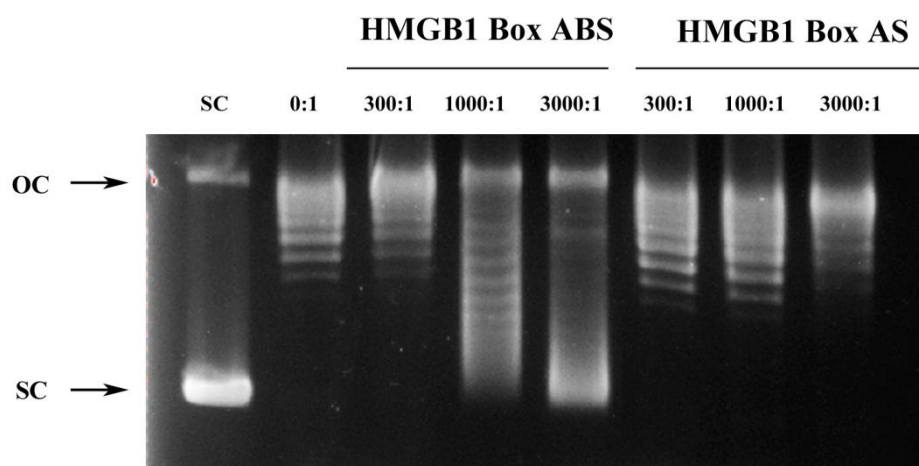


Fig. 4.107. Assaig d'introducció de *supercoiling* a baixa força iònica (~50 mM) i 0,6 µg/ml cloroquina. Indicat amb fletxes es mostren el DNA superenrotllat (SC, *supercoiled*) i el DNA *nicked* (OC, *open circular*).

Teo i col·laboradors [143] també observen introducció de *supercoiling* negatiu per a un fragment similar al nostre Box A però, mentre ells la detecten tant a alta com a baixa força iònica, en el nostre cas només en trobem a baixa força iònica. Carballo *et al* [141] i Nakamura *et al* [144] no obtenen tampoc increment de la superhelicitat treballant únicament a alta sal. Pel que fa al fragment contenint els dos dominis, Box ABS, mentre que Stros *et al* [142] observen el mateix perfil de bandes a alta o baixa força iònica, nosaltres veiem una variació en la introducció de superenrotllament negatiu a les dues forces iòniques.

Així, a més a més de la força iònica en la que tenen lloc les reaccions, la longitud del fragment del domini HMG-box amb què es treballa i les condicions específiques de l'experiment són factors determinants. No obstant això, podem confirmar la capacitat del domini Box A d'introduir superenrotllament negatiu pel desenrotllament de la doble hèlix del DNA, sent la seva acció incrementada amb la presència del domini Box B.

Hem realitzat també el mateix tipus d'assaig a alta força iònica per a la proteïna NHP6A. Al gel de la fig. 4.108 podem observar com a l'incrementar la concentració de la proteïna en la reacció amb la topoisomerasa incrementa la mobilitat de topoisòmers al gel i per tant el seu grau de superenrotllament. Si comparem aquest gel amb el corresponent per a la proteïna Box ABS a alta força iònica podem veure que el grau d'introducció de *supercoiling* és aproximadament del mateix ordre, sent una mica major per a l'NHP6A.

Això està en concordança amb els estudis de Paull i Johnson [145] que van determinar que l'efecte d'introducció de voltes de superhelicitat per a l'NHP6A és 2-3 cops major que per a l'HMGB1 sencera, produint aquesta menor introducció de superenrotllament que el fragment sense la cua acídica (Box A+Box B) [142].

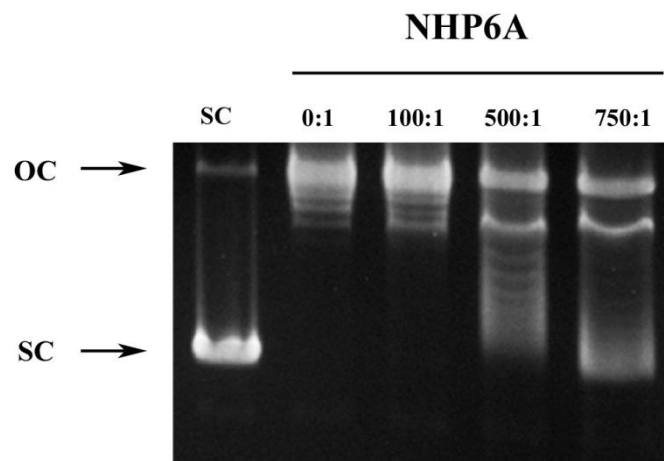


Fig. 4.108. Assaig d'introducció de *supercoiling* a alta força iònica (~140 mM) per l'NHP6A. Indicat amb fletxes es mostren el DNA superenrotllat (SC, *supercoiled*) i el DNA nicked (OC, *open circular*).

4.7. Estudis de la transició estructural de DNAs circulars amb inserts ATs

Degut al gran interès del nostre grup en les seqüències riques en ATs, en aquest treball s'ha volgut veure la formació d'estructures del tipus PX-DNA (v. fig. 4.109) en DNAs circulars contenint fragments amb homologia de seqüència d'ATs alternants. Amb aquesta finalitat s'han realitzat tant electroforesis bidimensionals en gels d'agarosa (apartat 4.7.1) com anàlisis de microscòpia de força atòmica (apartat 4.7.2) de plasmidis amb inserts AT. Els experiments d'electroforesis en gels d'agarosa s'han realitzat en col·laboració i amb l'assessorament del Dr. Joaquim Roca de l'Institut de Biologia Molecular de Barcelona (IBMB) - Consejo Superior de Investigaciones Científicas (CSIC). Els plasmidis amb els inserts ATs han estat construïts i facilitats per Xavier Fernández del grup del Dr. Joaquim Roca. Finalment, els experiments d'AFM s'han realitzat al Departament de Físicoquímica de la Facultat de Farmàcia de la UB amb la col·laboració del Dr. Oscar Domènech i el responsable del grup Dr. Jordi Hernández Borrell.

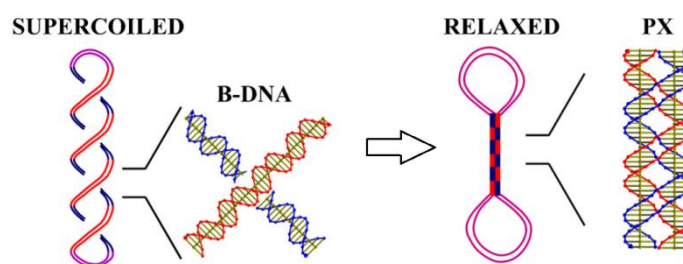


Fig. 4.109. Esquema de la formació d'un fragment del tipus PX-DNA en un DNA circular [68].

Per a més informació sobre aquestes estructures veure l'apartat 1.2.6.

S'ha descrit que una única volta de la forma PX amb homologia total de 22 nucleòtids és suficient per a aconseguir una visualització de la relaxació [68]. Per aquest motiu el plasmidi amb el que hem treballat (pSTAT-CEN2-S2) presenta dos inserts $(AT)_{11}$ en costats oposats. Com a controls negatius hem fet servir el plasmidi original (pSTAT-CEN2) i un plasmidi amb un únic insert (pSTAT-CEN2-S1). Tots tres plasmidis estan representats a la fig. 4.110.

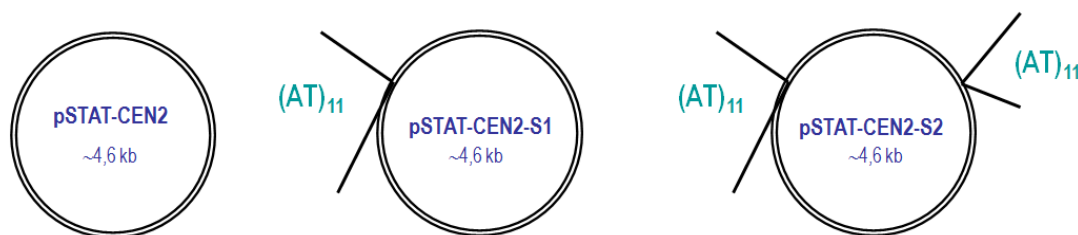


Fig. 4.110. Esquema dels plasmidis amb inserts ATs que s'han fet servir en aquest treball.

El plasmidi pSTAT-CEN2 no presenta cap insert, el plasmidi pSTAT-CEN2-S1 presenta un únic insert $(AT)_{11}$ i el plasmidi pSTAT-CEN2-S2 dos inserts $(AT)_{11}$ en localitzacions oposades.

4.7.1. Electroforesis bidimensionals

L'electroforesi en gels bidimensionals ens permet la visualització de transicions estructurals en DNAs circulars, en aquest cas la transició a la forma PX-DNA. Reproduint les mateixes condicions d'electroforesi que Seeman i col·laboradors [68], hem realitzat electroforesis bidimensionals pels plasmidis pSTAT-CEN2-S2 amb dos inserts (AT)₁₁ i el plasmidi original pSTAT-CEN2 com a control. Aquestes condicions electroforètiques s'expliquen en detall a l'apartat 3.10.4. El desenvolupament de la primera dimensió d'aquests gels bidimensionals s'ha realitzat amb un tampó TAE contenint 12,5 mM d'acetat de magnesi. Es creu que la presència d'ions Mg⁺² estabilitza els cruciforms del DNA [146].

El pas previ al desenvolupament d'una anàlisi electroforètica bidimensional d'aquest tipus és l'obtenció d'una mostra de topoisòmers amb un rang ampli de superenrotllaments. Amb aquest objectiu, els plasmidis han estat tractats amb topoisomerasa I a diferents concentracions de l'intercalant BrEt (bromur d'etidi) i a continuació precipitats amb etanol (v. apartat 3.10.3). Els topoisòmers resultants d'aquestes reaccions han estat visualitzats en gels d'agarosa petits (6 x 8 cm). Cada reacció produeix una distribució gaussiana de topoisòmers on el promig de superenrotllament depèn de la concentració d'intercalant. A la fig. 4.111 es mostren els gels amb les diferents reaccions de relaxació per als plasmidis pSTAT-CEN2 i pSTAT-CEN2-S2. En tots dos casos s'obtenen diferents mostres amb graus creixents de superenrotllament negatiu per a concentracions creixents de bromur d'etidi en les reaccions (d'entre 0 i ~1 µg/ml). La mostra SC (en el primer carril del gel) correspon a la mostra de partida del plasmidi superenrotllat (i per tant que no s'ha tractat amb topoisomerasa/BrEt).

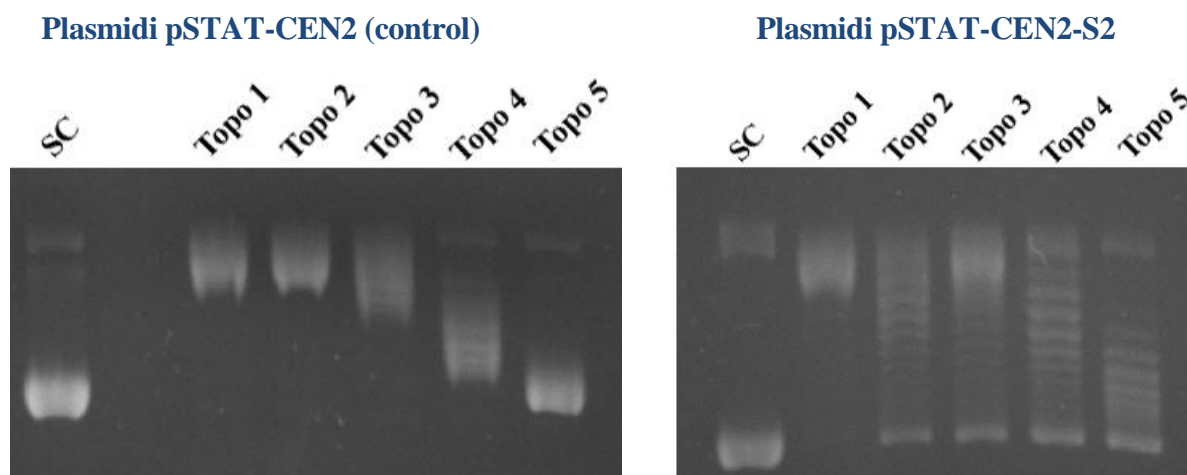


Fig. 4.111. Distribució de topoisòmers per a pSTAT-CEN2 (control) (esquerra) i pSTAT-CEN2-S2 (dreta).

Topo 1, 2, 3, 4 i 5 corresponen a les mostres obtingudes en les reaccions amb topoisomerasa I en presència de concentracions creixents de bromur d'etidi. SC és la mostra del DNA superenrotllat.

S'ha realitzat una barreja dels volums adients de cada reacció per a aconseguir una mostra contenint 1,2 µg de DNA circular amb diferents graus de superenrotllament per a carregar en un gel bidimensional gran (20 x 20 cm). Després de desenvolupar l'electroforesi en la primera dimensió (D1) amb el tampó 1x TAEMg i en la segona dimensió (D2) amb 1x TBE + 1,5 µl/ml cloroquina s'ha visualitzat el gel un cop tenyit amb BrEt. El resultat de l'electroforesi per al plasmidi pSTAT-CEN2-S2, que conté els dos inserts, es mostra a la fig. 4.112 (a). Podem observar clarament com la distribució de topoisòmers segueix el perfil de taques típic d'aquest tipus d'anàlisi electroforètiques però no observem cap discontinuïtat en aquest. Segons aquest resultat no ha tingut lloc la transició estructural per a passar d'un DNA circular superenrotllat a una estructura del tipus PX-DNA.

Per tal de confirmar que la distribució de topoisòmers s'ha realitzat correctament i que no hi falten topoisòmers on s'hauria de produir el salt, hem realitzat una electroforesi bidimensional en la qual s'ha afegit cloroquina tant en la primera dimensió (1x TBE + 1,5 µg/ml cloroquina) com a la segona dimensió (1x TBE + 8 µg/ml cloroquina). D'aquesta manera el perfil de taques canvia: en la primera dimensió els topoisòmers que presenten superenrotllament positiu, i per tant encara més positiu a la segona, queden a la dreta de r1. Els topoisòmers que presenten majors superenrotllaments negatius a D1, queden a la esquerra de r1 i es veuen millor separats en aquestes condicions d'electroforesi (veure fig. 4.112 (b)), confirmant així que la distribució de topoisòmers està completa.

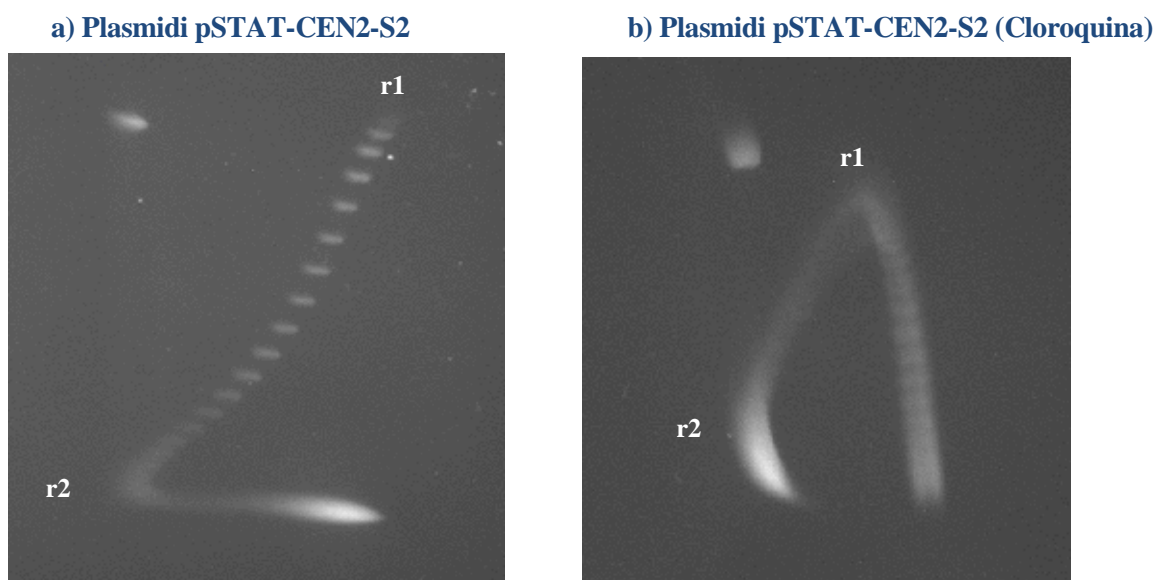


Fig. 4.112. Electroforesi bidimensional per a una distribució de topoisòmers de pSTAT-CEN2-S2 sense (a) i amb cloroquina (b) a la primera dimensió.

La posició r1 correspon al topoisòmer que es troba relaxat en la primera dimensió, i r2 al que es troba en la forma relaxada en la segona dimensió. a) Gel desenvolupat en condicions normals (1x TAEMg (D1) i 1x TBE + 1,5 µl/ml cloroquina (D2)). b) S'ha afegit 1,5 µg/ml cloroquina en D1 i 8 µg/ml cloroquina en D2 per observar millor la separació dels topoisòmers amb major superenrotllament negatiu.

Hem repetit les condicions de la primera electroforesi bidimensional (TAEMg//TBE+cloroquina) (les de l'article del Dr. Seeman [68]) però aquest cop carregant en el mateix gel una distribució de topoisòmers del plasmidi control i del plasmidi amb els dos inserts. Per al plasmidi amb els dos inserts podem observar el mateix resultat que a la fig. 4.113 (a) on no s'observa cap salt en el perfil de taques. Per al plasmidi control tampoc s'observa cap discontinuïtat (v. fig. 4.113 (b)), com és d'esperar.

El perfil de taques del plasmidi control en aquest últim gel (fig. 4.113 (b)) hagués quedat millor amb major quantitat de topoisòmers amb major superenrotllament. S'ha repetit aquesta electroforesi realitzant la barreja dels topoisòmers de les diferents reaccions de tal manera que quedin més topoisòmers amb majors Lk. A més a més s'ha afegit cloroquina en la primera dimensió (1x TAEMg + 0,6 µg/ml cloroquina) i major concentració en la segona (1x TBE + 3 µg/ml cloroquina) (v. fig. 4.113 (c) i (d)).

a) Plasmidi pSTAT-CEN2 (control)

b) Plasmidi pSTAT-CEN2-S2



c) Plasmidi pSTAT-CEN2 (control) (Cloroq.)

d) Plasmidi pSTAT-CEN2-S2 (Cloroq.)

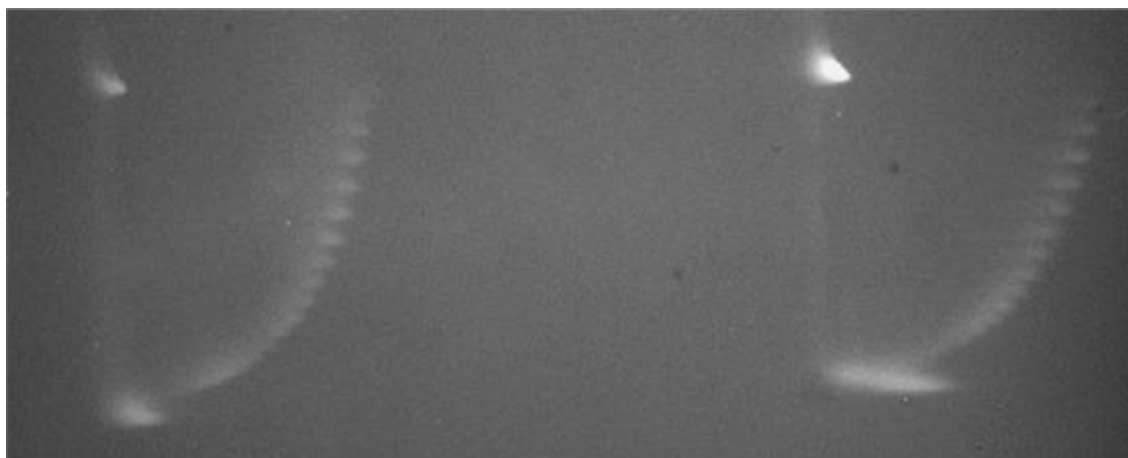


Fig. 4.113. Electroforesi bidimensionals per pSTAT-CEN2 (control) i pSTAT-CEN2-S2 (2 inserts) sense (a, b) i amb cloroquina (c, d) a la primera dimensió.

Per a poder assegurar-nos que la no observació d'una transició estructural a una estructura PX-DNA en el plasmidi amb dos inserts (AT)₁₁ reflectís la situació real, vam considerar important dur a terme un control positiu, és a dir, realitzar el mateix experiment utilitzant un plasmidi on estigués provat que es dona aquesta transició estructural. Amb aquesta finalitat vam contactar amb el Dr. Seeman, el qual gentilmente ens va facilitar els plasmidis pUC-19 220A i pUC-19 22A (v. fig. 1.27).

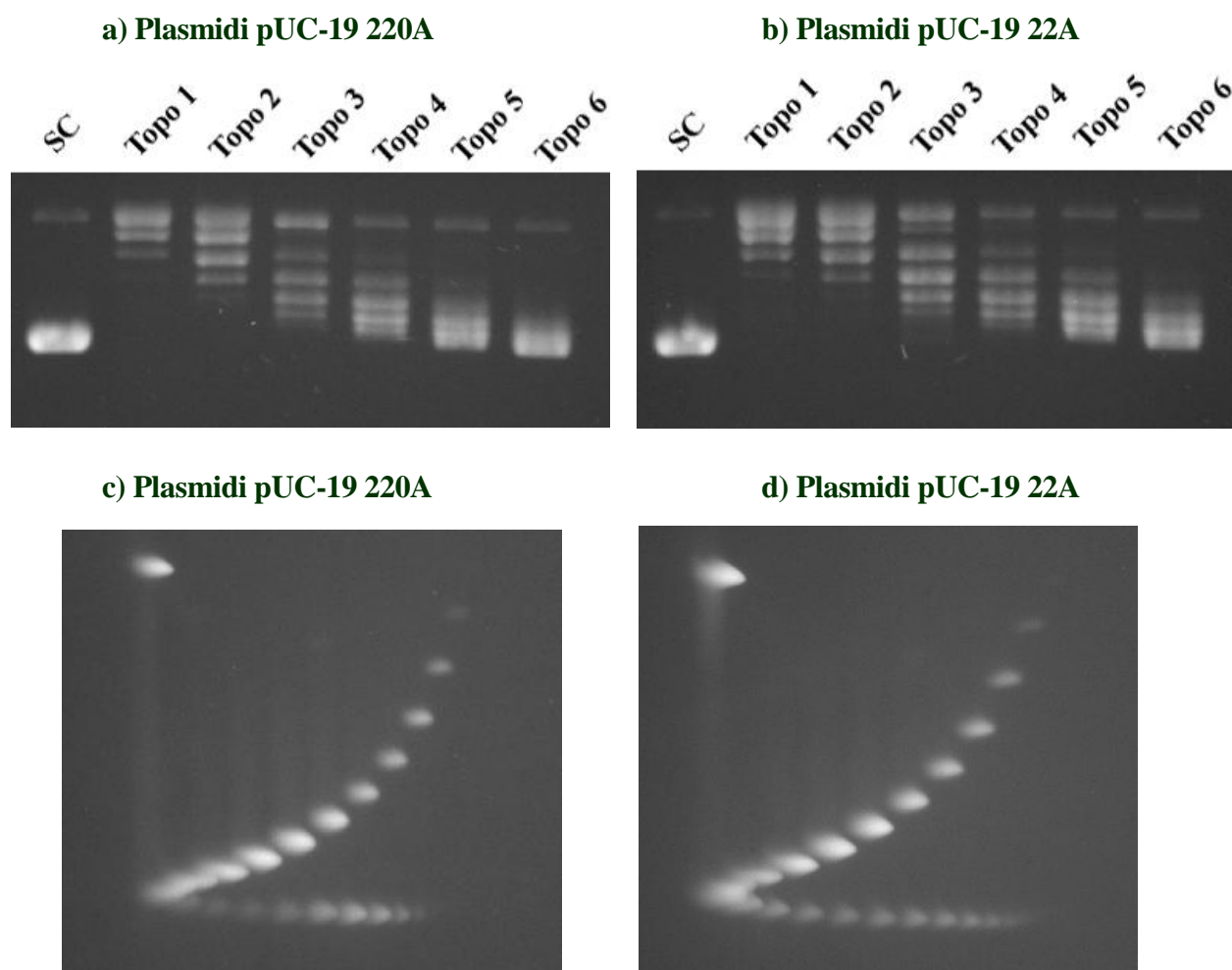


Fig. 4.114. Distribucions de topoisòmers (a, b) i electroforesis bidimensionals (b, c) per a pUC-19 220A i pUC-19 22A.

Per als plasmidis pUC-19 220A i pUC-19 22A s'ha realitzat una barreja de topoisòmers ajuntant entre 3 i 5 μ l de mostra de cada reacció (fig. 4.114 (a) i (b)), uns 3,5 μ g de DNA circular en total. Aquesta barreja s'ha carregat en un gel d'agarosa gran (20 x 20 cm) i s'ha desenvolupat l'electroforesi bidimensional de la forma habitual (v. fig. 4.114 (c) i (d)).

Tot i que el perfil de les taques s'observa perfectament en els dos casos, no s'arriba a observar cap discontinuïtat. Per tant, s'hauria d'intentar esbrinar el motiu pel qual l'experiment no resulta reproducible.

Només hem determinat dues diferències entre el nostre protocol i el seu. En primer lloc, per a l'eliminació de la topoisomerasa I i el bromur d'etidi de les mostres de DNA nosaltres fem servir una precipitació amb etanol, en canvi ells prèviament a aquesta precipitació realitzen una extracció fenol/cloroform. En segon lloc, ells fan servir una agarosa amb baix punt de fusió mentre que nosaltres usem una agarosa estàndard. Creiem que l'agarosa no és un factor determinant per a l'observació d'aquest tipus d'estructura ja que observem força ben separades les bandes. Caldria provar l'extracció fenol/cloroform, ja que el fet de no eliminar completament el bromur d'etidi de la mostra podria suposar un impediment per a l'obtenció d'aquest tipus de transició estructural segons el Dr. Seeman.

4.7.2. Microscòpia de força atòmica (AFM)

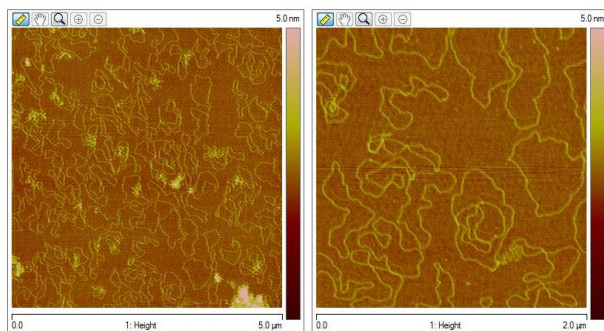
Les estructures en forma de manuelles (*dumbbell-like structures*) que formen els DNA relaxats per la formació d'estructures tipus PX-DNA poden ser visualitzades directament per microscòpia de força atòmica (AFM) (v. apartat 1.2.7). Aquesta tècnica permet l'obtenció dels perfils topològics de les mostres de DNA circular dipositades en un suport de mica (veure apartat 3.12).

En el nostre cas, hem analitzat per microscòpia de força atòmica les mostres dels plasmidis pSTAT-CEN2 (cap insert), pSTAT-CEN2-S1 (un insert (AT)₁₁) i plasmidi pSTAT-CEN2-S2 (dos inserts (AT)₁₁) (fig. 4.115). Com no hem arribat a veure les transicions estructurals del plasmidi pSTAT-CEN2-S2 en electroforesis bidimensionals, hem analitzat com a primeres proves les formes superenrotllades corresponents a les mostres de partida.

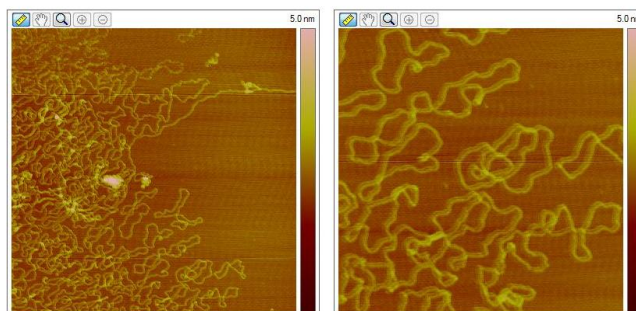
El procediment que hem seguit per a la preparació de les mostres està descrit a l'apartat 3.12 i les concentracions de les mostres van ser de 15 ng/ μ l, 13 ng/ μ l i 1,3 ng/ μ l per als plasmidis amb cap insert, un i dos, respectivament.

A la fig. 4.115 observem com hem pogut obtenir unes bones imatges d'AFM on es visualitzen clarament els DNAs circulars que hem analitzat. Podem observar l'alçada de les mostres analitzades a l'eix de la dreta de la imatge, obtenint-se una diferència entre el *background* (marró clar) i la mostra (groc) d'aproximadament 2 nm (corresponents al diàmetre de la doble hèlix). L'amplada dels DNAs circulars, mesurada a l'eix inferior, és d'entre 500 i 1000 nm depenent de si la molècula en què es mesura està més o menys compactada (aquest valor sempre s'obté una mica sobreestimat en les anàlisis d'AFM). No hem pogut veure en aquestes imatges, però, la formació de les estructures en forma de manuelles (*dumbbell-like structures*). Hem de tenir present que aquestes eren unes proves preliminars i que l'ideal seria poder analitzar les molècules que presentin una transició estructural en els gels bidimensionals, retallant les bandes corresponents del gel i recuperant-ne el DNA.

a) Plasmidi pSTAT-CEN2 (control)



b) Plasmidi pSTAT-CEN2-S1



c) Plasmidi pSTAT-CEN2-S2

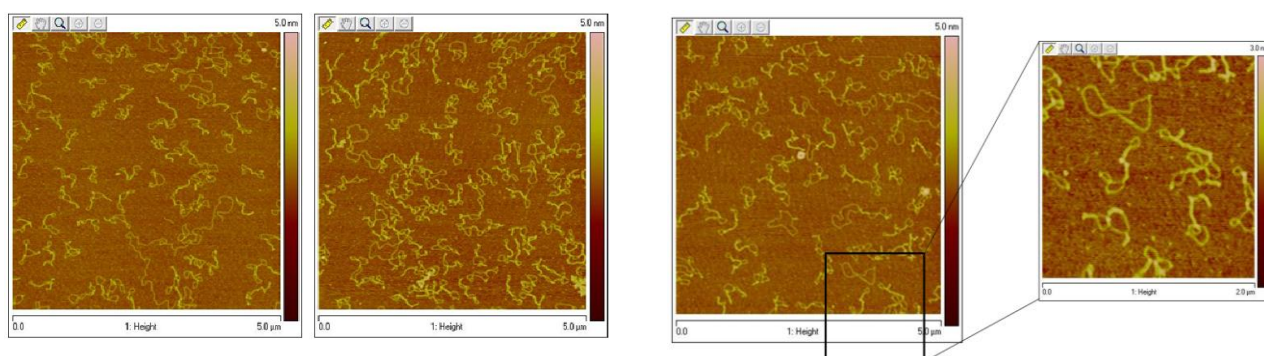


Fig. 4.115. Imatges d'AFM per als plasmidis pSTAT-CEN2 (a), pSTAT-CEN2-S1 (b) i pSTAT-CEN2-S2 (c).

5. Discussió general

HMGA vs HMGB

Les proteïnes HMGA i HMGB presenten característiques comunes com a membres de la superfamília de les HMGs: alteren l'estructura de la cromatina influenciant d'aquesta manera la transcripció, la replicació i la reparació del DNA. Són considerades com a factors arquitectònics que modulen la cromatina i regulen l'expressió de nombrosos gens de forma dinàmica. L'alta mobilitat d'aquestes proteïnes nuclears garanteix la seva disponibilitat d'unió a diferents llocs de forma ràpida i permanent [4]. Les proteïnes HMGA i HMGB també presenten afinitat d'unió a DNA no canònic com ara DNA doblegat, cruciforms o DNA superenrotllat..

Les proteïnes HMGA i HMGB es diferencien, però, en els seus motius d'unió característics, en les preferències pels seus substrats d'unió i en els canvis que provoquen en aquests. Les proteïnes HMGA, intrínsecament desordenades, presenten uns motius curts (els *AT-hooks*), que s'uneixen a seqüències riques en ATs del solc menor del DNA provocant certes distorsions, mentre que els HMG-box de les HMGBs, són dominis formats per tres hèlixs alfa que s'uneixen al DNA amb baixa o sense especificitat de seqüència i provoquen una gran distorsió d'aquest [4]. Les quantitats que es troben al nucli d'aquestes dues proteïnes són diferents i es troben expressades de forma diferent en diferents tipus de cèl·lules i/o estats d'aquesta. També, a part dels seus rols com a "xaperones del DNA", presenten diferents capacitats per a la unió a d'altres proteïnes, principalment factors de transcripció [5]. A més a més únicament la proteïna HMGB1 presenta una funció al citoplasma i al medi extracel·lular [28].

Tot i que aquestes proteïnes van ser descobertes fa 40 anys i es va coneixent poc a poc les seves activitats biològiques, encara molts dels mecanismes moleculars relacionats amb les proteïnes HMGA i HMGB segueixen sense ser completament entesos. En aquest treball s'ha volgut aportar una mica més d'informació sobre la interacció de les proteïnes HMGA i HMGB amb el DNA, els canvis que provoquen en aquest i estudiar els seus mecanismes de reconeixement. A continuació es discuteixen breument alguns d'aquests aspectes.

Flexibilitat intrínseca de les proteïnes HMGA

Les proteïnes HMGA no presenten una estructura secundària definida en solució i es consideren IDPs (*intrinsically disordered proteins*) [16]. Mitjançant estudis de dicroïsm circular es veu que aquesta proteïna quan no està unida al DNA mostra un 73% de *random-coil* (i un contingut del 27% restant en girs β) [147]. A l'unir-se al DNA, però, aquestes proteïnes s'estructuren parcialment, augmentant el seu contingut en girs β un 20-40% [147]. Els primers estudis d'RMN d'un fragment de l'HMGA1a en complex amb DNA (lloc d'unió del promotor de l'interferó β) mostraven que eren justament els motius *AT-hook* els que s'ordenaven. En aquesta estructura es va veure que el fragment PRGRP dels *AT-hooks* II i III mostra una configuració plana en forma de mitja lluna que s'insereix en el solc menor del DNA, posicionant-se les cadenes laterals de les arginines seguint el gir del solc menor [19].

Es creu que justament l'alt grau de flexibilitat degut a la falta d'estructura secundària juga un rol molt important en possibilitar la participació de les proteïnes HMGA en una gran varietat de processos biològics. A més a més cada cop es considera més la importància biològica de les transicions d'estructures desordenades a ordenades que presenten algunes proteïnes. Les transicions reversibles entre configuracions ordenades i desordenades de les proteïnes HMGA probablement proporcionen el suport necessari per a la regulació de la formació de complexos d'HMGA funcionals dins de les cèl·lules, controlant així l'activitat biològica d'aquestes proteïnes *in vivo* [14].

Gràcies a l'estructura resolta recentment en el nostre laboratori [20], s'han observat els ponts d'hidrogen poc específics entre els grups NH de les cadenes principals de la Gly37 i l'Arg38 amb els oxígens de tres timines (v. apartat 1.1.1). S'ha observat la mateixa preferència d'unió a una seqüència AATT que presentaven els *AT-hooks* II i III resolts per RMN [19]. En aquest cas, semblaria que la menor amplada del solc menor que presenten les regions AT seria la que contribuiria al reconeixement dels motius *AT-hook*. De fet és coneguda la capacitat de les arginines per a unir-se al solc menor de regions riques en ATs, que presenten una menor amplada [148]. Tot això ens proporciona una idea d'altres mecanismes de reconeixement en el solc menor addicionals als patrons de formació de ponts d'hidrogen que presenten especificitat limitada [149] (v. apartat 1.2.4).

Finalment cal destacar que la unió dels *AT-hooks* al DNA en regions riques en ATs, a diferència del que alguns estudis sostenien, modifica lleugerament la forma B del DNA (com s'ha vist a l'estructura de raigs X) eixamplant el solc menor en la zona d'unió, provocant un cert doblegament del DNA i desenrotllament de la doble hèlix [20].

En un primer moment, en aquest treball, ens vàrem proposar l'ambiciós objectiu de cristal·litzar les proteïnes HMGA1a i HMGA1b senceres i fragments d'elles, contenint els tres *AT-hooks*. Degut a aquesta flexibilitat intrínseca, per la pròpia falta d'una estructura definida en les zones entre *AT-hooks* d'aquestes proteïnes, obtenir cristalls ordenats dels complexos HMGA-DNA va suposar tot un repte. Tot i no poder millorar els diagrames de difracció dels cristalls HMGA1a-DNA que en estudis propis anteriors [128] mostraven la localització ordenada únicament de DNA dins dels cristalls, es va poder confirmar la presència proteica. Aquest fet indica la unió proteïna-DNA i confirma la dificultat d'obtenir una estructura ordenada de la proteïna dins de l'empaquetament del cristall.

Segueix, però, sent de gran interès l'obtenció d'una estructura que presenti tots tres *AT-hook*, tot i que potser la zona entre *AT-hooks* seria difícilment caracteritzable, i veure d'aquesta manera les distorsions que poden provocar en el DNA la interacció de tots tres *AT-hook* alhora i els seus mecanismes de reconeixement. A més a més seria també de gran importància poder caracteritzar la interacció d'aquests *AT-hooks* ja que moltes altres proteïnes no canòniques contenen també aquest motiu d'unió al DNA [150]. Per exemple, la proteïna MeCP2 (*methyl-CpG-binding protein 2*), relacionada amb la síndrome de Rett, a més del MBD (*methyl binding domain*), conté tres dominis *AT-hook*. Recentment s'ha publicat que justament un d'aquests *AT-hooks* determina el curs clínic d'aquesta síndrome i desordres associats [151].

Mecanismes de reconeixement de les proteïnes HMGB

L'estudi de les interaccions a partir de les estructures a nivell atòmic dels complexos entre les proteïnes i el DNA ens permet un major coneixement dels mecanismes de reconeixement que presenten aquestes proteïnes. En el cas de l'estructura resolta en aquest treball, corresponent al domini Box AS de la proteïna HMGB1 en complex amb DNA lineal, hem pogut observar la formació d'un parell de ponts d'hidrogen directes a les bases del DNA, juntament amb d'altres contactes no específics. Aquest fet és congruent amb l'esperada especificitat de seqüència limitada que presenten aquest tipus d'HMG-box. Però el reconeixement i la distorsió que les proteïnes d'unió al solc menor del DNA poden provocar pot dependre d'altres factors com canvis en l'amplada del solc menor (com s'observa amb les HMGAs) o bé canvis locals en la flexibilitat del DNA.

De fet, existeixen dos mecanismes de discriminació dels llocs d'unió al DNA que poden presentar les proteïnes: directe o indirecte [152]. En l'anomenada lectura directa (*direct readout*) residus específics formen ponts d'hidrogen amb bases específiques (A, T, C o G). Normalment aquest tipus d'unió té lloc en el solc major on el patró de donadors i acceptors de ponts d'hidrogen és

diferent per a cada parell de bases [149] (v. apartat 1.2.4). Un exemple podrien ser els factors de transcripció que reconeixen molt específicament el solc major del DNA.

D'altra banda, es troben els factors arquitectònics d'unió al solc menor del DNA, del que formen part les proteïnes d'estudi d'aquest treball, entre d'altres. El patró de reconeixement en la formació de ponts d'hidrogen del solc menor és més limitat [149]. En el solc menor és més difícil diferenciar la A de la T i la C de la G, i només és diferent el grup donador del mig en els parells de bases A:T i C:G (sent ADA el patró per A:T i T:A i AHA per C:G i G:C, segons nomenclatura de l'apartat 1.2.4). Així, en el mecanisme de reconeixement de lectura indirecte, les variacions de les propietats mecàniques (com la flexibilitat o l'amplada del solc, per exemple) relacionades amb la seqüència juguen un rol molt important. Per tant es creu que la manera d'aquestes proteïnes de "llegir" seqüències nucleotídiques depèn de seva facilitat en doblegar, enrotllar o deformar llocs d'unió potencials. Molt sovint, però, es troben tots dos tipus de reconeixement (directe i indirecte) en la unió de les proteïnes al DNA.

Les proteïnes HMG-box, com d'altres factors arquitectònics d'unió al solc menor del DNA, introdueixen doblegaments notables en el DNA deguts a la inserció parcial o la intercalació de residus hidrofòbics que provoquen un gran valor positiu de l'angle de *roll*. Majoritàriament la intercalació d'aquests residus té lloc als passos YR, RR o YY (sent R purina, G o A, i Y pirimidina, C o T) els quals, en estat lliure, presenten qualitats inherents per a ser sotmesos a deformacions i doblegaments (revisat a [153]). Però és el pas YR, pirimidina-purina, el més fàcilment deformable i el que ha presentat major flexibilitat a l'estudiar els passos més variables en complexos proteïna-DNA [154].

En el nostre cas, hem observat la intercalació de les Phe37 de dos dominis Box A justament en el pas CG de l'oligonucleòtid ATATCGATAT corresponent a un pas pirimidina-purina (YR). Una intercalació similar es pot observar en l'estructura de la proteïna TBP (*TATA-Box binding protein*). Aquesta proteïna, que és un factor de transcripció que intervé en la formació del complex PIC pas previ a la iniciació de la transcripció, presenta dos fenilalanines intercalades en el pas TA (pirimidina-purina) (en l'estructura de la TBP d'*A. thaliana* [155]). La interacció d'aquesta proteïna mitjançant interaccions hidrofòbiques i electrostàtiques amb el DNA provoca en aquest un doblegament continu i regular d'uns 80°.

Dins de la limitada especificitat de seqüència dels HMG-boxes de l'HMGB1, sembla que el domini Box A pugui presentar una certa preferència per a la intercalació del seu únic residu d'intercalació, la Phe37, entre els passos més fàcilment deformables pirimidina-purina. Tot això podria confirmar la preferència d'intercalació entre bases pirimidina-purina de les HMG-box que presenten només un únic lloc d'intercalació, com es va veure per al mutant HMGD-M13A [39]. L'estructura de raigs X de la forma *wild type* de l'HMGD presenta la intercalació dels residus del

segon lloc d'intercalació (Val-Thr) entre un pas purina-purina, però quan el primer lloc d'intercalació s'anul·la (mutant HMGD-M13A), deixant així a la proteïna amb un únic lloc d'intercalació funcional, s'observa el canvi de la intercalació dels residus del segon lloc d'intercalació a un pas pirimidina-purina [45].

La diferència entre especificitat de reconeixement de seqüència de les proteïnes HMG-box depèn dels ponts d'hidrogen específics que puguin formar amb el solc estret del DNA. Així s'ha vist que les proteïnes amb especificitat de seqüència (v. apartat 1.1.2), a diferència de les de limitada especificitat, formen ponts d'hidrogen específics amb les bases del DNA (com per exemple l'Asn13 de l'SRY que presenta ponts d'hidrogen amb tres àtoms de les bases [41]).

Gràcies a la resolució del domini Box AS de l'HMGB1 en complex amb DNA d'aquest treball s'ha pogut conèixer per primer cop els mecanismes d'unió i distorsió d'un DNA lineal per part d'aquest domini. Fins al moment només s'havia vist la unió d'un domini Box A a un DNA modificat amb cisplatí. A diferència del Box BS i de les proteïnes NHP6A i HMGD (proteïnes més similars al domini Box B), el domini Box AS presenta la intercalació d'un únic residu (Phe) entre un pas de la doble hèlix (ja que l'alanina corresponent al residu del primer lloc d'intercalació no arriba a intercalar-se). Aquest fet provoca que el doblegament del dúplex sigui en forma de colze, és a dir en un únic punt i no gradual. En la nostra estructura hem observat per primera vegada la intercalació de dos residus de dos dominis HMG-box en un mateix pas. Això provoca la major distorsió dels angles de *roll* i *twist* en un pas de la que tenim constància per a un domini HMG-box i crea un doblegament global d'uns 85°, comparable a l'observat per un únic domini d'HMG-box amb dos punts d'intercalació (per exemple el Box B de l'HMGB1 [44]).

L'estructura presentada en aquest treball ha aportat nova informació sobre els mecanismes de reconeixement i les distorsions que provoca el domini Box A de la proteïna HMGB1 en un DNA lineal. Però justament aquest model d'interacció planteja noves incògnites. Per exemple no se sap si la interacció de forma conjunta de tots dos dominis per a doblegar el dúplex de DNA es produeix de forma simultània o si la unió d'un domini Box A pot afavorir la interacció del segon (presentant així certa cooperativitat). També caldria estudiar si els dominis HMG-box d'altres proteïnes poden mostrar aquesta interacció conjunta intercalant-se un residu de cada domini en un mateix pas o si és una característica exclusiva del domini Box A de l'HMGB1. Finalment resta pendent aconseguir obtenir un model estructural de la proteïna HMGB1 sencera (amb tots dos dominis, Box A i B) en complex amb DNA per a tenir un coneixement global de la seva implicació en la distorsió del DNA i un enteniment de les diferències entre els dominis Box A i Box B en el reconeixement i en els canvis que provoquen en el DNA.

6. Conclusions

1. HMG expression and purification:

1.1. HMGA1b(1-79) mutant was constructed using site directed mutagenesis. This mutant as well as HMGA1a protein, HMGB1 Box AS, Box BS and Box ABS fragments were expressed in *E. coli* strain Rosetta(DE3)pLysS and purified.

1.2. NHP6A was expressed in *E. coli* strain BL21(DE3) and purified.

2. Crystallization of HMGA-DNA complexes:

2.1. The quality of the HMGA1a-CGCATTAATGCG crystals was not good enough to determine the presence of protein in the structure. We did not obtain good crystals of DNA and HMGA1b(1-79) complexes either.

2.2. Crystals of different morphologies of two AT-hook peptides and the CGCATTAATGCG oligonucleotide were obtained. However, their diffraction patterns showed disorder.

2.3. A crystal of the PRGRP AT-hook peptide and the AATAATTATT oligonucleotide diffracted to 3.1 Å resolution. The space group was C2 with $a=149$ Å, $b=252$ Å, $c=51$ Å; $\alpha=90^\circ$, $\beta=104^\circ$, $\gamma=90^\circ$. The asymmetric unit comprises 8 parallel duplexes.

3. Crystallization of HMGB-DNA complexes:

3.1. NHP6A with DNA only crystallized at high salt and high pH conditions which seems to indicate the presence of only DNA in the crystal. In the case of the complexes of Box ABS and DNA, we could not find any appropriate crystallization condition.

3.2. Crystals grown from a precipitate as well as needles were obtained for the HMGB1 Box BS fragment and AATAAATTTATT. However, the quality of the diffraction was not as good as needed to determine their structure. By contrast, a rod shaped crystal, which was grown from a drop containing needles, diffracted to 3.3 Å resolution. This crystal belonged to P222 space group with the cell parameters $a=39$ Å, $b=68$ Å, $c=91$ Å; $\alpha=90^\circ$, $\beta=90^\circ$, $\gamma=90^\circ$.

3.3. We have obtained diffraction-quality crystals for the HMGB1 Box AS fragment in complex with both AATAAATTTATT and ATATCGATAT. The first oligonucleotide produced thin plates whereas the latter needles and rod shaped crystals.

3.3.1. Box AS and AATAAATTTATT plates diffracted to 2.8 Å resolution and belonged to P1 space group with $a=39$ Å, $b=39$ Å, $c=91$ Å; $\alpha=90^\circ$, $\beta=90^\circ$, $\gamma=90^\circ$. The preliminary analysis of the structure indicated the placement of two 60° crossed duplexes into the asymmetric unit and the possible location of the protein in the cavities.

3.3.2. Three unit cells were calculated for the crystals of Box AS and ATATCGATAT (C, D and E cells in table 4.8). The C cell parameters are $a=43$ Å, $b=84$ Å, $c=94$ Å; $\alpha=90^\circ$, $\beta=90^\circ$, $\gamma=90^\circ$ in P222 and could contain 2 complexes in the asymmetric unit according to the Matthews coefficient probability. The D cell ($a=94$ Å, $b=94$ Å, $c=100$ Å; $\alpha=90^\circ$, $\beta=90^\circ$, $\gamma=90^\circ$ in C222) could have 3 complexes and the E cell ($a=105$ Å, $b=163$ Å, $c=89$ Å; $\alpha=90^\circ$, $\beta=90^\circ$, $\gamma=90^\circ$ in C222) 5 complexes.

4. Atomic structure of the Box AS-DNA:

4.1. We have focused on the determination, using molecular replacement, of the structure corresponding to the C cell, which showed the highest resolution among all the Box AS crystals. The diffraction data were finally processed in P2₁2₁2₁ to 2 Å resolution. The final model showed two Box AS domains and two ATATCGATAT duplexes in the asymmetric unit.

4.2. Two Box AS domains interact symmetrically in a tail-to-tail fashion with one of the duplexes intercalating the Phe37 of each domain at the central CG base pair step. This interaction of both Box AS kinks the DNA by 85°, underwinds the double helix and induces a remarkable distortion of the DNA duplex that shows some characteristic features of the A-form of the DNA. The other duplex remains straight almost without interacting with any residue of the protein.

4.3. In the kinked duplex, regions of helix I and helix II of Box AS bind to one DNA strand, the central region of helix I together with the N-terminal stretch primarily contacts the other DNA strand. Helix III does not interact with the DNA at all.

- 4.4. The side chain of Ser 41 and the main chain of Phe37 make direct hydrogen bonds with DNA bases, while Ser13 makes a water mediated hydrogen bond. The structure also shows multiple hydrogen bonds between the protein and the sugar-phosphate backbone, Lys7 and Arg9 participating in two contacts that were not observed previously.
- 4.5. The DNA distortion in this structure is slightly greater than the observed for the pre-bent cisplatin DNA in complex with Box A. As for the comparison of the Box A domains of our structure with the latter, the main differences are found in the location of loop 1 and in the fact that the N-terminal of helix I is straighter in the cisplatin-modified structure. Also the orientation of the phenyl ring is different in both structures.
- 4.6. We have found that the interaction of two Box AS domains together creates the larger distortion of the roll and the twist angles in a base pair step observed until now for HMG-box complexes.

5. DNA topology studies:

- 5.1. We have observed the introduction of negative supercoiling by Box AS, Box ABS and NHP6A in circular DNAs in the presence of topoisomerase I. This negative supercoiling introduction is related with the ability of these proteins to distort DNA, for example by unwinding the double helix.
- 5.2. Under the applied conditions, we have not seen a structural transition of the plasmid with two (AT)₁₁ segments to a PX-DNA structure.

7. Bibliografia

7.1. Referències bibliogràfiques

- [1] Goodwin, G.H., Sanders, C. & Johns, E.W. (1973). A new group of chromatin-associated proteins with a high content of acidic and basic amino acids. *Eur. J. Biochem.* **38**, 14-19.
- [2] Bianchi, M.E. & Agresti, A. (2005). HMG proteins: dynamic players in gene regulation and differentiation. *Curr. Opin. Genet. Dev.* **15**, 496-506.
- [3] Bustin, M. & Reeves, R. (1996). High-mobility-group chromosomal proteins: architectural components that facilitate chromatin function. *Prog. Nucleic Acid Res. Mol. Biol.* **54**, 35-100.
- [4] Catez, F. & Hock, R. (2010). Binding and interplay of HMG proteins on chromatin: lessons from live cell imaging. *Biochim. Biophys. Acta.* **1799**, 15-27.
- [5] Reeves, R. (2010). Nuclear functions of the HMG proteins. *Biochim. Biophys. Acta.* **1799**, 3-14.
- [6] Bustin, M. (2001). Revised nomenclature for high mobility group (HMG) chromosomal proteins. *Trends. Biochem. Sci.* **26**, 152-153.
- [7] Hock, R., Furusawa, T., Ueda, T. & Bustin, M. (2007). HMG chromosomal proteins in development and disease. *Trends. Cell. Biol.* **17**, 72-79.
- [8] Kato, H. *et al.* (2011). Architecture of the high mobility group nucleosomal protein 2-nucleosome complex as revealed by methyl-based NMR. *Proc. Natl. Acad. Sci. USA.* **108**, 12283-12288.
- [9] Catez, F., Brown, D.T., Misteli, T. & Bustin, M. (2002). Competition between histone H1 and HMGN proteins for chromatin binding sites. *EMBO Rep.* **3**, 760-766.
- [10] Catez, F. *et al.* (2004). Network of dynamic interactions between histone H1 and high-mobility-group proteins in chromatin. *Mol. Cell. Biol.* **24**, 4321-4328.
- [11] Fusco, A. & Fedele, M. (2007). Roles of HMGA proteins in cancer. *Nat. Rev. Cancer*, **7**, 899-910.
- [12] Reeves, R. & Beckerbauer, L. (2001). HMGI/Y proteins: flexible regulators of transcription and chromatin structure. *Biochim. Biophys. Acta.* **1519**, 13-29.
- [13] Reeves, R. (2000). Structure and Function of the HMGI(Y) Family of Architectural Transcription Factors. *Environ. Health Perspect.* **108**, 803-809.

- [14] Reeves, R. (2001). Molecular biology of HMGA proteins: hubs of nuclear function. *Gene*. **277**, 63-81.
- [15] Uversky, V.N. (2011). Intrinsically disordered proteins from A to Z. *Int. J. Biochem. Cell. Biol.* **43**, 1090-1103.
- [16] Dyson, H.J. & Wright, P.E. (2005). Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell. Biol.* **6**, 197-208.
- [17] Cleynen, I. & Van de Ven, W.J. (2008). The HMGA proteins: a myriad of functions. *Int. J. Oncol.*, **32**, 289-305.
- [18] Geierstanger, B.H., Volkman, B.F., Kremer, W. & Wemmer, D.E. (1994). Short peptide fragments derived from HMG-I/Y proteins bind specifically to the minor groove of DNA. *Biochemistry*. **33**, 5347-5355.
- [19] Huth, J.R. *et al.* (1997). The solution structure of an HMG-I(Y)-DNA complex defines a new architectural minor groove binding motif. *Nat. Struct. Biol.* **4**, 657-665.
- [20] Fonfría-Subirós, E. *et al.* (2012). Crystal structure of a complex of DNA with one AT-hook of HMGA1. *PLoS One*. **7**.
- [21] Tallini, G. & Dal Cin, P. (1999). HMGI(Y) and HMGI-C dysregulation: a common occurrence in human tumors. *Adv. Anat. Pathol.* **6**, 237-246.
- [22] Giancotti, V. *et al.* (1991). Comparison of multiple forms of the high mobility group I proteins in rodent and human cells. *Eur. J. Biochem.* **198**, 211-216.
- [23] Giancotti, V. *et al.* (1989). Analysis of the HMGI nuclear proteins in mouse neoplastic cells induced by different procedures. *Exp. Cell. Res.* **184**, 538-545.
- [24] Thomas, J.O. & Travers, A.A. (2001). HMG1 and 2, and related 'architectural' DNA-binding proteins. *Trends. Biochem. Sci.* **26**, 167-174.
- [25] Stros, M. (2010). HMGB proteins: Interactions with DNA and chromatin. *Biochim. Biophys. Acta.* **1799**, 101-113.
- [26] Stros, M., Launholt, D. & Grasser, K.D. (2007). The HMG-box: a versatile protein domain occurring in a wide variety of DNA-binding proteins. *Cell Mol. Life Sci.* **64**, 2590-2606.
- [27] Ivics, Z. *et al.* (2004). The Sleeping Beauty Transposable Element: Evolution, Regulation and Genetic Applications. *Curr. Issues Mol. Biol.* **6**, 43-56.
- [28] Tang, D., Kang, R., Zeh, H.J. 3rd & Lotze, M.T. (2010). High-mobility group box 1 and cancer. *Biochim. Biophys. Acta.* **1799**, 131-140.

- [29] Klune, J.R., Dhupar, R., Cardinal, J., Billiar, T.R. & Tsung, A. (2008). HMGB1: endogenous danger signaling. *Mol. Med.* **14**, 476-484.
- [30] Sims, G.P., Rowe, D.C., Rietdijk, S.T., Herbst, R. & Coyle, A.J. (2010). HMGB1 and RAGE in inflammation and cancer. *Annu. Rev. Immunol.* **28**, 367-388.
- [31] Malarkey, C.S. & Churchill, M.E. (2012). The high mobility group box: the ultimate utility player of a cell. *Trends. Biochem. Sci.* **37**, 553-562.
- [32] Hardman, C.H. *et al.* (1995). Structure of the A-domain of HMG1 and its interaction with DNA as studied by heteronuclear three- and four-dimensional NMR spectroscopy. *Biochemistry.* **34**, 16596-16607.
- [33] Weir, H.M. *et al.* (1993). Structure of the HMG box motif in the B-domain of HMG1. *EMBO J.* **12**, 1311-1319.
- [34] Teo, S.H., Grasser, K.D. & Thomas, J.O. (1995). Differences in the DNA-binding properties of the HMG-box domains of HMG1 and the sex-determining factor SRY. *Eur. J. Biochem.* **230**, 943-950.
- [35] Travers, A.A. (2000). Recognition of distorted DNA structures by HMG domains. *Curr. Opin. Struct. Biol.* **10**, 102-109.
- [36] Muller, S., Bianchi, M.E. & Knapp, S. (2001). Thermodynamics of HMGB1 interaction with duplex DNA. *Biochemistry*, **40**, 10254-10261.
- [37] Dragan, A.I. *et al.* (2004). DNA binding and bending by HMG boxes: energetic determinants of specificity. *J. Mol. Biol.* **343**, 371-393.
- [38] Klass, J. *et al.* (2003). The role of intercalating residues in chromosomal high-mobility-group protein DNA binding, bending and specificity. *Nucleic Acids Res.* **31**, 2852-64.
- [39] Churchill, M.E., Klass, J. & Zoetewey, D.L. (2010). Structural analysis of HMGB1-DNA complexes reveals influence of intercalation on sequence selectivity and DNA bending. *J. Mol. Biol.* **403**, 88-102.
- [40] Murphy, F.V. 4th & Churchill, M.E. (2000). Nonsequence-specific DNA recognition: a structural perspective. *Structure.* **8**, 83-89.
- [41] Werner, M.H., Huth, J.R., Gronenborn, A.M. & Clore, G.M. (1995). Molecular basis of human 46X,Y sex reversal revealed from the three-dimensional solution structure of the human SRY-DNA complex. *Cell.* **81**, 705-714.
- [42] Love, J.J. *et al.* (1995). Structural basis for DNA bending by the architectural transcription factor LEF-1. *Nature.* **376**, 791-795.

- [43] Ohndorf, U.M., Rould, M.A., He, Q., Pabo, C.O. & Lippard, S.J. (1999). Basis for recognition of cisplatin-modified DNA by high-mobility-group proteins. *Nature*. **399**, 708-712.
- [44] Stott, K., Tang, G.S.F., Lee, K.-B. & Thomas, J.O. (2006). Structure of a Complex of Tandem HMG Boxes and DNA. *J. Mol. Biol.* **360**, 90-104.
- [45] Murphy, F.V. 4th, Sweet, R.M. & Churchill, M.E.A. (1999). The structure of a chromosomal high-mobility group protein-DNA complex reveals sequence-neutral mechanisms important for non-sequencespecific DNA recognition. *EMBO J*, **18**, 6610-6618.
- [46] Allain, F.H. *et al.* (1999). Solution structure of the HMG protein NHP6A and its interaction with DNA reveals the structural determinants for nonsequence-specific binding. *EMBO J*. **18**, 2563-2579.
- [47] Ngo, H.B., Kaiser, J.T. & Chan, D.C. (2011). The mitochondrial transcription and packaging factor Tfam imposes a U-turn on mitochondrial DNA. *Nat. Struct. Mol. Biol.* **18**, 1290-1296.
- [48] Rubio-Cosials, A. *et al.* (2011). Human mitochondrial transcription factor A induces a U-turn structure in the light strand promoter. *Nat. Struct. Mol. Biol.* **18**, 1281-1289.
- [49] Choi, Y.R. *et al.* (2003). Overexpression of high mobility group box 1 in gastrointestinal stromal tumors with KIT mutation. *Cancer Res.* **63**, 2188–2193.
- [50] Volp, K. *et al.* (2006). Increased expression of high mobility group box 1 (HMGB1) is associated with an elevated level of the antiapoptotic c-IAP2 protein in human colon carcinomas. *Gut*. **55**, 234–242.
- [51] Hao, Q., Du, X.Q., Fu, X. & Tian, J. (2008). Expression and clinical significance of HMGB1 and RAGE in cervical squamous cell carcinoma. *Zhonghua Zhongliu Zazhi*. **30**, 292–295.
- [52] Watson, J.D & Crick, F.H. (1953). Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature*. **171**, 737-738.
- [53] Watson, J.D. *et al.* (2008). *Molecular biology of the gene*. 6th edition. (San Francisco Cold Spring Harbor, Pearson/Benjamin Cummings; Cold Spring Harbor Laboratory Press, New York).
- [54] Saenger, W. (1984). Principles of nucleic acid structure. *Springer-Verlag, New York*.
- [55] Berg, J.M., Tymoczko, J.L. & Stryer, L. (2002). Biochemistry. 5th edition. *W. H. Freeman, New York*.

- [56] Franklin, R.E. & Gosling, R.G. (1953). Molecular configuration in sodium thymonucleate. *Nature*. **171**, 740-741.
- [57] Bates, A.D. & Maxwell, A. (1993). DNA topology. *IRL Press at Oxford University Press, Oxford*.
- [58] Wang, A.H. *et al.* (1979). Molecular structure of a left-handed double helical DNA fragment at atomic resolution. *Nature*. **282**, 680-686.
- [59] Wing, R. *et al.* (1980). Crystal structure analysis of a complete turn of B-DNA. *Nature*. **287**, 755-758.
- [60] Dickerson, R.E. *et al.* (EMBO Workshop). (1989). Definitions and nomenclature of nucleic acid structure parameters. *EMBO J.* **8**, 1-4.
- [61] Neidle, S. (2007). Principles of Nucleic Acid Structure. *1st Edition, Academic Press, London*.
- [62] Lu, X.J. & Olson, W.K. (2003). 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res.* **31**, 5108-5121.
- [63] Olson, W.K. *et al.* (2001). A standard reference frame for the description of nucleic acid base-pair geometry. *J. Mol. Biol.* **313**, 229-237.
- [64] Murchie, A.I. & Lilley, D.M. (1992). Supercoiled DNA and cruciform structures. *Methods Enzymol.* **211**, 158-180.
- [65] Matek, C., Ouldrige, T.E., Levy, A., Doye, J.P. & Louis, A.A. (2012). DNA cruciform arms nucleate through a correlated but asynchronous cooperative mechanism. *J. Phys. Chem. B.* **116**, 11616-11625.
- [66] Lilley, D.M. (1980). The inverted repeat as a recognizable structural feature in supercoiled DNA molecules. *Proc. Natl. Acad. Sci. USA.* **77**, 6468-6472.
- [67] Rich, A., Nordheim, A. & Wang, A.H. (1984). The chemistry and biology of left-handed Z-DNA. *Annu. Rev. Biochem.* **53**, 791-846.
- [68] Wang, X., Zhang, X., Mao, C. & Seeman, N.C. (2010). Double-stranded DNA homology produces a physical signature. *Proc. Natl. Acad. Sci. USA.* **107**, 12547-12552.
- [69] Shen, Z., Yan, H., Wang, T. & Seeman, N.C. (2004). Paranemic crossover DNA: a generalized Holliday structure with applications in nanotechnology. *J. Am. Chem. Soc.* **126**, 1666-1674.

- [70] Campos, L. *et al.* (2006). Overview of the structure of all-AT oligonucleotides: organization in helices and packing interactions. *Biophys. J.* **91**, 892-903.
- [71] Field, D. & Wills, C. (1996). Long, polymorphic microsatellites in simple organisms. *Proc. Biol. Sci.* **263**, 209-215.
- [72] Tóth, G., Gáspári, Z. & Jurka, J. (2000). Microsatellites in different eukaryotic genomes: survey and analysis. *Genome Res.* **10**, 967-981.
- [73] Katti, M.V., Ranjekar, P.K. & Gupta, V.S. (2001). Differential distribution of simple sequence repeats in eukaryotic genome sequences. *Mol. Biol. Evol.* **18**, 1161-1167.
- [74] Eichinger, L. *et al.* (2005). The genome of social amoeba *Dictyostelium discoideum*. *Nature.* **435**, 43-57.
- [75] Leslie, A.G., Arnott, S., Chandrasekaran, R. & Ratliff, R.L. (1980). Polymorphism of DNA Double Helices. *J. Mol. Biol.* **143**, 49-72.
- [76] Zuo, E.T. *et al.* (1990). Effect of base-pair sequence on the conformations and thermally induced transitions in oligodeoxyribonucleotides containing only AT base pairs. *Biochemistry.* **29**, 4446-4456.
- [77] Guzikevich-Guerstein, G. & Shakked, Z. (1996). A novel form of the DNA double helix imposed on the TATA-box by the TATA-binding protein. *Nature Struct. Biol.* **3**, 32-37.
- [78] Mahendrasingam, A. *et al.* (1986). Time-resolved X-ray diffraction studies of the B \rightleftharpoons D structural transition in the DNA double helix. *Science.* **233**, 195-197.
- [79] Abrescia, N.G. *et al.* (2004). X-ray and NMR studies of the DNA oligomer d(ATATAT): Hoogsteen base pairing in duplex DNA. *Biochemistry.* **43**, 4092-4100.
- [80] Nikolova, E.N. *et al.* (2011). Transient Hoogsteen base pairs in canonical duplex DNA. *Nature.* **470**, 498-502.
- [81] Reeves, R. & Nissén, M.S. (1999). Purification and assays for high mobility group HMG-I(Y) protein function. *Methods Enzymol.* **304**, 155-188.
- [82] Roemer, S.C., Adelman, J., Churchill, M.E. & Edwards, D.P. (2008). Mechanism of high-mobility group protein B enhancement of progesterone receptor sequence-specific DNA binding. *Nucleic Acids Res.* **36**, 3655-3666.
- [83] Paull, T.T. & Johnson, R.C. (1995). DNA looping by *Saccharomyces cerevisiae* High Mobility Group NHP6A/B. *J. Biol. Chem.* **270**, 8744-8754.
- [84] Yen, Y.-M., Wong, B. & Johnson, R.C. (1998). Determinants of DNA Binding and Bending by the *Saccharomyces cerevisiae* High Mobility Group Protein NHP6A That Are Important for Its Biological Activities. *J. Biol. Chem.* **273**, 4424-4435.

- [85] Jones, D.N. *et al.* (1994). The solution structure and dynamics of the DNA-binding domain of HMG-D from *Drosophila melanogaster*. *Structure*. **2**, 609-627.
- [86] Wingfield, P. (2001). Protein Precipitation Using Ammonium Sulfate. *Curr. Protoc. Protein Sci.* **13**, A.3F.1–A.3F.8.
- [87] Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of head of bacteriophage T4. *Nature*. **227**, 680-685.
- [88] Panyim, S. & Chalkley, R. (1969). High resolution acrylamide gel electrophoresis of histones. *Arch. Biochem. Biophys.* **230**, 337-346.
- [89] Stephano, J.L., Gould, M. & Rojasgalicia, L. (1986) Advantages of picrate fixation for staining polypeptides in polyacrylamide gels. *Anal. Biochem.* **152**, 308-313.
- [90] Rabilloud, T., Vuillard, L., Gilly, C. & Lawrence, J.J. (1994). Silver-staining of proteins in polyacrylamide gels: a general overview. *Cell. Mol. Biol.* **40**, 57-75.
- [91] Cantor, C.R., Warshaw, M.M. & Shapiro, H. (1970). Oligonucleotide interactions. III. Circular dichroism studies of the conformation of deoxyoligonucleotides. *Biopolymers*. **9**, 1059-1077.
- [92] Chayen, N.E. (2004). Turning protein crystallization from an art into a science. *Curr. Opin. Struct Biol.* **14**, 577–583.
- [93] Jancarik, J. & Kim, S.H. (1991). Sparse matrix sampling: a screening method for crystallization of proteins. *J. Appl. Cryst.* **24**, 409-411.
- [94] Scott, W.G. *et al.* (1995). Rapid crystallization of chemically synthesized hammerhead RNAs using a double screening procedure. *J. Mol. Biol.* **250**, 327-332.
- [95] Lesley, S.A. & Wilson, I.A. (2005). Protein production and crystallization at the joint center for structural genomics. *J. Struct. Funct. Genomics*. **6**, 71-79.
- [96] International Union of Crystallography. (1969). *International Tables for X-Ray Crystallography*. Volume 1, 3rd Edition (Kynoch Press, Birmingham, England).
- [97] McPherson, A. (2009). *Introduction to macromolecular crystallography*. (2nd Edition. Wiley-Liss. New Jersey).
- [98] Wlodawer, A., Minor, W., Dauter, Z. & Jaskolski, M. (2013). Protein crystallography for aspiring crystallographers or how to avoid pitfalls and traps in macromolecular structure determination. *FEBS J.* **280**, 5705-5736.
- [99] Otwinowski, Z. & Minor, W. (1997). Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods Enzymol.* **276**, 307-326.

- [100] Leslie, A.G.W. & Powell, H.R. (2007). Processing Diffraction Data with Mosflm. *Evolv. Methods Macromol. Crystallogr.* **245**, 41-51.
- [101] Evans, P.R. (2005). Scaling and assessment of data quality. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 72-82.
- [102] French, G.S. & Wilson, K.S. (1978). On the treatment of negative intensity observations. *Acta. Cryst. A.* **34**, 517-525.
- [103] Winn, M.D. *et al.* (2011). Overview of the CCP4 suite and current developments. *Acta. Cryst. D.* **67**, 235-242.
- [104] Kantardjieff, K.A. & Rupp, B. (2003). Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. *Protein Sci.* **12**, 1865-1871.
- [105] Roussel, A., Inisan, A., Knoops-Mouthuy, E. & Cambillau, E. (1998). Turbo-Frodo version OpenGL.1. *University of Marseille.*
- [106] Delano, W.L. (2003). The PyMol molecular graphic system. *DeLano Scientific LLC, San Carlos, CA, USA.*
- [107] Accelrys, Inc. (2005). Cerius2 Modeling Environment. Release 4.8. *San Diego: Accelrys Software Inc.*
- [108] Vagin, A.A. & Teplyakov, A. (1997). MOLREP: an automated program for molecular replacement. *J. Appl. Cryst.* **30**, 1022-1025.
- [109] Navaza, J. (1994). AMoRe: an automated package for molecular replacement. *Acta Cryst. A.* **50**, 157-163.
- [110] McCoy, A.J. *et al.* (2007). Phaser crystallographic software. *J. Appl. Cryst.* **40**, 658-674.
- [111] Vagin, A.A. *et al.* (2004). REFMAC5 dictionary: "Organization of Prior Chemical Knowledge and Guidelines for its Use". *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2184-2195.
- [112] Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126-2132.
- [113] Carugo, O. & Djinovic-Carugo, K. (2013). Half a century of Ramachandran plots. *Acta Crystallogr. D Biol. Crystallogr.* **69**, 1333-1341.
- [114] Chen, V.B. *et al.* (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12-21.

- [115] Parkinson, G., Vojtechovsky, J., Clowney, L., Brünger, A.T. & Berman, H.M. (1996) New parameters for the refinement of nucleic acid-containing structures. *Acta Crystallogr. D Biol. Crystallogr.* **52**, 57-64.
- [116] Engh, R.A. & Huber, R. (2001). Structure quality and target parameters. A: *International Tables for Crystallography*, Vol. F (eds. Rossmann M.G. & Arnold, E.) 382–392 (Kluwer Academic Publishers, Dordrecht).
- [117] García-Segura, J.M. *et al.* (1996). *Técnicas instrumentales de análisis en bioquímica*. (Ed. Síntesis, Madrid).
- [118] Hellman, L.M. & Fried, M.G. (2007) Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions. *Nat. Protoc.* **2**, 1849-1861.
- [119] Fried, M. & Crothers, D.M. (1981). Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. *Nucleic Acids Res.* **9**, 6505-6525.
- [120] Garner, M.M. & Revzin, A. (1981). A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of the Escherichia coli lactose operon regulatory system. *Nucleic Acids Res.* **9**, 3047-3060.
- [121] Sambrook, J. & Russell D.W. (2001). *Molecular Cloning: A Laboratory Manual*. 3rd Edition. vol. 1, 5.1. (Cold Spring Harbor Laboratory Press, New York).
- [122] Zeman, S.M., Depew, K.M., Danishefsky, S.J. & Crothers, D.M. (1998). Simultaneous determination of helical unwinding angles and intrinsic association constants in ligand-DNA complexes: the interaction between DNA and calichearubicin B. *Proc. Natl. Acad. Sci. USA.* **95**, 4327-4332.
- [123] Bowater, R., Aboul-ela, F. & Lilley, D.M.J. (1992). Two-dimensional gel electrophoresis of circular DNA topoisomers. *Methods Enzym.* **212**, 105-120.
- [124] Peck, L.J. & Wang, J.C. (1983). Energetics of B-to-Z transition in DNA. *Proc. Natl. Acad. Sci. USA.* **80**, 6206-10.
- [125] Clark, D.J. & Leblanc, B. (2009). Analysis of DNA Supercoiling Induced by DNA–Protein Interactions. A: *DNA-Protein Interactions* (eds. Moss, T. & Leblanc, B.) *Methods Mol Biol.* **543**. 523-535 (Humana Press, a part of Springer Science + Business Media, New York).
- [126] Haugstad, G. (2012). *Atomic Force Microscopy: Understanding Basic Modes and Advanced Applications*. (Ed. John Wiley & Sons, New Jersey).
- [127] Sawyer, L., Grubb, D. & Meyers, G.F. (2008). *Polymer Microscopy*. 46-50. (Springer Science + Business Media, New York).

- [128] Sánchez-Giraldo, R. (2009). Projecte Final de Carrera: *Expressió, purificació i cristal·lització de la proteïna HMGA1a humana amb oligonucleòtids rics en ATs*. ETSEIB, Barcelona.
- [129] Murphy, F.V. 4th, Sehy, J.V., Dow, L.K., Gao, Y.G. & Churchill, M.E. (1999). Co-crystallization and preliminary crystallographic analysis of the high mobility group domain of HMG-D bound to DNA. *Acta Crystallogr. D Biol. Crystallogr.* **55**, 1594-1597.
- [130] Giegé, R. & Ducruix, A. (1992). An introduction to the crystallogenesis of biological macromolecules. A: *Crystallization of Nucleic Acids and Proteins*. (eds. Ducruix, A. & Giegé, R.) 1-15 (Oxford University Press, New York).
- [131] Ng, J.D. *et al.* (1996). The crystallization of biological macromolecules from precipitates: evidence for Ostwald ripening. *J. Crystal Growth*, **168**, 50-62.
- [132] Carter, E.S. 2nd & Tung, C.S. (1996). NAMOT2--a redesigned nucleic acid modeling tool: construction of non-canonical DNA structures. *Comput Appl Biosci.* **12**, 25-30.
- [133] Lovell, S.C. *et al.* (2003). Structure validation by C α geometry: phi,psi and C β deviation. *Proteins.* **50**, 437-450.
- [134] Luscombe, N.M., Laskowski, R.A. & Thornton, J.M. (1997). NUCPLOT: a program to generate schematic diagrams of protein-nucleic acid interactions. *Nucleic Acids Res.* **25**, 4940-4955.
- [135] Takahara, P.M., Frederick, C.A. & Lippard, S.J. (1996). Crystal Structure of the Anticancer Drug Cisplatin Bound to Duplex DNA. *J. Am. Chem. Soc.* **118**, 12309-12321.
- [136] Wang, J. *et al.* (2013). Redox-sensitive structural change in the A-domain of HMGB1 and its implication for the binding to cisplatin modified DNA. *Biochem Biophys Res Commun.* **441**, 701-706.
- [137] Maiti, R., Van Domselaar, G.H., Zhang, H. & Wishart, D.S. (2004). SuperPose: a simple server for sophisticated structural superposition. *Nucleic Acids Res.* **32**, (Web Server issue) W590-4.
- [138] Park, S. & Lippard, S.J. (2011) Redox state-dependent interaction of HMGB1 and cisplatin-modified DNA. *Biochemistry.* **50**, 2567-2574.
- [139] Venereau, E. *et al.* (2012) Mutually exclusive redox forms of HMGB1 promote cell recruitment or proinflammatory cytokine release. *J. Exp. Med.* **209**, 1519-1528
- [140] Javaherian, K., Liu, J.F. & Wang, J.C. (1978). Nonhistone proteins HMG1 and HMG2 change the DNA helical structure. *Science.* **199**, 1345-1346.

- [141] Carballo, M., Puigdomènech, P., Tancredi, T. & Palau, J. (1984). Interaction between domains in chromosomal protein HMG-1. *EMBO J.* **3**, 1255-1261.
- [142] Stros, M., Stokrová, J. & Thomas, J.O. (1994). DNA looping by the HMG-box domains of HMG1 and modulation of DNA binding by the acidic C-terminal domain. *Nucleic Acids Res.* **22**, 1044-1051.
- [143] Teo, S.H., Grasser, K.D. & Thomas, J.O. (1995). Differences in the DNA-binding properties of the HMG-box domains of HMG1 and the sex-determining factor SRY. *Eur. J. Biochem.* **230**, 943-950.
- [144] Nakamura, Y., Yoshioka, K., Shirakawa, H. & Yoshida, M. (2001). HMG box A in HMG2 protein functions as a mediator of DNA structural alteration together with box B. *J. Biochem.* **129**, 643-651.
- [145] Paull, T.T. & Johnson, R.C. (1995). DNA looping by *Saccharomyces cerevisiae* high mobility group proteins NHP6A/B. Consequences for nucleoprotein complex assembly and chromatin condensation. *J. Biol. Chem.* **270**, 8744-8754.
- [146] Vologodskaja, M.Y. & Vologodskii, A.V. (1999). Effect of magnesium on cruciform extrusion in supercoiled DNA. *J. Mol. Biol.* **289**, 851-859.
- [147] Slama-Schwok, A. *et al.* (2000). Structural changes induced by binding of the high-mobility group I protein to a mouse satellite DNA sequence. *Biophys. J.* **78**, 2543-2559.
- [148] Rohs, R. *et al.* (2009). The role of DNA shape in protein-DNA recognition. *Nature.* **461**, 1248-53.
- [149] Seeman, N.C., Rosenberg, J.M. & Rich, A. (1976). Sequence-specific recognition of double helical nucleic acids by proteins. *Proc. Natl. Acad. Sci. USA.* **73**, 804-808.
- [150] Aravind, L. & Landsman, D. (1998). AT-hook motifs identified in a wide variety of DNA-binding proteins. *Nucleic Acids Res.* **26**, 4413-4421.
- [151] Baker, S.A. *et al.* (2013) An AT-hook domain in MeCP2 determines the clinical course of Rett syndrome and related disorders. *Cell.* **152**, 984-996.
- [152] Ansari, A. & Kuznetsov, S.V. (2011). Dynamics and Mechanism of DNA-Bending Proteins in Binding Site Recognition. A: *Biophysics of DNA-Protein Interactions. Biological and Medical Physics, Biomedical Engineering* (eds. Williams, M.C. & Maher, L.J.) 107-142 (Springer Science + Business Media, New York).
- [153] Bewley, C.A., Gronenborn, A.M. & Clore, G.M. (1998). Minor groove-binding architectural proteins: structure, function, and DNA recognition. *Annu. Rev. Biophys. Biomol. Struct.* **27**, 105-131.

- [154] Olson, W.K., Gorin, A.A., Lu, X.J., Hock, L.M. & Zhurkin, V.B. (1998). DNA sequence-dependent deformability deduced from protein-DNA crystal complexes. *Proc. Natl. Acad. Sci. USA*. **95**, 11163-11168.
- [155] Kim, J.L., Nikolov, D.B. & Burley, S.K. (1993). Co-crystal structure of TBP recognizing the minor groove of a TATA element. *Nature*. **365**, 520-527.
- [156] Sawaya, M.R. (2007). Characterizing a Crystal From an Initial Native Dataset. A: *Macromolecular Crystallography Protocols. Structure determination*. (ed. Doubl  , S.) Methods Mol. Biol., **364** (Humana Press, New Jersey).
- [157] Horton, H.R., Moran, L.A., Scrimgeour, K.G., Perry, M.D., & Rawn J.D. (2006) *Principles of Biochemistry*. 4th edition. (Pearson Prentice Hall, New Jersey).
- [158] Alberts, B. *et al.* (2002). *Molecular Biology of the Cell*. 4th edition. (Garland Science, New York).

7.2. Bibliografia complementària

Llibres

- Barceló Mairata, F. (2003) *Técnicas instrumentales en bioquímica y biología*. (Universitat de les Illes Balears, Palma).
- Bates, A.D. & Maxwell, A. (1993). *DNA topology*. (IRL Press at Oxford University Press, Oxford).
- Berg, J.M., Tymoczko, J.L. & Stryer, L. (2002). *Biochemistry*. 5th edition. (W.H. Freeman, New York).
- Ducruix, A. & Giegé, R. (1992). *Crystallization of Nucleic Acids and Proteins*. (Oxford University Press, New York).
- García-Segura, J.M. *et al.* (1996). *Técnicas instrumentales de análisis en bioquímica*. (Ed. Síntesis, Madrid).
- GE Healthcare Handbooks. (2000). *Gel Filtration. Principles and Methods*. (General Electric Company).
- GE Healthcare Handbooks. (2002). *GST Gene Fusion System. Handbook*. (General Electric Company).
- Gòdia, F. & López, J. (1998). *Ingeniería Bioquímica*. (Ed. Síntesis, Madrid).
- Gómez-Moreno, C. & Sancho Sanz, J. (2003). *Estructura de proteínas*. (Ed. Ariel, Barcelona).
- McPherson, A. (2003). *Introduction to macromolecular crystallography*. (Wiley-Liss. New Jersey).
- Neidle, S. (2007). *Principles of Nucleic Acid Structure*. 1st edition. (Academic Press, London).
- Peretó, J., Sendra, R., Pamblanco, M & Bañó, C. (1997). *Fonaments de bioquímica*. (Universitat de València, València).
- Rhodes, G. (2000). *Crystallography made crystal clear: a guide for users of macromolecular models*. 2nd edition (Academic Press, San Diego).
- Sambrook, J. & Russell D.W. (2001). *Molecular Cloning: A Laboratory Manual*. 3rd edition. (Cold Spring Harbor Laboratory Press, New York).
- Subirana, J.A. (1985). *Estructura del ADN*. (Alhambra, Madrid).
- Watson, J.D. *et al.* (2008). *Molecular biology of the gene*. 6th edition. (Pearson/Benjamin Cummings; Cold Spring Harbor Laboratory Press, New York).

Pàgines web

- ALBA SYNCHROTRON [[http:// www.cells.es](http://www.cells.es)]
- BERNHARD RUPP'S MACROMOLECULAR CRYSTALLOGRAPHY WEBSITE [<http://www.ruppweb.org>]
- BIO-RAD LABORATORIES [<http://www.bio-rad.com/>]
- CCP4 WIKI [<http://ccp4wiki.org/>]
- CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS (CSIC). *Cristalografía de Rayos X*. [<http://www.xtal.iqfr.csic.es/Cristalografia/index.html>]
- CRYO-CRYSTALLOGRAPHY AND DATA COLLECTION [<http://www.doe-mpi.ucla.edu/~sawaya/m230d/Data/data.html>]
- EUROPEAN SYNCHROTRON RADIATION FACILITY. [<http://www.esrf.eu/>]
- HAMPTON RESEARCH [<http://www.hamptonresearch.com/>]
- MILLIPORE BIOMANUFACTURING AND LIFE SCIENCE RESEARCH [<http://www.millipore.com/>]
- MOLPROBITY: MAIN PAGE [<http://molprobity.biochem.duke.edu/>]
- RCSB PROTEIN DATA BANK - RCSB PDB [<http://www.rcsb.org/>]
- X-RAY DATA COLLECTION COURSE [<http://xray0.princeton.edu/~phil/Facility/Guides/XrayDataCollection.html>]
- 3DNA HOMEPAGE -- NUCLEIC ACID STRUCTURES [<http://x3dna.org>]

Annexos

A. Seqüència, massa molecular i coef. d'extinció de les proteïnes utilitzades

Taula A.1. Seqüència aminoacídica, massa molecular, coeficient d'extinció i nombre de residus per a les diferents proteïnes estudiades en el present treball.

Nom	Massa molecular (Da)	Coef. Extinció (cm ⁻¹ M ⁻¹)	Coef. Extinció (ml/mg cm)	Nombre residus	Seqüència
HMGA1a	11676	-	6,23 (a 220 nm)	107	MSESSSKSSQ PLASKQEKDG TEKRGRGRPR KQPPVSPGTA LVGSQKEPSE VPTPKRPRGR PKGSKNKGAA KTRKTTTTTPG RKPRGRPKKL EKEEEEGISQ ESSEEEQ
HMGA1b	10679	-	6,37 (a 220 nm)	96	MSESSSKSSQ PLASKQEKDG TEKRGRGRPR KQPPKEPSEV PTPKRPRGRP KGSKNKGAAK TRKTTTTTPGR KPRGRPKKLE KEEEEEGISQE SSEEEQ
HMGA1b (1-79)	8701	-	6,37 (a 220 nm)	79	MSESSSKSSQ PLASKQEKDG TEKRGRGRPR KQPPKEPSEV PTPKRPRGRP KGSKNKGAAK TRKTTTTTPGR KPRGRPKKL
HMGB1 Box AS	8929	9770 (a 280 nm)	-	74+2*	GS KPRGKMSS YAFFVQTCRE EHKKKHPDAS VNFSEFSKCC SERWKTMSAK EKGKFEDMAK ADKARYEREM KTYIPP
HMGB1 Box BS	8430	10930 (a 280 nm)	-	73+2*	GS NAPKRPPS AFFLFCSEYR PKIKGEHPGL SIGDVAKKLG EMWNNTAADD KQPYEKAAK LKEKYEKDIA AYRAK
HMGB1 Box ABS	18484	20700 (a 280 nm)	-	158+2*	GS KPRGKMSS YAFFVQTCRE EHKKKHPDAS VNFSEFSKCC SERWKTMSAK EKGKFEDMAK ADKARYEREM KTYIPPKGET KKKFKDPNAP KRPPSAFFLF CSEYRPKIKG EHPGLSIGDV AKKLGEMWNN TAADDKQPYE KKAACLKEY EKDIAAYRAK
NHP6A	10802	10810 (a 280 nm)	-	93	MVTPREPKKR TTRKKKDPNA PKRALSAYMF FANENRDIVR SENPDITFGQ VGKKLGEKWK ALTPEEKQPY EAKAQADKKR YESEKELYNA TLA

Els coeficients d'extinció de les proteïnes HMGA1 han estat extrets de Reeves i Nissen [81]; en el cas del fragment HMGA1b(1-79) hem fet servir el mateix valor que per a la forma sencera. En canvi per a les proteïnes HMGB1 el coeficient d'extinció a 280 nm s'ha estimat amb una eina de càlcul de propietats de proteïnes (*Peptide Calculator*, Northwestern University). *Els dos aminoàcids addicionals a l'extrem N-terminal (ressaltats a la seqüència) són producte del clonatge.

B. Medis de cultiu

Els medis de cultiu que hem fet servir en aquest treball són els que es recullen en la taula B.1. Cal remarcar que els medis de cultiu s'han d'esterilitzar abans d'utilitzar-los. A la taula B.2 es descriu la preparació dels antibiòtics ampicil·lina i cloramfenicol i el reactiu IPTG.

Taula B.1. Medis de cultiu fets servits al treball.

Medis de cultiu			
LB	LB-agar	NZY+Broth	SOC*
<u>Per a 1 litre</u> 10 g de triptona 5 g extracte de llevat 10 g NaCl O bé 25 g LB Broth, Miller	<u>Per a 500 ml</u> (~40 càpsules) 500 ml LB 7,5 g agar	<u>Per a 1 litre</u> 10 g NZ amina 5 g extracte de llevat 5 g NaCl Suplement per a afegir just abans d'utilitzar: 12,5 ml MgCl ₂ 1 M 12,5 ml MgSO ₄ 1 M 20 ml glucosa 20%	<u>Composició</u> 2% triptona 0.5% extracte de llevat 10 mM NaCl 2,5 mM KCl 10 mM MgCl ₂ 10 mM MgSO ₄ 20 mM glucosa

*En aquest cas es va utilitzar medi comercial ja estèril (Invitrogen™).

Taula B.2. Preparació dels estocs d'antibiòtics i IPTG.

Antibiòtics		IPTG
Ampicil·lina (Amp)	Cloramfenicol (Cam)	
<u>Estoc mare</u> 50 mg/ml (0,5 g en 10 ml H ₂ O)	<u>Estoc mare</u> 34 mg/ml (0,34 g en 10 ml EtOH)	<u>Estoc mare</u> 1 M (2,38 g en 10 ml H ₂ O)
<u>Estoc de treball</u> 50 µg/ml	<u>Estoc de treball</u> 34 µg/ml	<u>Estoc de treball</u> 0,5 o 1 mM
Tots els estocs s'esterilitzen per filtració a 0,2 µm i es guarden aliquidats a -20°C		

C. Inducció amb IPTG

L'enzim T7 RNA polimerasa és el responsable d'iniciar la transcripció del gen clonat al plasmidi pET però normalment la seva acció es veu inhibida per la presència d'una molècula repressora, el repressor lac. L'isopropil- β -D-tiogalactopiranosid (IPTG) és un anàleg sintètic de la lactosa capaç d'unir-se al repressor lac inactivant-lo. D'aquesta manera es pot iniciar la transcripció del gen clonat i finalment la síntesi de la proteïna d'interès (veure fig. A.1).

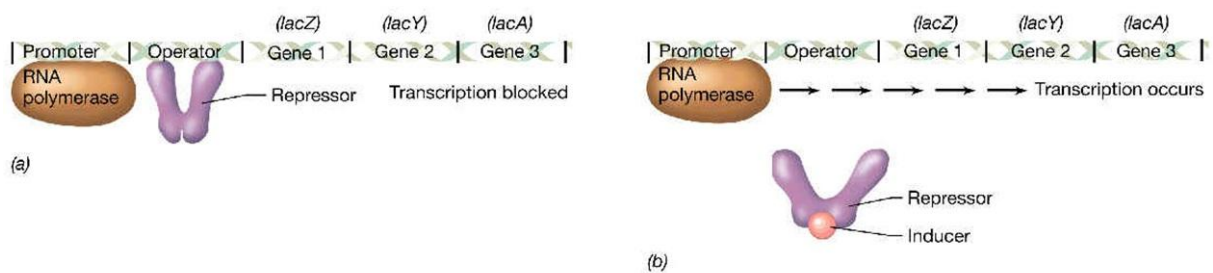


Fig. C.1. Mecanisme d'acció de l'inductor IPTG.

- a) La transcripció està inhibida per la presència del repressor. b) L'inductor IPTG s'uneix al repressor inactivant-lo i d'aquesta manera la RNA polimerasa inicia la transcripció.

D. Tampons comercials de cristal·lització

A continuació es detalla la composició dels tampons comercials utilitzats per a la realització dels assajos cristal·logràfics.

a) *Natrix HR2-116 (Hampton Research)*

Taula D.1. Composició dels tampons Natrix HR2-116.

Tampons matriu Natrix HR2-116	
1.	0.01 M Mg Chloride, 0.05 M MES pH 5.6, 2.0 M Lithium Sulfate
2.	0.01 M Mg Acetate, 0.05 M MES pH 5.6, 2.5 M Ammonium Sulfate
3.	0.1 M Mg Acetate, 0.05 M MES pH 5.6, 20% MPD
4.	0.2 M K Chloride, 0.01 M Mg Sulfate, 0.05 M MES pH 5.6, 10% PEG 400
5.	0.2 M K Chloride, 0.01 M Mg Chloride, 0.05 M MES pH 5.6, 5% PEG 8000
6.	0.1 M Ammonium Sulfate, 0.01 M Mg Chloride, 0.05 M MES pH 5.6, 20% PEG 8000
7.	0.02 M Mg Chloride, 0.05 M MES pH 6.0, 15% iso-Propanol
8.	0.005 M Mg Sulfate, 0.1 M Ammonium Acetate, 0.05 M MES pH 6.0, 0.6 M NaCl
9.	0.1 M K Chloride, 0.01 M Mg Chloride, 0.05 M MES pH 6.0, 10% PEG 400
10.	0.005 M Mg Sulfate, 0.05 M MES pH 6.0, 5% PEG 4000
11.	0.01 M Mg Chloride, 0.05 M Na Cacodylate pH 6.0, 1.0 M Lithium Sulfate
12.	0.01 M Mg Sulfate, 0.05 M Na Cacodylate pH 6.0, 1.8 M Lithium Sulfate
13.	0.015 M Mg Acetate, 0.05 M Na Cacodylate pH 6.0, 1.7 M Ammonium Sulfate
14.	0.1 M K Chloride, 0.025 M Mg Chloride, 0.05 M Na Cacodylate pH 6.0, 15% iso-Propanol
15.	0.04 M Mg Chloride, 0.05 M Na Cacodylate pH 6.0, 5% MPD
16.	0.04 M Mg Acetate, 0.05 M Na Cacodylate pH 6.0, 30% MPD
17.	0.2 M K Chloride, 0.01 M Ca Chloride, 0.05 M Na Cacodylate pH 6.0, 10% PEG 4000
18.	0.01 M Mg Acetate, 0.05 M Na Cacodylate pH 6.5, 1.3 M Lithium Sulfate
19.	0.01 M Mg Sulfate, 0.05 M Na Cacodylate pH 6.5, 2.0 M Ammonium Sulfate
20.	0.1 M Ammonium Acetate, 0.015 M Mg Acetate, 0.05 M Na Cacodylate pH 6.5, 10% iso-Propanol
21.	0.2 M K Chloride, 0.005 M Mg Chloride, 0.05 M Na Cacodylate pH 6.5, 10% 1,6 Hexanediol
22.	0.08 M Mg Acetate, 0.05 M Na Cacodylate pH 6.5, 15% PEG 400
23.	0.2 M K Chloride, 0.01 M Mg Chloride, 0.05 M Na Cacodylate pH 6.5, 10% PEG 4000
24.	0.01 M Mg Acetate, 0.01 M Ca Chloride, 0.05 M Na Cacodylate pH 6.5, 10% PEG 4000
25.	0.08 M Mg Acetate, 0.05 M Na Cacodylate pH 6.5, 30% PEG 4000
26.	0.2 M K Chloride, 0.1 M Mg Acetate, 0.05 M Na Cacodylate pH 6.5, 10% PEG 8000
27.	0.2 M Ammonium Acetate, 0.01 M Mg Acetate, 0.05 M Na Cacodylate pH 6.5, 30% PEG 8000
28.	0.05 M Mg Sulfate, 0.05 M HEPES - Na pH 7.0, 1.6 M Lithium Sulfate
29.	0.01 M Mg Chloride, 0.05 M HEPES - Na pH 7.0, 4.0 M Lithium Chloride
30.	0.01 M Mg Chloride, 0.05 M HEPES - Na pH 7.0, 1.6 M Ammonium Sulfate
31.	0.005 M Mg Chloride, 0.05 M HEPES - Na pH 7.0, 25% PEG Monomethyl Ether 550
32.	0.2 M KCl, 0.01 M Mg Chloride, 0.05 M HEPES - Na pH 7.0, 20% 1,6 Hexanediol
33.	0.2 M Ammonium Chloride, 0.01 M Mg Chloride, 0.05 M HEPES - Na pH 7.0, 30% 1,6 Hexanediol
34.	0.1 M K Chloride, 0.005 M Mg Sulfate, 0.05 M HEPES - Na pH 7.0, 15% MPD
35.	0.1 M K Chloride, 0.01 M Mg Chloride, 0.05 M HEPES - Na pH 7.0, 5% PEG 400
36.	0.1 M K Chloride, 0.01 M Ca Chloride, 0.05 M HEPES - Na pH 7.0, 10% PEG 400
37.	0.2 M K Chloride, 0.025 M Mg Sulfate, 0.05 M HEPES - Na pH 7.0, 20% PEG 200
38.	0.2 M Ammonium Acetate, 0.15 M Mg Acetate, 0.05 M HEPES - Na pH 7.0, 5% PEG 4000
39.	0.1 M Ammonium Acetate, 0.02 M Mg Chloride, 0.05 M HEPES - Na pH 7.0, 5% PEG 8000
40.	0.01 M Mg Chloride, 0.05 M Tris HCl pH 7.5, 1.6 M Ammonium Sulfate
41.	0.1 M KCl, 0.015 M Mg Chloride, 0.05 M Tris HCl pH 7.5, 10% PEG Monomethyl Ether 550
42.	0.01 M Mg Chloride, 0.05 M Tris HCl pH 7.5, 5% iso-Propanol
43.	0.01 M Mg Chloride, 0.05 M Ammonium Acetate, 0.05 M Tris HCl pH 7.5, 10% MPD
44.	0.2 M K Chloride, 0.05 M Mg Chloride, 0.05 M Tris HCl pH 7.5, 10% PEG 4000
45.	0.025 M Mg Sulfate, 0.05 M Tris HCl pH 8.5, 1.8 M Ammonium Sulfate
46.	0.005 M Mg Sulfate, 0.05 M Tris HCl pH 8.5, 35% 1,6 Hexanediol
47.	0.1 M K Chloride, 0.01 M Mg Chloride, 0.05 M Tris HCl pH 8.5, 30% PEG 400
48.	0.01 M Ca Chloride, 0.2 M Ammonium Chloride, 0.05 M Tris HCl pH 8.5, 30% PEG 4000

b) JSCG Core Suite I (Qiagen)

Taula D.2. Composició del tampons de la matriu JSCG Core Suite I (1a Part).

Tampons matriu JSCG Core Suite I (1a part)	
1	A01 0.1 M CHES pH 9.5 20 % (w/v) PEG 8000
2	A02 0.1 M Bicine pH 8.5 20 % (w/v) PEG 6000
3	A03 0.05 M Lithium sulfate 0.05 M Sodium sulfate 0.05 M Tris-HCl pH 8.5 30 % (w/v) PEG 400
4	A04 0.2 M Ammonium dihydrogen phosphate 0.1 M Tris pH 8.5 50 % (v/v) MPD
5	A05 0.2 M Magnesium chloride 0.1 M Tris pH 8.5 3.4 M 1,6 Hexanediol
6	A06 0.05 M Magnesium chloride 0.1M Tris pH 8.5 40% (v/v) Ethanol
7	A07 0.2 M tri-Potassium citrate 20% (w/v) PEG 3350
8	A08 0.2 M tri-Sodium citrate 20% (w/v) PEG 3350
9	A09 0.2 M tri-Lithium citrate 20% (w/v) PEG 3350
10	A10 0.2 M Calcium acetate 0.1 M Imidazole pH 8.0 20 % PEG 1000
11	A11 0.2 M Potassium acetate 20% (w/v) PEG 3350
12	A12 0.2 M Magnesium acetate 20% (w/v) PEG 3350
13	B01 0.2 M Sodium chloride 0.1 M HEPES pH 7.5 20 % (w/v) PEG 3000
14	B02 0.1 M HEPES pH 7.5 20 % (w/v) PEG 8000
15	B03 0.1 M HEPES pH 7.5 10 % (w/v) PEG 8000
16	B04 0.19 M Calcium chloride 0.095 M HEPES pH 7.5 26.6% (v/v) PEG 400 5 % (v/v) Glycerol
17	B05 0.1 M HEPES pH 7.5 20 % (w/v) PEG 4000 10 % (v/v) Isopropanol
18	B06 0.8 M di-Sodium hydrogen phosphate 0.8 M di-Potassium hydrogen phosphate 0.1 M HEPES pH 7.5
19	B07 0.2 M di-Sodium tartrate 20% (w/v) PEG 3350
20	B08 0.2 M Calcium acetate hydrate 20% (w/v) PEG 3350
21	B09 0.2 M Potassium formate 20% (w/v) PEG 3350
22	B10 0.2 M Potassium Sodium tartrate 20% (w/v) PEG 3350
23	B11 0.2 M Sodium formate 20% (w/v) PEG 3350
24	B12 0.2 M Potassium fluoride 20% (w/v) PEG 3350
25	C01 0.2 M Ammonium acetate 20% (w/v) PEG 3350
26	C02 0.2 M Lithium nitrate 20% (w/v) PEG 3350
27	C03 0.1M Sodium cacodylate pH 6.5 5% (w/v) PEG 8000 40% (v/v) MPD
28	C04 0.2 M Magnesium chloride 0.1 M Tris pH 7.0 10 % (w/v) PEG 8000
29	C05 0.2 M Calcium acetate 0.1 M Tris pH 7.0 20 % (w/v) PEG 3000
30	C06 0.2 M Magnesium chloride 0.1 M Tris pH 7.0 2.5 M Sodium chloride
31	C07 0.1 M Tris pH 7.0 20 % (w/v) PEG 2000 MME
32	C08 0.2 M Sodium acetate 20% (w/v) PEG 3350
33	C09 0.2 M Potassium thiocyanate 20% (w/v) PEG 3350
34	C10 0.1 M HEPES pH 6.5 20 % (w/v) PEG 6000
35	C11 0.2 M Potassium nitrate 20% (w/v) PEG 3350
36	C12 0.2 M Sodium thiocyanate 20% (w/v) PEG 3350
37	D01 0.2 M Sodium iodide 20% (w/v) PEG 3350
38	D02 0.2 M Potassium chloride 20% (w/v) PEG 3350
39	D03 0.2 M Sodium chloride 20% (w/v) PEG 3350
40	D04 0.2 M Potassium iodide 20% (w/v) PEG 3350
41	D05 0.2 M Lithium chloride 20% (w/v) PEG 3350
42	D06 0.2 M Magnesium chloride 0.1M Sodium cacodylate pH 6.5 50% (v/v) PEG 200
43	D07 0.2 M di-Ammonium tartrate 20% (w/v) PEG 3350
44	D08 0.2 M Sodium sulfate 20% (w/v) PEG 3350
45	D09 0.2 M Ammonium formate 20% (w/v) PEG 3350
46	D10 0.1 M HEPES pH 7.5 10 % (w/v) PEG 6000 5 % (v/v) MPD
47	D11 1.6 M Sodium citrate pH 6.5
48	D12 0.2 M Magnesium acetate 0.1 M Sodium cacodylate pH 6.5 20 % (w/v) PEG 8000

Taula D.3. Composició del tampons de la matriu JSCG Core Suite I (2a Part).

Tampons matriu JSCG Core Suite I (2a part)

49	E01	0.2 M Ammonium nitrate	20% (w/v) PEG 3350
50	E02	0.2 M Ammonium chloride	20% (w/v) PEG 3350
51	E03	0.2 M Sodium chloride	0.1 M Na/K phosphate pH 6.2 10 % (w/v) PEG 8000
52	E04	0.2 M Ammonium iodide	20% (w/v) PEG 3350
53	E05	0.2 M Ammonium fluoride	20% (w/v) PEG 3350
54	E06	0.1M MES pH 6.0	5% (w/v) PEG 3000 30% (v/v) PEG 200
55	E07	0.2 M Calcium acetate	0.1 M MES pH 6.0 20 % (w/v) PEG 8000
56	E08	0.2 M Lithium sulfate	0.1 M MES pH 6.0 35 % (v/v) MPD
57	E09	0.2 M Ammonium sulfate	20% (w/v) PEG 3350
58	E10	0.1 M MES pH 5.0	40 % (v/v) MPD
59	E11	0.1 M MES pH 5.0	20 % (v/v) MPD
60	E12	0.1 M MES pH 5.0	20 % (w/v) PEG 6000
61	F01	0.1 M MES pH 5.0	10 % (w/v) PEG 6000
62	F02	0.2 M Magnesium sulfate	20% (w/v) PEG 3350
63	F03	0.2 M Magnesium formate	20% (w/v) PEG 3350
64	F04	0.2 M Magnesium nitrate	20% (w/v) PEG 3350
65	F05	0.2 M Magnesium chloride	20% (w/v) PEG 3350
66	F06	0.095 M Sodium citrate pH 5.6	19 % (v/v) Isopropanol 19 % (w/v) PEG 4000 5 % (v/v) Glycerol
67	F07	0.1 M Sodium citrate pH 5.6	20 % (v/v) Isopropanol 20 % (w/v) PEG 4000
68	F08	0.1 M Sodium citrate pH 5.5	20 % (w/v) PEG 3000
69	F09	0.2 M Sodium chloride	0.1M Phosphate-citrate pH 4.2 50% (v/v) PEG 200
70	F10	0.1M Phosphate-citrate pH 4.2	5% (w/v) PEG 1000 40% Ethanol
71	F11	0.2 M Lithium sulfate	0.1M Sodium acetate pH 4.5 50% (v/v) PEG 400
72	F12	0.1M Phosphate-citrate pH 4.2	40% (v/v) MPD
73	G01	0.18 M tri-Ammonium citrate	20% (w/v) PEG 3350
74	G02	0.1 M Sodium acetate	pH 5.0 20 % (v/v) MPD
75	G03	1.0 M Lithium chloride	0.1 M Citric Acid pH 5.0 10 % (w/v) PEG 6000
76	G04	0.1 M Citric Acid pH 4.0	20 % (w/v) PEG 6000
77	G05	0.1 M Citric Acid	10 % (w/v) PEG 6000
78	G06	0.1 M Citric Acid pH 4.0	5 % (w/v) PEG 6000
79	G07	0.2 M Potassium dihydrogen phosphate	20% (w/v) PEG 3350
80	G08	0.2 M Ammonium dihydrogen phosphate	20% (w/v) PEG 3350
81	G09	0.2 M Ammonium sulfate	0.1 M Sodium acetate pH 4.6 30 % (w/v) PEG 2000 MME
82	G10	0.1 M Sodium acetate	pH 4.6 8 % (w/v) PEG 4000
83	G11	0.2 M Ammonium sulfate	0.1 M Sodium acetate pH 4.6 25 % (w/v) PEG 4000
84	G12	0.02 M Calcium chloride	0.1 M Sodium acetate pH 4.6 30 % (v/v) MPD
85	H01	0.1 M Sodium acetate	pH 4.5 35 % (v/v) MPD
86	H02	0.1 M Sodium acetate	pH 4.5 20 % (w/v) PEG 3000
87	H03	0.2 M Sodium dihydrogen phosphate	20% (w/v) PEG 3350
88	H04	0.05 M Potassium dihydrogen phosphate	20 % (w/v) PEG 8000
89	H05	0.2 M Sodium chloride	0.1 M Phosphate-citrate pH 4.2 10 % (w/v) PEG 3000
90	H06	0.1 M Phosphate/citrate	pH 4.2 2.0 M Ammonium sulfate
91	H07	0.2 M Lithium sulfate	0.1 M Phosphate-citrate pH 4.2 20 % (w/v) PEG 1000
92	H08	0.1 M Citric Acid	pH 2.5 20 % (v/v) MPD
93	H09	0.1 M Citric Acid	pH 3.5 0.8 M Ammonium sulfate
94	H10	1.0 M Lithium chloride	0.1 M Citric Acid pH 4.0 20 % (w/v) PEG 6000
95	H11	1.0 M Lithium chloride	0.1 M Citric Acid pH 4.0 10 % (w/v) PEG 6000
96	H12	0.1 M Citric Acid	pH 4.0 5 % (w/v) PEG 6000

c) *Matrius PAC (Plataforma Automatitzada de Cristal·lografia, PCB)**Matriu PAC9m***Taula D.4. Composició del tampons de la matriu PAC9m (1a Part).**

Tampons matriu PAC9m (1a part)	
1	0.01 M Magnesium Chloride 0.05 M MES 5.6 2 M Lithium Sulfate Monohydrate
2	0.01 M Magnesium Acetate Tetrahydrate 0.05 M MES 5.6 2.5 M Ammonium Sulfate
3	0.1 M Magnesium Acetate Tetrahydrate 0.05 M MES 5.6 20 % V/V MPD
4	0.2 M Potassium Chloride 0.01 M Magnesium Sulfate Hydrate 0.05 M MES 5.6 10 % V/V PEG 400
5	0.2 M Potassium Chloride 0.01 M Magnesium Chloride 0.05 M MES 5.6 5 % W/V PEG 8000
6	0.1 M Ammonium Sulfate 0.01 M Magnesium Chloride 0.05 M MES 5.6 20 % W/V PEG 8000
7	0.02 M Magnesium Chloride 0.05 M MES 6 15 % V/V Isopropanol
8	0.1 M Ammonium Acetate 5 Mm Magnesium Sulfate Hydrate 0.05 M MES 6 0.6 M Sodium Chloride
9	0.1 M Potassium Chloride 0.01 M Magnesium Chloride 0.05 M MES 6 10 % V/V PEG 400
10	5 Mm Magnesium Sulfate Hydrate 0.05 M MES 6 5 % W/V PEG 4000
11	0.01 M Magnesium Chloride 0.05 M Cacodilate 6 1 M Lithium Sulfate Monohydrate
12	0.01 M Magnesium Sulfate Hydrate 0.05 M Cacodilate 6 1.8 M Lithium Sulfate Monohydrate
13	15 Mm Magnesium Acetate Tetrahydrate 0.05 M Cacodilate 6 1.7 M Ammonium Sulfate
14	0.1 M Potassium Chloride 25 Mm Magnesium Chloride 0.05 M Cacodilate 6 15 % V/V Isopropanol
15	0.04 M Magnesium Chloride 0.05 M Cacodilate 6 5 % V/V MPD
16	0.04 M Magnesium Acetate Tetrahydrate 0.05 M Cacodilate 6 30 % V/V MPD
17	0.2 M Potassium Chloride 0.01 M Calcium Chloride 0.05 M Cacodilate 6 10 % W/V PEG 4000
18	0.01 M Magnesium Acetate Tetrahydrate 0.05 M Cacodilate 6.5 1.3 M Lithium Sulfate Monohydrate
19	0.01 M Magnesium Sulfate Hydrate 0.05 M Cacodilate 6.5 2 M Ammonium Sulfate
20	0.1 M Ammonium Acetate 15 Mm Magnesium Acetate Tetrahydrate 0.05 M Cacodilate 6.5 10 % V/V Isopropanol
21	0.2 M Potassium Chloride 5 Mm Magnesium Chloride 0.05 M Cacodilate 6.5 10 % W/V 1,6-Hexanediol
22	0.08 M Magnesium Acetate Tetrahydrate 0.05 M Cacodilate 6.5 15 % V/V PEG 400
23	0.2 M Potassium Chloride 0.01 M Magnesium Chloride 0.05 M Cacodilate 6.5 10 % W/V PEG 4000
24	0.2 M Ammonium Acetate 0.01 M Calcium Chloride 0.05 M Cacodilate 6.5 10 % W/V PEG 4000
25	0.08 M Magnesium Acetate Tetrahydrate 0.05 M Cacodilate 6.5 30 % W/V PEG 4000
26	0.2 M Potassium Chloride 0.1 M Magnesium Acetate Tetrahydrate 0.05 M Cacodilate 6.5 10 % W/V PEG 8000
27	0.2 M Ammonium Acetate 0.01 M Magnesium Acetate Tetrahydrate 0.05 M Cacodilate 6.5 30 % W/V PEG 8000
28	0.05 M Magnesium Sulfate Hydrate 0.05 M HEPES 7 1.6 M Lithium Sulfate Monohydrate
29	0.01 M Magnesium Chloride 0.05 M HEPES 7 4 M Lithium Chloride
30	0.01 M Magnesium Chloride 0.05 M HEPES 7 1.6 M Ammonium Sulfate
31	5 Mm Magnesium Chloride 0.05 M HEPES 7 25 % V/V PEG MME 500
32	0.2 M Potassium Chloride 0.01 M Magnesium Chloride 0.05 M HEPES 7 20 % W/V 1,6-Hexanediol
33	0.2 M Ammonium Chloride 0.01 M Magnesium Chloride 0.05 M HEPES 7 30 % W/V 1,6-Hexanediol
34	0.1 M Potassium Chloride 5 Mm Magnesium Sulfate Hydrate 0.05 M HEPES 7 15 % V/V MPD
35	0.1 M Potassium Chloride 0.01 M Magnesium Chloride 0.05 M HEPES 7 5 % V/V PEG 400
36	0.1 M Potassium Chloride 0.01 M Calcium Chloride 0.05 M HEPES 7 10 % V/V PEG 400
37	0.2 M Potassium Chloride 25 Mm Magnesium Sulfate Hydrate 0.05 M HEPES 7 20 % V/V PEG 200
38	0.2 M Ammonium Acetate 0.15 M Magnesium Acetate Tetrahydrate 0.05 M HEPES 7 5 % W/V PEG 4000
39	0.1 M Ammonium Acetate 0.02 M Magnesium Chloride 0.05 M HEPES 7 5 % W/V PEG 8000
40	0.01 M Magnesium Chloride 0.05 M TRIS 7.5 1.6 M Ammonium Sulfate
41	0.1 M Potassium Chloride 15 Mm Magnesium Chloride 0.05 M TRIS 7.5 10 % V/V PEG MME 500
42	0.01 M Magnesium Acetate Tetrahydrate 0.05 M TRIS 7.5 5 % V/V Isopropanol
43	0.05 M Ammonium Acetate 0.01 M Magnesium Chloride 0.05 M TRIS 7.5 10 % V/V MPD
44	0.2 M Potassium Chloride 0.05 M Magnesium Chloride 0.05 M TRIS 7.5 10 % W/V PEG 4000
45	25 Mm Magnesium Sulfate Hydrate 0.05 M TRIS 8.5 1.8 M Ammonium Sulfate
46	5 Mm Magnesium Sulfate Hydrate 0.05 M TRIS 8.5 35 % W/V 1,6-Hexanediol
47	0.1 M Potassium Chloride 0.01 M Magnesium Chloride 0.05 M TRIS 8.5 30 % V/V PEG 400
48	0.2 M Ammonium Chloride 0.01 M Calcium Chloride 0.05 M TRIS 8.5 30 % W/V PEG 4000

Taula D.5. Composició del tampons de la matriu PAC9m (2a Part).

Tampons matriu PAC9m (2a part)	
49	0,1 M Sodium Acetate 4,6 6 %W/V PEG 8000
50	0,1 M Calcium Chloride 0,1 M Sodium Chloride 0,1 M MES 6 16 %W/V PEG 8000
51	0,05 M Magnesium Chloride 0,1 M HEPES 7,5 11 %W/V PEG 4000
52	0,1 M Calcium Chloride 0,1 M HEPES 7,5 20 %W/V PEG 8000
53	0,05 M Calcium Chloride 0,1 M Sodium Acetate 4,6 6 %W/V PEG 4000
54	0,05 M Magnesium Chloride 0,1 M Sodium Chloride 0,1 M HEPES 7,5 16 %W/V PEG 4000
55	0,1 M Sodium Chloride 0,1 M MES 6 11 %W/V PEG 2000
56	0,05 M Calcium Chloride 0,1 M Sodium Chloride 0,1 M MES 6 20 %W/V PEG 2000
57	0,1 M Calcium Chloride 0,1 M Sodium Chloride 0,1 M Sodium Acetate 4,6 16 %W/V PEG 2000
58	0,1 M Calcium Chloride 0,1 M Sodium Chloride 0,1 M HEPES 7,5 25 %W/V PEG 2000
59	0,1 M Magnesium Chloride 0,1 M MES 6 20 %W/V PEG 4000
60	0,1 M Sodium Acetate 4,6 11 %W/V PEG 8000
61	0,1 M Magnesium Chloride 0,1 M Sodium Chloride 0,1 M MES 6 25 %W/V PEG 4000
62	M 0,1 M HEPES 7,5 20 %W/V PEG 2000
63	0,1 M Calcium Chloride 0,1 M Sodium Acetate 4,6 11 %W/V PEG 8000
64	0,1 M MES 6 6 %W/V PEG 2000
65	0,05 M Magnesium Chloride 0,1 M Sodium Chloride 0,1 M Sodium Acetate 4,6 11 %W/V PEG 2000
66	0,1 M Sodium Chloride 0,1 M Sodium Acetate 4,6 20 %W/V PEG 4000
67	0,05 M Magnesium Chloride 0,1 M MES 6 16 %W/V PEG 4000
68	0,1 M Calcium Chloride 0,1 M Sodium Chloride 0,1 M MES 6 6 %W/V PEG 8000
69	M Magnesium Chloride 0,1 M Sodium Chloride 0,1 M Sodium Acetate 4,6 20 %W/V PEG 8000
70	0,1 M Calcium Chloride 0,1 M HEPES 7,5 11 %W/V PEG 4000
71	0,1 M HEPES 7,5 16 %W/V PEG 2000
72	0,1 M Sodium Chloride 0,1 M Sodium Acetate 4,6 6 %W/V PEG 4000
73	0,05 M Calcium Chloride 0,1 M Sodium Chloride 0,1 M Sodium Acetate 4,6 16 %W/V PEG 2000
74	0,05 M Calcium Chloride 0,1 M Sodium Chloride 0,1 M MES 6 20 %W/V PEG 4000
75	0,1 M Calcium Chloride 0,1 M Sodium Chloride 0,1 M MES 6 6 %W/V PEG 2000
76	0,1 M Calcium Chloride 0,1 M HEPES 7,5 11 %W/V PEG 8000
77	0,05 M Calcium Chloride 0,1 M HEPES 7,5 20 %W/V PEG 8000
78	0,1 M Calcium Chloride 0,1 M HEPES 7,5 16 %W/V PEG 4000
79	0,1 M MES 6 11 %W/V PEG 8000
80	0,1 M HEPES 7,5 6 %W/V PEG 8000
81	0,1 M Magnesium Chloride 0,1 M Sodium Chloride 0,1 M MES 6 16 %W/V PEG 8000
82	0,1 M Calcium Chloride 0,1 M Sodium Chloride 0,1 M Sodium Acetate 4,6 20 %W/V PEG 2000
83	0,1 M Sodium Chloride 0,1 M Sodium Acetate 4,6 11 %W/V PEG 4000
84	0,05 M Calcium Chloride 0,1 M Sodium Chloride 0,1 M MES 6 6 %W/V PEG 4000
85	0,05 M Calcium Chloride 0,1 M HEPES 7,5 6 %W/V PEG 4000
86	0,05 M Calcium Chloride 0,1 M Sodium Chloride 0,1 M Sodium Acetate 4,6 16 %W/V PEG 8000
87	0,1 M Calcium Chloride 0,1 M MES 6 20 %W/V PEG 2000
88	0,1 M Magnesium Chloride 0,1 M Sodium Chloride 0,1 M Sodium Acetate 4,6 11 %W/V PEG 2000
89	0,1 M Calcium Chloride 0,1 M Sodium Chloride 0,1 M Sodium Acetate 4,6 11 %W/V PEG 4000
90	0,1 M MES 6 6 %W/V PEG 2000
91	0,1 M Calcium Chloride 0,1 M Sodium Chloride 0,1 M HEPES 7,5 20 %W/V PEG 8000
92	0,05 M Magnesium Chloride 0,1 M Sodium Chloride 0,1 M HEPES 7,5 16 %W/V PEG 4000
93	0,05 M Calcium Chloride 0,1 M HEPES 7,5 6 %W/V PEG 2000
94	0,05 M Magnesium Chloride 0,1 M Sodium Acetate 4,6 6 %W/V PEG 8000
95	0,05 M Calcium Chloride 0,1 M Sodium Chloride 0,1 M MES 6 11 %W/V PEG 8000
96	0,1 M Calcium Chloride 0,1 M HEPES 7,5 25 %W/V PEG 4000

Matriu PAC10

Taula D.6. Composició del tampons de la matriu PAC10 (1a Part).

Tampons matriu PAC10 (1a part)	
1	A1 1 M Lithium Sulfate Monohydrate 0.05 M Tri-Sodium Citrate 5.5 5 Mm Magnesium Chloride
2	A2 10 % V/V Butanol 0.05 M Tri-Sodium Citrate 5.5
3	A3 20 % V/V MPD 0.05 M Tri-Sodium Citrate 5.5 5 Mm Magnesium Chloride
4	A4 30 % V/V Glycerol 0.05 M Tri-Sodium Citrate 5.5 5 Mm Magnesium Chloride
5	A5 15 % V/V 2,3-Butanediol 0.05 M Tri-Sodium Citrate 5.5 5 Mm Magnesium Chloride 0.15 M Sodium Chloride
6	A6 25 % V/V DMSO 0.05 M Tri-Sodium Citrate 5.5 5 Mm Magnesium Chloride
7	A7 10 % V/V PEG 200 0.05 M Tri-Sodium Citrate 5.5 5 Mm Magnesium Chloride
8	A8 10 % V/V PEG 400 0.05 M Tri-Sodium Citrate 5.5 0.2 M Potassium Chloride
9	A9 10 % W/V PEG MME 500 0.05 M Tri-Sodium Citrate 5.5 5 Mm Magnesium Chloride
10	A10 10 % W/V PEG 1000 0.05 M Tri-Sodium Citrate 5.5 5 Mm Magnesium Chloride 0.2 M Potassium Chloride
11	A11 5 % W/V PEG 8000 0.05 M MES 6 0.01 M Magnesium Chloride
12	A12 10 % W/V PEG 8000 0.05 M Tri-Sodium Citrate 5.5 0.1 M Magnesium Chloride
13	B1 1.75 M Ammonium Sulfate 0.05 M MES 6 0.01 M Magnesium Chloride
14	B2 20 % V/V Ethanol 0.05 M MES 6 0.01 M Magnesium Chloride
15	B3 10 % W/V PEG 3350 0.05 M MES 6 0.01 M Magnesium Chloride 0.2 M Potassium Chloride
16	B4 10 % V/V PEG 400 0.05 M MES 6 5 Mm Barium Chloride
17	B5 15 % W/V PEG MME 2000 0.05 M MES 6
18	B6 20 % W/V PEG 8000 0.05 M MES 6
19	B7 20 % V/V Propanediol 0.05 M Cacodilate 6.5 0.01 M Magnesium Chloride
20	B8 1.5 M Lithium Sulfate Monohydrate 0.05 M Cacodilate 6.5
21	B9 10 % V/V Ethanol 0.05 M Cacodilate 6.5 5 Mm Barium Chloride
22	B10 20 % V/V PEG 200 0.05 M Cacodilate 6.5 0.01 M Magnesium Chloride
23	B11 15 % V/V Isopropanol 0.05 M Cacodilate 6.5 0.1 M Potassium Chloride
24	B12 10 % V/V MPD 0.05 M Cacodilate 6.5 0.1 M Magnesium Chloride
25	C1 20 % W/V PEG 3350 0.05 M Cacodilate 6.5 0.01 M Magnesium Chloride
26	C2 15 % V/V Glycerol 0.05 M Cacodilate 6.5 0.05 M Cesium Chloride
27	C3 10 % V/V PEG 400 0.05 M Cacodilate 6.5 0.01 M Magnesium Chloride
28	C4 20 % V/V Ethylene Glycol 0.05 M Cacodilate 6.5
29	C5 10 % W/V PEG 1000 0.05 M Cacodilate 6.5 0.01 M Magnesium Chloride
30	C6 12.5 % W/V PEG MME 5000 0.05 M Cacodilate 6.5 0.01 M Magnesium Chloride 0.1 M Potassium Chloride
31	C7 1.5 M Sodium Chloride 0.05 M Cacodilate 7 5 % V/V Ethylene Glycol
32	C8 20 % V/V Ethylene Glycol 0.05 M Cacodilate 7 5 Mm Barium Chloride
33	C9 20 % V/V 2,3-Butanediol 0.05 M Cacodilate 7 0.01 M Magnesium Chloride
34	C10 10 % W/V PEG MME 500 0.05 M Cacodilate 7 0.01 M Magnesium Chloride
35	C11 30 % V/V MPD 0.05 M Cacodilate 7 0.1 M Magnesium Chloride
36	C12 15 % V/V Isopropanol 0.05 M Cacodilate 7 0.01 M Magnesium Chloride 0.1 M Potassium Chloride
37	D1 30 % W/V PEG MME 350 0.05 M Cacodilate 7 0.01 M Magnesium Chloride
38	D2 10 % V/V PEG 400 0.05 M Cacodilate 7 0.01 M Magnesium Chloride 0.1 M Potassium Chloride
39	D3 10 % W/V PEG 1000 0.05 M Cacodilate 7
40	D4 20 % W/V PEG 3000 0.05 M Cacodilate 7 0.2 M Potassium Chloride
41	D5 2.5 % W/V PEG 8000 0.05 M Cacodilate 7 0.4 M Potassium Chloride
42	D6 10 % W/V PEG 8000 0.05 M Cacodilate 7 0.01 M Barium Chloride
43	D7 15 % V/V Ethylene Glycol 0.05 M TRIS 7.5
44	D8 25 % W/V PEG MME 500 0.05 M TRIS 7.5
45	D9 30 % V/V Propanediol 0.05 M TRIS 7.5
46	D10 30 % V/V Glycerol 0.05 M TRIS 7.5 0.2 M Lithium Chloride
47	D11 15 % V/V Isopropanol 0.05 M TRIS 7.5 0.01 M Magnesium Chloride 0.2 M Potassium Chloride
48	D12 20 % V/V MPD 0.05 M TRIS 7.5 0.1 M Cesium Chloride

Taula D.7. Composició del tampons de la matriu PAC10 (2a Part).

Tampons matriu PAC10 (2a part)	
49	E1 20 % V/V PEG 400 0.05 M TRIS 7.5 0.01 M Magnesium Chloride 0.1 M Potassium Chloride
50	E2 25 % W/V PEG 1000 0.05 M TRIS 7.5
51	E3 40 % W/V PEG 6000 0.05 M TRIS 7.5 0.2 M Potassium Chloride
52	E4 20 % W/V PEG MME 2000 0.05 M TRIS 7.5 0.01 M Magnesium Chloride
53	E5 12.5 % W/V PEG 3350 0.05 M TRIS 7.5 0.2 M Potassium Chloride
54	E6 10 % W/V PEG 8000 0.05 M TRIS 7.5 5 % V/V Ethylene Glycol
55	E7 20 % V/V DMSO 0.05 M TRIS 8
56	E8 6 % V/V Glycerol 0.05 M TRIS 8 2.5 M Sodium Chloride
57	E9 20 % W/V PEG 3350 0.05 M TRIS 8 0.05 M Magnesium Chloride
58	E10 10 % V/V MPD 0.05 M TRIS 8 0.2 M Magnesium Chloride
59	E11 30 % V/V Ethylene Glycol 0.05 M TRIS 8 5 Mm Barium Chloride
60	E12 10 % V/V PEG 200 0.05 M TRIS 8 0.2 M Magnesium Chloride
61	F1 30 % V/V Tert-Butanol 0.05 M TRIS 8 0.2 M Lithium Chloride
62	F2 20 % V/V PEG 600 0.05 M TRIS 8 5 % V/V MPD
63	F3 10 % W/V PEG 1000 0.05 M TRIS 8 0.01 M Magnesium Chloride
64	F4 20 % W/V PEG 3350 0.05 M TRIS 8 0.01 M Magnesium Chloride
65	F5 5 % W/V PEG MME 2000 0.05 M TRIS 8 10 % V/V Ethylene Glycol
66	F6 20 % W/V PEG 8000 0.05 M TRIS 8 0.1 M Cesium Chloride
67	F7 25 % V/V 2,3-Butanediol 0.05 M TRIS 8 0.01 M Magnesium Chloride
68	F8 1 M Lithium Sulfate Monohydrate 0.05 M TRIS 8 5 % V/V Ethylene Glycol
69	F9 30 % V/V Ethylene Glycol 0.05 M TRIS 8 0.01 M Magnesium Chloride
70	F10 12.5 % W/V PEG 3000 0.05 M TRIS 8 0.1 M Magnesium Chloride
71	F11 25 % W/V PEG MME 5000 0.05 M TRIS 8 0.01 M Zinc Acetate
72	F12 30 % V/V Glycerol 0.05 M TRIS 8 0.01 M Magnesium Chloride
73	G1 15 % V/V Ethanol 0.05 M TRIS 8.5 6 % V/V Glycerol
74	G2 20 % V/V MPD 0.05 M TRIS 8.5 0.2 M Potassium Chloride
75	G3 20 % W/V PEG MME 2000 0.05 M TRIS 8.5 0.1 M Magnesium Chloride
76	G4 35 % V/V PEG 400 0.05 M TRIS 8.5 0.01 M Magnesium Chloride
77	G5 10 % V/V PEG 600 0.05 M TRIS 8.5 0.5 M Sodium Chloride
78	G6 20 % W/V PEG MME 350 0.05 M TRIS 8.5 0.05 M Cesium Chloride
79	G7 10 % W/V PEG 1000 0.05 M TRIS 8.5 0.01 M Zinc Acetate
80	G8 20 % V/V PEG 4000 0.05 M TRIS 8.5 0.2 M Potassium Chloride
81	G9 15 % W/V PEG MME 2000 0.05 M TRIS 8.5 0.05 M Cesium Chloride
82	G10 20 % W/V PEG 3350 0.05 M TRIS 8.5 0.01 M Magnesium Chloride
83	G11 5 % W/V PEG MME 5000 0.05 M TRIS 8.5 0.01 M Magnesium Chloride 12 % V/V Glycerol
84	G12 5 % W/V PEG 8000 0.05 M TRIS 8.5
85	H1 1.5 M Sodium Chloride 0.05 M CHES 9 5 % V/V MPD
86	H2 10 % W/V PEG MME 2000 0.05 M CHES 9 0.05 M Magnesium Chloride
87	H3 10 % V/V 2,3-Butanediol 0.05 M CHES 9
88	H4 12 % V/V Glycerol 0.05 M CHES 9 0.1 M Cesium Chloride
89	H5 15 % V/V Isopropanol 0.05 M CHES 9 0.05 M Magnesium Chloride
90	H6 20 % V/V MPD 0.05 M CHES 9 0.05 M Magnesium Chloride
91	H7 25 % V/V PEG 200 0.05 M CHES 9
92	H8 10 % W/V PEG MME 500 0.05 M CHES 9 0.5 M Magnesium Chloride
93	H9 20 % W/V PEG MME 5000 0.05 M CHES 9 0.01 M Magnesium Chloride
94	H10 25 % W/V PEG MME 350 0.05 M CHES 9 5 % V/V Isopropanol
95	H11 10 % W/V PEG 8000 0.05 M CHES 9 6 % V/V Glycerol
96	H12 12.5 % W/V PEG 8000 0.05 M CHES 9 0.2 M Magnesium Chloride

d) *Test de precrystal·lització (Hampton Research)*

El test de precrystal·lització (*Precrystallization Test*, PCT™, Hampton Research) es fa servir per a determinar la concentració apropiada de proteïna per als assajos cristal·logràfics. El PCT *kit* conté 4 reactius (veure taula D.8) per a avaluar aquesta concentració. En un primer moment es realitza el test amb els tampons A1 i A2 i si és necessari es repeteix amb B1 i B2. La concentració estàndard de la mostra de proteïna en els assajos de cristal·lització és de 15 mg/ml però aquesta pot variar des de 5 mg/ml fins a 100 mg/ml.

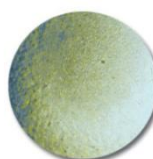
Taula D.8. Composició del tampons del test de precrystal·lització.

Tampons del test de precrystal·lització	
A1	0.1 M Tris HCl pH 8.5, 2.0 M Ammonium Sulfate
B1	0.1 M Tris HCl pH 8.5, 1.0 M Ammonium Sulfate
A2	0.2 M Mg Chloride, 0.1 M Tris HCl pH 8.5, 30% PEG 4000
B2	0.2 M Mg Chloride, 0.1 M Tris HCl pH 8.5, 15% PEG 4000

Per a l'assaig es parteix d'una mostra de proteïna (5-20 mg/ml) i es preparen 2 gotes amb els tampons A1 i A2 del *kit*. Les gotes es realitzen barrejant de 0,05 fins a 1 µl de la proteïna i un volum equivalent del tampó corresponent amb 0,5-1 ml del tampó al reservori. Després de 30 minuts es visualitza la gota amb un microscopi òptic i es segueixen les recomanacions de la figura D.1 segons l'aspecte de la gota.

PCT Reagent A1/B1 Results	PCT Reagent A2/B2 Results	Recommended Action
Heavy Amorphous Precipitate	Heavy Amorphous Precipitate	Dilute sample 1:1, repeat PCT
Clear	Clear	Concentrate sample to half the original volume, repeat PCT
Light granular precipitate	Clear	Perform Screen
Clear	Light granular precipitate	Perform Screen
Heavy Amorphous Precipitate	Light granular precipitate	Perform Screen
Heavy Amorphous Precipitate	Clear	Perform PCT with B1 & B2 / perform diagnostic testing
Clear	Heavy Amorphous Precipitate	Perform PCT with B1 & B2 / perform diagnostic testing

Light Precipitate



Heavy Amorphous Precipitate

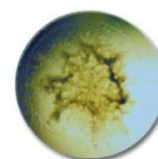


Fig. D.1. Resultats i recomanacions del PCT i imatges de dues gotes (una poc precipitada i l'altra molt).

E. La unitat cristal·lina

En les dues figures següents es detallen els sistemes cristal·lins, les xarxes de Bravais i els grups espacials. Els 7 sistemes cristal·lins s'obtenen a partir de la classificació dels paràmetres de la cel·la unitària. Segons la localització dels punts reticulars obtenim els 14 tipus de cel·les anomenades xarxes de Bravais (v. fig. E.1 a). I finalment els 65 grups espacials componen el conjunt d'operacions de simetria que s'han d'aplicar a la unitat asimètrica per a generar la cel·la unitària (v. fig. E.1 b).

a

The seven crystal systems	The fourteen Bravais Lattices
Triclinic $a \neq b \neq c$ $\alpha \neq \beta \neq \gamma$	$\alpha, \beta, \gamma \neq 90^\circ$
Monoclinic $a \neq b \neq c$ $\alpha = \gamma = 90^\circ \neq \beta$	$\alpha \neq 90^\circ, \beta, \gamma = 90^\circ$ and $\alpha \neq 90^\circ, \beta, \gamma = 90^\circ$
Orthorhombic $a \neq b \neq c$ $\alpha = \beta = \gamma = 90^\circ$	$a \neq b \neq c$ (repeated four times)
Tetragonal $a = b \neq c$ $\alpha = \beta = \gamma = 90^\circ$	$a \neq c$ (repeated two times)
Hexagonal $a = b \neq c$ $\alpha = \beta = 90^\circ, \gamma = 120^\circ$ Rhombohedral $a = b = c$ $\alpha = \beta = \gamma$	$a \neq c$ and $\alpha, \beta, \gamma \neq 90^\circ$
Cubic $a = b = c$ $\alpha = \beta = \gamma = 90^\circ$	

b

Triclinic	P1
Monoclinic	P2, P2 ₁ , C2
Orthorhombic	P222, P2 ₁ 2 ₁ 2 ₁ , P2 ₁ 2 ₁ 2 P222 ₁ , C222, C222 ₁ , F222 I222, I2 ₁ 2 ₁ 2 ₁
Tetragonal	P4, P4 ₁ , P4 ₂ , P4 ₃ , I4, I4 ₁ P422, P4 ₂ 2, P4 ₁ 22 P4 ₁ 2 ₁ 2, P4 ₂ 22, P4 ₂ 2 ₁ 2 P4 ₃ 2 ₁ 2, P4 ₃ 22, I4 ₁ 22, I422
Trigonal and rhombohedral	P3, P3 ₁ , P3 ₂ , R3 P312, P321, P3 ₁ 21, P3 ₁ 12 P3 ₂ 21, P3 ₂ 12, R32
Hexagonal	P6, P6 ₅ , P6 ₁ , P6 ₄ , P6 ₃ P6 ₂ , P622, P6 ₁ 22, P6 ₂ 22 P6 ₃ 22, P6 ₄ 22, P6 ₅ 22
Cubic	P23, F23, I23, P2 ₁ 3, I2 ₁ 3 P432, P4 ₁ 32, P4 ₂ 32, P4 ₃ 32 F432, F4 ₁ 32, I432, I4 ₁ 32

Fig. E.1. a) Classificació dels sistemes cristal·lins i de les xarxes de Bravais. b) Classificació dels grups espacials [97].

La nomenclatura P, C, I i F es refereix a la localització dels punts reticulars:

P: Cel·la primitiva. Conté un punt reticular situat un octau a cada vèrtex.

C: Centrada en una cara. Conté dos punts reticulars situats un als vèrtexs i l'altre amb dues meitats a cada cara d'un únic eix.

I: Centrada al cos. Conté dos punts reticulars situats un als vèrtexs i l'altre al centre de la cel·la.

F: Centrada a les cares. Conté quatre punts reticulars situats un als vèrtexs i els altres tres a meitats en el centre de totes les cares.

F. Funció de Patterson i funció d'autorotació

a) Funció de Patterson

La funció de Patterson suposa una simplificació de la informació continguda a la funció de densitat electrònica, degut a què suprimeix la informació de les fases i es substitueixen els mòduls dels factors d'estructura pels seus quadrats. Per tant, aquesta funció pot calcular-se directament a partir de la informació experimental (les intensitats corresponents a les reflexions dels diagrames de difracció). Així la funció de Patterson, introduïda per Arthur Lindo Patterson, va suposar la primera solució al problema de les fases.

$$\text{a) } \rho(xyz) = \frac{1}{V} \sum_h \sum_k \sum_l |F(hkl)| \cos(2\pi(hx + ky + lz - \phi(hkl)))$$

$$\text{b) } P(uvw) = \frac{1}{V} \sum_h \sum_k \sum_l |F(hkl)|^2 \cos(2\pi(hx + ky + lz))$$

Fig. F.1. Comparació entre la funció de la densitat electrònica i la funció de Patterson.

a) Funció de densitat electrònica que representa la transformada de Fourier entre l'espai cristal·lí, definit per $\rho(xyz)$, i l'espai recíproc, definit pels factors d'estructura $F(hkl)$. b) Funció de Patterson calculada en un punt (u,v,w) que presenta una simplificació de l'equació a).

La informació que proporciona la funció de Patterson és un mapa de vectors de posició entre àtoms (és a dir, posicions relatives). Els màxims de la funció de Patterson són majors com més gran sigui el nombre d'electrons dels àtoms implicats. Amb el mètode de resolució MIR es fa servir justament la introducció d'àtoms pesats per tal de calcular les fases de la funció de densitat electrònica mitjançant l'obtenció de mapes de Patterson.

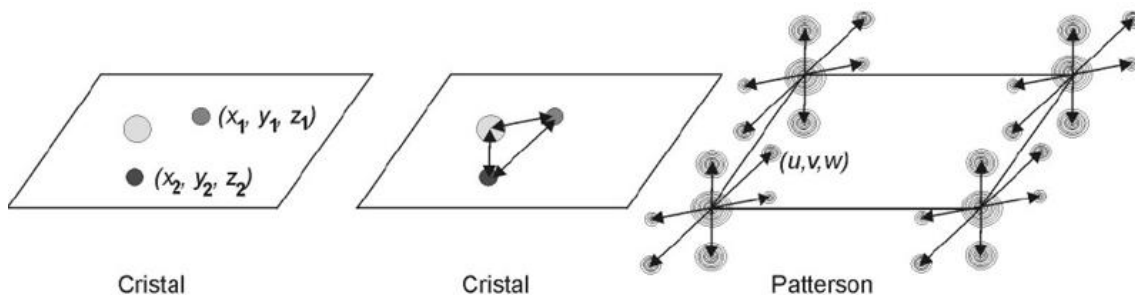


Fig. F.2. Exemple d'una projecció de la funció de Patterson derivada d'un cristall amb tres àtoms.

Per a l'obtenció de la funció de Patterson gràficament a partir d'una estructura coneguda d'un cristall (esquerra) en primer lloc es dibuixen tots els vectors possibles entre cada parella d'àtoms (centre). A continuació tots aquests vectors es traslladen a l'origen de la cel·la de la funció de Patterson (dreta). Els extrems d'aquests vectors corresponen als valors màxims de la funció de Patterson i les seves alçades són proporcionals al producte dels nombres atòmics dels àtoms implicats. A l'origen de la cel·la s'observa un màxim molt alt corresponent al vector interatòmic de cada àtom amb si mateix. Imatge extreta de [www.xtal.iqfr.csic.es].

L'espai cristal·lí està definit pel valor de la funció de densitat electrònica (ρ) en cada punt de la cel·la amb coordenades x,y,z , en canvi la funció de Patterson està definida en un nou espai (anomenat espai de Patterson) corresponent a la mateixa cel·la però on els seus punts venen definits per les coordenades genèriques u,v,w . Així cada parella d'àtoms, situats a les coordenades (x_1,y_1,z_1) i (x_2,y_2,z_2) , tindrà un vector de Patterson interatòmic (màxim de la funció de Patterson) amb les coordenades: $(u=x_1-x_2, v=y_1-y_2, w=z_1-z_2)$ (v. fig. F.2).

b) Funció d'autorotació

La funció d'autorotació (*self-rotation function*) aporta informació de les relacions d'orientació entre molècules dins de la cel·la unitària. És a dir, amb aquesta funció podem conèixer la presència d'eixos de simetria rotacionals i helicoidals, saber l'orientació d'aquests eixos i obtenir el grau de rotació sobre aquests eixos necessari per superposar una parella de molècules relacionades simètricament.

Aquesta funció revela la simetria rotacional del mapa de Patterson per la comparació de dos còpies idèntiques però no orientades del mapa de Patterson. Les dues còpies del mapa coincideixen quan una d'elles ha rotat un cert angle respecte el seu origen mantenint-se l'altra estacionària. L'espai de rotació està definit per tres angles de rotació (en coordenades polars): ϕ i ψ que indiquen la orientació dels eixos de simetria i κ que correspon al valor de la rotació necessària per fer coincidir dues còpies relacionades simètricament.

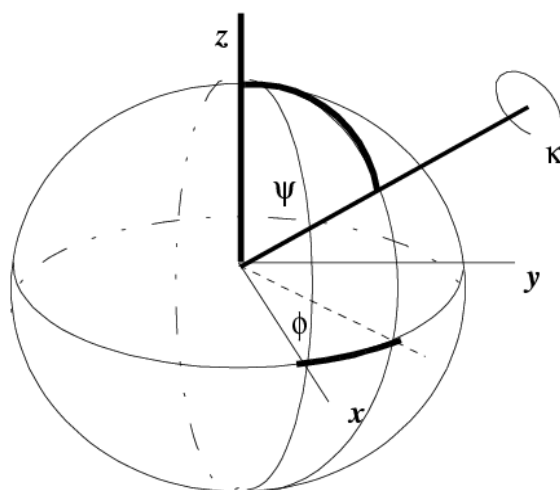


Fig. F.3. Representació dels angles de rotació (ϕ, ψ, κ).

L'eix representat comença al centre de l'esfera i talla la superfície d'aquesta en un punt definit per les coordenades polars ϕ, ψ (seria equivalent a definir una localització a la Terra amb la seva latitud i altitud). L'angle necessari per a rotar un objecte sobre aquest eix i per a fer-lo coincidir amb un simètricament relacionat és κ .

La funció de autorotació està definida de la següent forma:

$$R(\phi\psi\kappa) = \int_v P(u) \cdot P_r(u_r) du$$

Sent:

$P(u)$: la funció de Patterson per a un punt u

$P_r(U_r)$: la versió rotada de la funció de Patterson

V : el volum de l'esfera.

Els màxims (pics) del mapa corresponent a la funció de rotació són els angles que porten el mapa de Patterson en coincidència. Normalment el mapa de rotació es representa en seccions per a valor de l'angle κ constants. L'obtenció de pics sobre la secció corresponent a $\kappa=180^\circ$ indica la presència de eixos binaris, l'obtenció en $\kappa=120^\circ$ indica la presència d'eixos terciaris i així successivament. Cal pensar en cada secció de κ com una projecció esfèrica on els pics indiquen les localitzacions on els eixos de simetria entren i surten de l'esfera. També cal tenir en compte que les orientacions dels eixos a l'esfera corresponen als eixos del cristall.

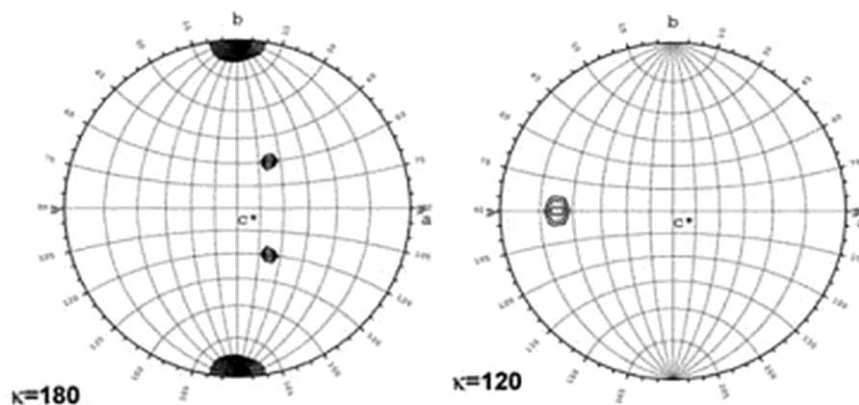


Fig. F.4. Exemple de funció de rotació [156].

Seccions per a valors de κ igual a 180° i 120° .

En aquest exemple s'observa la presència de tres eixos binaris ($\kappa=180^\circ$) i un terciari ($\kappa=120^\circ$).

G. Aminoàcids i enllaç peptídic

a) Aminoàcids

A la fig. G.1 es mostren els 20 aminoàcids que conformen les proteïnes classificats segons la naturalesa de les seves cadenes laterals.

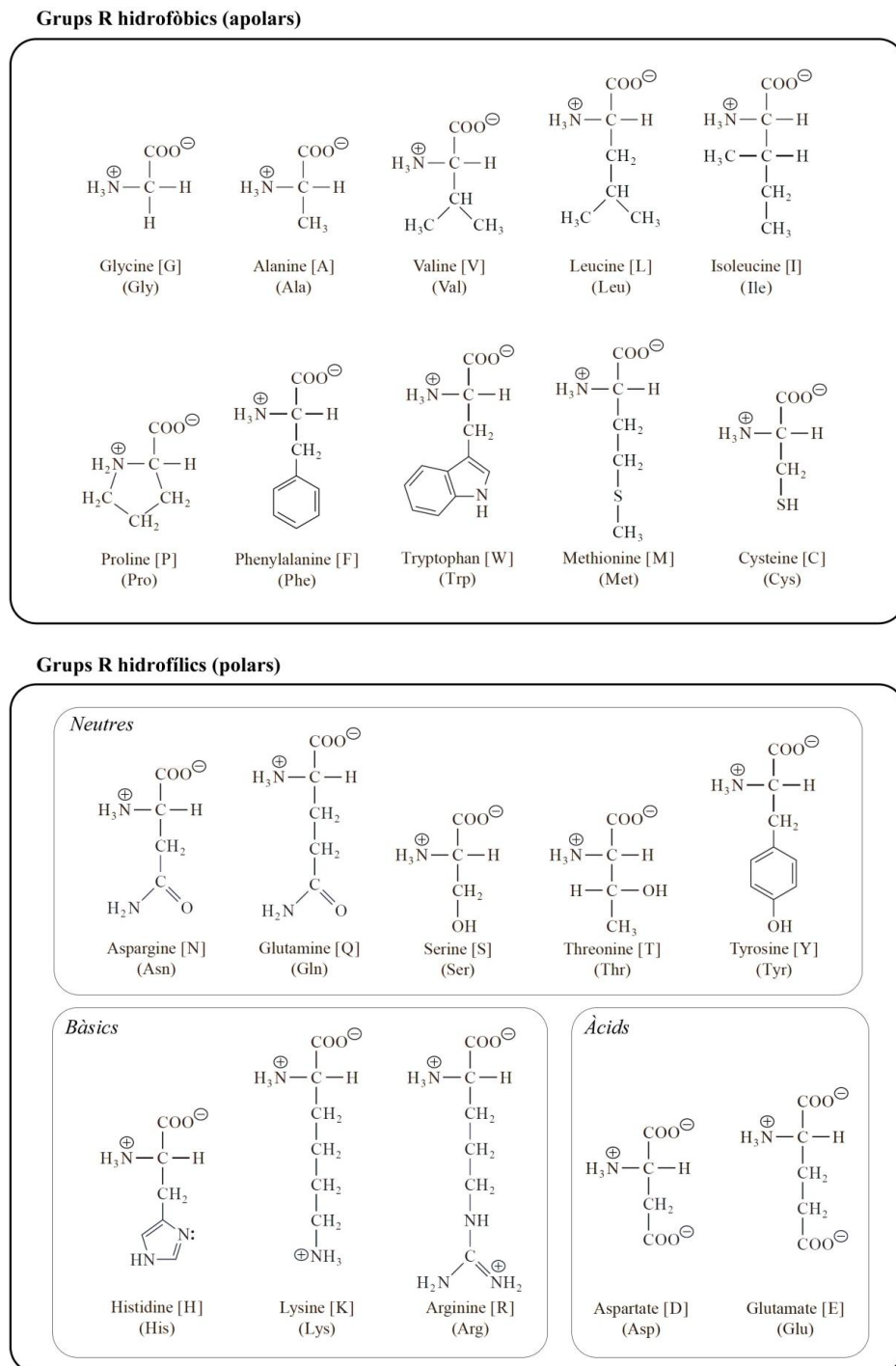


Fig. G.1. Classificació, estructura, nom i codis d'una i tres lletres dels aminoàcids [157].

b) Geometria de l'enllaç peptídic

L'enllaç peptídic és el resultat de la condensació del grup carbonil d'un aminoàcid amb el grup amino d'un altre amb la corresponent alliberació d'una molècula d'aigua. L'enllaç peptídic presenta ressonància entre dues formes fet que li confereix propietats de doble enllaç parcial. Per aquest motiu, la rotació al voltant de l'enllaç peptídic (ω) es veu restringida i els àtoms C carbonil, O, N, H i els dos C_α (carbonis α) queden en un mateix pla (pla gris a la fig. G.2).

L'enllaç peptídic pot ser *cis* quan els dos carbonis α es troben en el mateix semiplà ($\omega=0^\circ$) o *trans* quan cada carboni α queda en un costat del semiplà ($\omega=180^\circ$). En general, la forma *trans* és la de menor energia ja que situa les cadenes laterals dels residus (més voluminoses) més allunyades. La majoria dels enllaços peptídics presenten aquesta conformació *trans*, a excepció dels enllaços que forma la prolina que poden trobar-se en *cis* (conformació una mica més estable per a l'enllaç X-Pro).

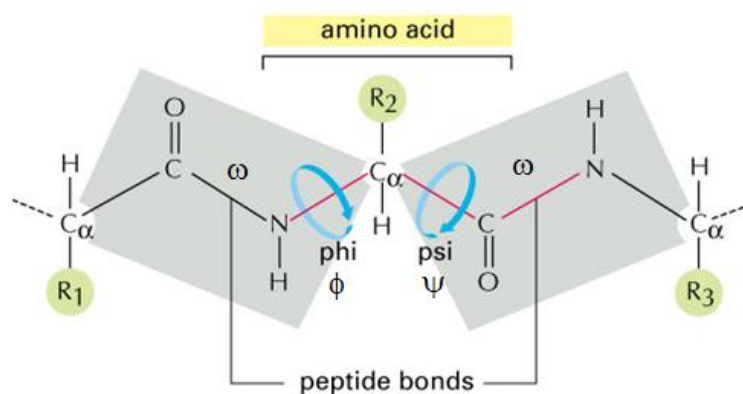


Fig. G.2. Geometria de l'enllaç peptídic [158].

La conformació global d'una proteïna queda definida pels valors dels seus angles ϕ (phi) i ψ (psi) (fig. G.2). L'angle ϕ mesura el gir entorn l'enllaç que formen el carboni α amb el nitrogen del mateix residu, mentre que l'angle ψ mesura el gir de l'enllaç que uneix el carboni α amb el carboni carbonil del mateix residu. Els valors d'aquests angles no estan fixats *a priori*, però, no poden adoptar qualsevol valor a la pràctica degut als possibles xocs estèrics entre els àtoms dels plans anterior i posterior o les cadenes laterals dels residus. Ramachandran i col·laboradors van determinar els valors permesos d'aquests angles i van proposar els diagrames de *Ramachandran* per a estudiar-los. El diagrama de *Ramachandran* és molt similar per als diferents aminoàcids excepte per a la glicina, que presenta una major llibertat degut a la seva curta cadena lateral que no provoca impediments estèrics.

