



Universitat de Girona

RESOLUCIÓ DE L'ESTRUCTURA TRIDIMENSIONAL DE L'HELICASA HEXAMÈTRICA DNAB

Raquel ARRIBAS BOSACOMA

ISBN: 978-84-692-6415-7
Dipòsit legal: GI-1068-2009



UNIVERSITAT DE GIRONA
ÀREA DE BIOQUÍMICA I BILOGIA MOLECULAR
DEPARTAMENT DE BIOLOGIA

Resolució de l'estructura tridimensional de l'helicasa hexamèrica DnaB

Memòria presentada per optar al grau de Doctor.

Aquest treball s'ha realitzat en el marc del Programa de Doctorat de *Ciències: Química i Física dels Àtoms, les Biomolècules i els Materials (Itinerari de Ciències de la Salut)*, de la Universitat de Girona.

Aquest treball ha estat realitzat al laboratori de cristal·lografia de proteïnes d'unió al DNA de l'institut de Biologia Molecular de Barcelona (Consell Superior d'Investigacions Científiques) i l'Institut de Recerca Biomèdica.

Treball realitzat per Raquel Arribas Bosacoma,
sota la direcció del Professor Miquel Coll i Capella
i la tutoria del Dr. Rafael de Llorens Duran.

Doctorand

Director

Tutor

Raquel
Arribas Bosacoma

Prof. Miquel
Coll i Capella

Dr. Rafael de
Llorens Duran

Girona, gener de 2009

TAULA D'ABREVIATURES

A: adenina

Å: angstrom

AAA+: *ATPases associated with various cellular activities*

AAV: virus adeno-associat

ABC: complex borà-dimetilamina

ADP: difosfat d'adenina

Aq_DnaB: proteïna DnaB d'*Aquifex aeolicus* sencera

Aq_DnaB[1-439]: proteïna DnaB d'*Aquifex aeolicus* sense els 29 residus finals

Aq_DnaB-CT: proteïna que compren només el domini C-terminal de la DnaB d'*A. aeolicus*.

ASCE: família *additional strand, conserved E*

ATP: trifosfat d'adenosina

BPV: virus del papil·loma boví

C: citosina

Cam: cloramfenicol

C-Terminal: extrem carboxi-terminal

Da: dalton

DLS: *dynamic light scattering*

DO: densitat òptica

dsDNA: àcid desoxiribonucleic de cadena doble

DTT: ditiotreitòl

EDTA: àcid etilendiaminotetraacètic

ESRF: sincrotró del *European Synchrotron Radiation Facility*

g: gram

G: guanidina

h: hora

HCV: virus de l'hepatitis C

Hg: mercuri

IPTG: isopropil- β -D-thio-galactòsid.

Kan: kanamicina
Kb: quilobase de nucleòtid
l: litre
LB: medi de cultiu Luria-Bertani
LTag: antigen T gran
μ: micro
m: metre
m: mil·li
M: molar
MAD: dispersió anòmala de longitud d'ona múltiple
MANT-ADP: 3'-O-(N-metiln-traniloil) 5'-difosfat
MCM: complex proteic de manteniment de mini cromosomes
MIR: reemplaçament isomòrfic múltiple
mRNA: RNA missatger
MWCO: (*molecular weight cut off*) mida d'exclusió molecular
n: nano
NMWL: (*nominal molecular weight limit*) límit de pes molecular
NTP: nucleòtid trifosfat
o/n: durant la nit (*over night* en anglès)
p/v: pes/volum
PAC: Plataforma Automatitzada de Cristal·lografia
pb: parells de bases de nucleòtid
PBS: tampó *phosphate buffered saline*
PDB: *protein data bank*. Base de dades on es dipositen les coordenades atòmiques dels models estructurals de les proteïnes.
pI: punt isoelèctric
Pol III: holoenzim polimerasa de DNA III
RNA: àcid ribonucleic
rpm: revolucions per minut
S/V: ràtio entre la superfície i el volum
s: segon
SAD: dispersió anòmala a una sola longitud d'ona
SDS: detergent aniònic dodecil sulfat sòdic

SDS-PAGE: electroforesi en gel de poliacrilamida en condicions desnaturalitzants (presència de SDS)

SeMet: selenometionina

SF: superfamília (d'helicases)

SIRAS: reemplaçament isomòrfic senzill amb senyal anòmala.

SLS: sincrotró *Swiss Light Source*

snRNA: RNA nucleolar petit

SSB: proteïnes d'unió a DNA de cadena senzilla (sigles de l'anglès: *single stranded DNA binding proteins*)

ssDNA: àcid desoxiribonucleic de cadena senzilla

SV40: virus de simi 40

T: timina

Ta: tàntal

TB: medi de cultiu Terrific Broth

TCEP: Tris[2-carboxietil] fosfina

Tet: tetraciclina

Th_DnaB: proteïna DnaB de *Thermotoga maritima* sencera

Thr: treonina

TNP-ADP: 2'(3')-O-(2,4,6-trinitrofenil) adenosina 5'-difosfat

TNP-AMP: 2'(3')-O-(2,4,6-trinitrofenil) adenosina 5'-monofosfat

TNP-ATP: 2'(3')-O-(2,4,6-trinitrofenil) adenosina 5'-trifosfat

Tris: tris(hidroximetil) metilamina

v/v: volum/volum

εADP: 1,N6-etenoadenosina difosfat

CODI D'AMINOÀCIDS:

| Codi d'una lletra | Codi de tres lletres | Aminoàcid |
|-------------------|----------------------|---------------|
| Ala | A | alanina |
| Arg | R | arginina |
| Asn | N | asparagina |
| Asp | D | àcid aspàrtic |
| Cys | C | cisteïna |
| Glu | E | àcid glutàmic |
| Gln | Q | glutamina |
| Gly | G | glicina |
| His | H | histidina |
| Ile | I | isoleucina |
| Leu | L | leucina |
| Lys | K | lisina |
| Met | M | metionina |
| Phe | F | fenilalanina |
| Pro | P | prolina |
| Ser | S | serina |
| Thr | T | treonina |
| Try | W | triptòfan |
| Tyr | Y | tirosina |
| Val | V | valina |

TAULA DE CONTINGUTS

| | |
|--|----------|
| TAULA D'ABREVIATURES | I |
| TAULA DE CONTINGUTS | V |
| RESUM | 1 |
| INTRODUCCIÓ | 5 |
| I. Helicases i translocases | 7 |
| 1. Classificació | 8 |
| 1.1 En funció de la seqüència aminoacídica | 8 |
| 1.2 Característiques comunes a helicases/translocases | 11 |
| 1.3 Classificacions mecàniques | 12 |
| 2. Helicases no hexamèriques: SF1 i SF2 | 15 |
| 2.1 Forma funcional | 15 |
| 2.2 Mecanismes de translocació del ssDNA | 16 |
| 2.2.1 Models pas a pas (<i>stepping</i>) | 16 |
| 2.2.2 Model Brownià | 17 |
| 2.3 Mecanismes de desenrotllament d'àcids nucleics de cadena doble | 18 |
| 2.3.1 Diferenciació mecanismes actius i passius | 18 |
| 2.3.2 Models per a SF1 i SF2 | 19 |
| 3. Helicases hexamèriques: superfamílies SF3, SF4, SF5, SF6 i SF7. | 21 |
| 3.1 Forma funcional | 21 |
| 3.1.2 Variacions respecte a un homohexàmer | 23 |
| 3.2 Mecanismes d'hidròlisi d'ATP | 25 |
| 3.2.1 Model dels tres llocs seqüencials (basat en l'ATPasa F1) | 25 |
| 3.2.2 Model de tots els llocs seqüencialment | 26 |
| 3.2.3 Model concertat | 27 |
| 3.3 Acoblament de l'activitat NTPasa a la translocació | 29 |
| 3.3.1 Model de transferència entre subunitats | 29 |
| 3.3.2 Model concertat (tot o res) | 31 |
| 3.3.3 Model d'escorta coordinada | 31 |
| 3.3.4 Consideracions finals | 34 |
| 3.4 Mecanismes de desenrotllament del dúplex | 34 |

| | |
|--|-----------|
| 3.4.1 Model d'exclusió estèrica | 35 |
| 3.4.2 Altres models..... | 36 |
| 4.4.3 Desenrotllament actiu o passiu? | 37 |
| II. La proteïna DnaB | 41 |
| 1. Caracterització històrica | 41 |
| 2. Organització en dominis | 43 |
| 3. Forma funcional | 44 |
| 3.1 Interaccions amb altres proteïnes | 47 |
| 4. Caracterització bioquímica | 47 |
| 4.1 Llocs d'unió a nucleòtid | 47 |
| 4.2 Interacció amb ssDNA..... | 48 |
| 4.3 Interacció amb substrats en forma de forqueta | 49 |
| 4.4 Activitat de desenrotllament de dúplexs..... | 49 |
| 4.4.1 Anàlisi quantitatiu de la reacció de desenrotllament catalitzada per la DnaB d' <i>E. coli</i> | 50 |
| 4.5 Translocació al llarg de dsDNA..... | 53 |
| 5. Estructures atòmiques de DnaB | 54 |
| 5.1 Domini N-terminal de la DnaB d' <i>E. coli</i> | 54 |
| 5.2 Monòmer de la DnaB de <i>Thermus aquaticus</i> | 56 |
| 5.3 Hexàmer de la DnaB de <i>Bacillus stearothermophilus</i> i de G40P del fag SPP1 de <i>Bacillus sp</i> .. | 58 |
| 5.3.1 Hexàmer de DnaB de <i>Bacillus stearothermophilus</i> | 59 |
| 5.3.2 Hexàmer de la G40P del fag SPP1 de <i>Bacillus sp</i> | 61 |
| III. Funcions de DnaB a la cèl·lula bacteriana..... | 63 |
| 1. Iniciació de la replicació bacteriana | 64 |
| 1.1 La proteïna iniciadora: DnaA..... | 64 |
| 1.2 L'origen de replicació bacterià: <i>oriC</i> | 66 |
| 1.3 Procés d'iniciació | 68 |
| 1.3.1 Carregament de DnaB a un <i>oriC</i> obert | 71 |
| 1.3.1.1 Procés | 72 |
| 1.3.1.2 Model de carregament de dues DnaB a una forqueta bidireccional..... | 73 |
| 1.4 Regulació de la iniciació: evitar sobreiniciacions | 75 |
| 2. Elongació de la replicació bacteriana | 77 |
| 2.1 Components del replisoma | 77 |
| 2.1.1. Holoenzim polimerasa de DNA III (Pol III) | 77 |
| 2.1.2 L'helicasa replicativa: DnaB | 81 |
| 2.1.3 La primasa: DnaG..... | 81 |
| 2.2 Arquitectura del replisoma | 83 |

| | |
|---|------------|
| 2.3 Interacció entre l'helicasa i la primasa..... | 84 |
| 2.3.1 Model estructural..... | 84 |
| 2.3.2 Control del cicle de la síntesi de fragments d'Okazaki..... | 86 |
| 2.3.3 Dissociació del complex i transferència de l'encebador al complex γ | 87 |
| 2.4 Carregament de la pinça β i entrada de Pol III..... | 87 |
| 2.5 Mecanismes per a la síntesi de la cadena líder i la retardada..... | 88 |
| OBJECTIUS..... | 91 |
| MATERIALS I MÈTODES..... | 93 |
| A. TÈCNiques DE BIOLOGIA MOLECULAR..... | 95 |
| I. Clonatge de les proteïnes d'interès..... | 97 |
| MATERIALS..... | 98 |
| DNA motlle..... | 98 |
| Vectors..... | 98 |
| Enzims..... | 98 |
| Encebadors..... | 99 |
| Kits comercials per a la purificació de DNA..... | 100 |
| Soca de clonatge..... | 100 |
| MÈTODES..... | 101 |
| 1. Determinació de les seqüències dels diferents constructes..... | 101 |
| 2. Disseny dels encebadors (<i>primers</i>)..... | 102 |
| 3. Obtenció del gen: amplificació mitjançant PCR..... | 103 |
| 4. Lligació dels productes gènics amplificats als vectors d'expressió..... | 105 |
| 4.1 Lligació a través d'un vector de clonatge..... | 105 |
| 4.2 Lligació al vector d'expressió..... | 106 |
| 5. Cribatge dels clons positius..... | 108 |
| II. Expressió heteròloga dels diferents constructes de DnaB a <i>Escherichia coli</i>..... | 109 |
| MATERIALS..... | 111 |
| A. GENERALS..... | 111 |
| Vectors d'expressió..... | 111 |
| Medis de cultiu bacterians:..... | 111 |
| Solucions d'antibiòtics..... | 112 |

| | |
|--|------------|
| Solució d'IPTG | 113 |
| Tampó PBS (phosphate buffered saline) | 113 |
| Soques bacterianes d'expressió..... | 113 |
| B. ESPECÍFICS PER PRODUIR PROTEÏNA AMB SELENOMETIONINA..... | 115 |
| Solucions: | 115 |
| Medis de cultiu bacterians:..... | 116 |
| MÈTODES..... | 117 |
| 1. Proves d'expressió a petita escala | 117 |
| 2. Proves d'expressió pilot a mitjana escala | 118 |
| 3. Sobreexpressió a gran escala | 119 |
| 4. Expressió de proteïna amb selenometionina | 120 |
| 4.1 Consideracions prèvies..... | 120 |
| 4.2 Protocol | 120 |
| III. PURIFICACIÓ DE LES PROTEÏNES D'INTERÈS | 123 |
| MATERIALS | 126 |
| Inhibidors de proteases. | 126 |
| Solució de DNasa. | 126 |
| Membrana de diàlisi: | 126 |
| Concentradors | 126 |
| Columnes de cromatografia líquida (FPLC)..... | 127 |
| MÈTODES..... | 128 |
| 1. Determinació del tampó de resuspensió cel·lular òptim..... | 128 |
| 2. Determinació del protocol de xoc tèrmic òptim..... | 129 |
| 3. Protocol de purificació | 130 |
| 3.1 Purificació d'Aq_DnaB i d'Aq_DnaB[1-439]:..... | 131 |
| 3.1.2 Primer dia. | 131 |
| 3.1.2.1 Extracció i clarificació | 131 |
| 3.1.2.2 Etapa de captura mitjançant un tractament de xoc tèrmic | 132 |
| 3.1.2.3 Preparació de la mostra per a la cromatografia de bescanvi aniònic | 132 |
| 3.1.3 Segon dia: bescanvi aniònic..... | 132 |
| 3.1.4 Tercer dia: gel filtració..... | 133 |
| 3.1.5 Característiques diferencials de la purificació d'Aq_DnaB[1-439] | 134 |
| 3.2 Purificació de Th_DnaB | 134 |
| 3.3 Purificació d'Aq_dnaB CT..... | 135 |

| | |
|---|------------|
| B. MÈTODES DE CRISTAL·LOGRAFIA DE RAIGS X | 137 |
| I. Cristal·lització | 139 |
| MATERIALS | 140 |
| Plaques de cristal·lització:..... | 140 |
| Solucions de cribatge de les condicions de cristal·lització (<i>screens</i>)..... | 140 |
| Cobreobjectes OptiClear™ | 141 |
| Vials de congelació..... | 142 |
| MÈTODES..... | 143 |
| 1. Cristal·logènesi..... | 143 |
| 1.1 El diagrama de fases de la cristal·logènesi | 144 |
| 1.2 Determinació de les condicions de cristal·lització..... | 145 |
| 1.3 Mètodes de cristal·lització de proteïnes. | 145 |
| 1.3.1 Cristal·lització per difusió de vapor | 146 |
| 1.4 Utilització de sistemes robòtics..... | 147 |
| 1.5 Detecció de les condicions de cristal·lització positives | 147 |
| 2. Tècniques de millora de la qualitat dels cristalls | 148 |
| 2.1 Mètodes previs a la cristal·lització. | 148 |
| 2.1.2 Separació física de les fases de nucleació i creixement: <i>seeding</i> | 148 |
| 2.1.2.1 Protocol <i>macroseeding</i> | 149 |
| 2.1.2.2 Protocol de <i>microseeding</i> | 150 |
| 2.1.2.3 Millorar la qualitat de les llavors: <i>seeding</i> serial. | 151 |
| 2.1.3 Limitar el nombre de nuclis a la prova de cristal·lització: separació dinàmica de la nucleació i el creixement..... | 152 |
| 2.1.3.1 Protocol de dilució per gota penjant..... | 152 |
| 2.1.3.2 Feeding i weeding..... | 153 |
| 2.1.4 Influenciar la cinètica de cristal·lització: utilització d'olis per alentir la difusió de vapor. | 154 |
| 2.2 Tècniques de millora postcristal·lització | 155 |
| 2.2.1 Annealing..... | 155 |
| 2.2.2 Deshidratació dels cristalls | 156 |
| 2.2.3 Submersió a major concentració de precipitant sense deshidratació | 157 |
| 2.2.4 Entrecreuament químic de cristalls per difusió de vapor..... | 158 |
| 2.2.5 Metilació reductiva de les lisines exposades..... | 158 |
| 3. Derivatització dels cristalls amb àtoms pesats | 160 |
| 3.1. Tractaments precristal·lització..... | 160 |
| 3.2. Tractaments postcristal·lització: <i>soaking</i> dels cristalls en solucions d'àtoms pesats..... | 161 |
| 4. Criocristal·lografia | 162 |

| | |
|--|------------|
| 4.1 El dany per radiació malmet els cristalls | 162 |
| 4.2 Objectius de la criocristal·lografia | 163 |
| 4.3 Crioprotectors | 164 |
| 4.4 Protocols de crioprotecció | 165 |
| 4.4.1 Transferència directa a la solució crioprotectora final | 165 |
| 4.4.2 Addició gradual de l'agent crioprotector. | 166 |
| 4.4.3 Equilibrat per pressió de vapor enfront de la solució crioprotectora final. | 167 |
| 4.4.4 Incorporació de l'agent crioprotector mitjançant diàlisi. | 167 |
| 4.5 Muntatge dels cristalls | 167 |
| II. DIFRACCIÓ DE RAIGS X I RESOLUCIÓ DE L'ESTRUCTURA..... | 171 |
| 1. Experiment de difracció de raigs X | 171 |
| 1.2 Recollida de dades de difracció..... | 172 |
| 1.3 Processament de les dades de difracció | 173 |
| 1.3.1 Indexat, integració i escalat..... | 173 |
| 2. Resolució de l'estructura | 174 |
| 2.1 Mètodes directes | 174 |
| 2.2 Reemplaçament molecular: | 175 |
| 2.3 Reemplaçament isomòrfic | 176 |
| 2.4 Dispersió anòmala (<i>anomalous scattering</i>) | 177 |
| 2.4.1 Dispersió anòmala de múltiple longitud d'ona (MAD) | 178 |
| 2.4.2 Dispersió anòmala a una sola longitud d'ona: SAD | 179 |
| 2.4.3 Combinar reemplaçament isomòrfic i dispersió anòmala..... | 179 |
| 3. Càlcul del mapa de densitat electrònica | 179 |
| 4. Millora de les fases | 180 |
| 4.1 Aplanament de solvent | 181 |
| 4.2 Mitjanat | 181 |
| 4.3 Histogram matching..... | 181 |
| 5. Construcció del model atòmic | 182 |
| 5.1 Fitejat automàtic i refinat | 182 |
| 5.1.1 Paràmetres estadístics | 183 |
| RESULTATS..... | 185 |
| 1. DnaB d' <i>Aquifex aeolicus</i> sencera: Aq_DnaB | 186 |
| 1.1 Clonatge | 186 |
| 1.2 Expressió | 187 |
| 1.3 Purificació | 188 |
| 1.3 Cristal·lització | 191 |

| | |
|--|------------|
| 2. DnaB de <i>Thermotoga maritima</i> : Th_DnaB..... | 193 |
| 2.1 Clonatge | 193 |
| 2.2 Expressió | 193 |
| 2.3 Purificació | 196 |
| 2.4 Cristal·lització..... | 198 |
| 3. Domini C-terminal de la DnaB d' <i>A. aeolicus</i> : Aq_DnaB-CT | 201 |
| 3.1 Estratègia de clonatge | 201 |
| 3.2 Expressió | 203 |
| 3.3 Purificació | 203 |
| 3.4 Cristal·lització..... | 206 |
| 4. Forma curta de la DnaB d' <i>A. aeolicus</i> : Aq_DnaB [1-439]..... | 208 |
| 4.1 Estratègia de clonatge | 208 |
| 4.2 Expressió | 212 |
| 4.3 Purificació | 212 |
| 4.4 Cristal·lització..... | 215 |
| 4.4.1 Cristalls amb àtoms pesants..... | 219 |
| 4.5 Recollida i processament de dades difracció..... | 220 |
| 4.5.1 Qualitat de les dades de difracció de les diferents formes cristal·lines | 220 |
| 4.5.1 Millor conjunt de dades natiu | 221 |
| 4.5.2 Cocristalls amb ssDNA | 223 |
| 4.5.3 Cristalls derivats amb àtoms pesant..... | 224 |
| 4.6 Resolució de l'estructura tridimensional | 227 |
| 4.6.1 Mètodes no experimentals: reemplaçament molecular | 227 |
| 4.6.1.1 Amb models atòmics | 227 |
| 4.6.1.2 Amb models de microscòpia electrònica | 230 |
| 4.6.1.2.1 Reconstruccions de criomicroscòpia electrònica | 230 |
| 4.6.1.2.2 Dades de difracció de molt baixa resolució..... | 233 |
| 4.6.2 Tècniques experimentals..... | 237 |
| 4.6.2.1 Resolució de l'estructura per SIRAS..... | 238 |
| 4.7 Descripció preliminar de l'estructura | 242 |
| 4.7.1 Estructura global de l'helicasa hexamèrica | 242 |
| 4.7.2 Estructura del monòmer | 243 |
| DISCUSSIÓ | 245 |
| 1. Producció de les proteïnes i obtenció de cristalls de qualitat | 245 |
| 2. Resolució de l'estructura i construcció d'un model per a l'hexàmer d'Aq_DnaB..... | 249 |
| 3. Estructura tridimensional i moviment | 251 |
| 3.1 Estructura dels glòbuls de l'hexàmer..... | 251 |

| | |
|-----------------------------|------------|
| 3.1.1 Capa N-terminal | 251 |
| 3.1.2 Capa C-terminal | 254 |
| 3.2 Centre actiu..... | 256 |
| 3.3 Unió a ssDNA..... | 256 |
| CONCLUSIONS | 257 |
| BIBLIOGRAFIA | 259 |

RESUM

Les helicases i translocases són una extensa superfamília molt diversa i ubiqua a tots els regnes de la vida, on realitzen funcions vitals per a la supervivència doncs formen part essencial dels complexos proteics que catalitzen el metabolisme dels àcids nucleics. La seva habilitat única per convertir l'energia química de la hidròlisi de nucleòtids trifosfat en treball mecànic de moviment direccional i separació de cadena les classifica com a proteïnes motores d'àcids nucleics.

La proteïna DnaB, objecte d'estudi d'aquest treball, és la principal helicasa replicativa bacteriana. Com a tal, la seva principal funció és la separació de les dues cadenes del DNA dúplex al front de la cadena retardada de la forqueta de replicació, en una reacció acoblada a la hidròlisi d'ATP. Igualment, DnaB juga un paper molt important en la progressió de l'etapa d'iniciació de la replicació cap a l'etapa d'elongació, doncs és el primer membre del replisoma a arribar a l'origen de replicació activat i participa activament en el reclutament de la resta de proteïnes que conformaran la maquinària replicativa cel·lular. Ja a l'etapa d'elongació, a part de l'activitat helicasa principal, DnaB coordina la seva activitat amb la de la polimerasa replicativa i la primasa, mitjançant interaccions directes amb ambdues proteïnes.

DnaB pertany a la superfamília 4 de les helicases, grup del qual n'és el representant prototípic. Com a membre del gran grup de les helicases hexamèriques, la seva forma funcional és un homohexàmer en forma d'anell, el qual té uns 320 kDa. Les sis butxaques d'unió a nucleòtid es situen a les interfícies entre monòmers de l'hexàmer, amb elements aportats per les dues subunitats veïnes. Això possibilita un cicle d'hidròlisi de l'ATP coordinat entre les sis subunitats. La separació de cadena es dona mitjançant un mecanisme d'exclusió estèrica en el qual la cadena retardada es transloca a través de la cavitat central de l'helicasa en el sentit de 5' a 3', mentre que la cadena líder queda exclosa. La translocació del ssDNA es dona gràcies a la coordinació dels canvis conformacionals dels llaços d'unió a DNA de cadascuna de les sis subunitats, els

quals es troben projectats cap al canal central. El mecanisme estructural exacte de translocació encara no ha estat descrit per a DnaB, però sí proposat per altres membres de la mateixa superfamília (proteïna gp4 del bacteriòfag de T7) i de la superfamília 3, també d'helicases hexamèriques (proteïna E1 del virus del papil·loma).

En aquest treball s'ha caracteritzat estructuralment, emprant les tècniques de cristal·lografia de raigs X, la proteïna DnaB procedent de la bactèria termòfila *Aquifex aeolicus*. Partint del DNA genòmic bacterià, el gen codificant per aquesta proteïna fou aïllat i la proteïna sobreexpressada heteròlogament a *Escherichia coli*, purificada i cristal·litzada, generant unes dades de difracció al voltant dels 3.5 Å de resolució. Concretament, es treballà amb una forma lleugerament més curta de la proteïna sencera, Aq_DnaB[1-439], a la qual li manquen els darrers 29 residus. Els intents de resoldre l'estructura tridimensional de la proteïna sencera, així com només del domini C-terminal d'aquesta, foren infructuosos degut a, respectivament, la baixa qualitat de les dades de difracció obtingudes o a l'absència de cristalls. També es tractà de determinar l'estructura de la DnaB de la bactèria termòfila *Thermotoga maritima*, però, tot i que la proteïna s'aïllà del genoma bacterià, s'expressà heteròlogament a *E. coli* i es purificà satisfactòriament, no s'aconseguí obtenir cristalls.

Aq_DnaB[1-439] cristal·litzà en el grup espacial C2, amb un hexàmer per unitat asimètrica. L'estructura es determinà mitjançant SIRAS i es generà un model atòmic per a l'hexàmer complet, que actualment es troba a 4.5 Å de resolució. Encara que l'etapa d'afinament de l'estructura encara no està acabada, en aquest treball es descriu la forma global de l'hexàmer d'Aq_DnaB i es discuteixen les diferències amb les altres proteïnes homòlogues a DnaB que han estat recentment resoltes també per cristal·lografia de raigs X. Aq_DnaB[1-439] forma un hexàmer amb dues capes de simetria clarament diferenciada, amb unes dimensions totals de 100 Å d'amplada i 80 Å d'alçada. La capa formada pels dominis N-terminals es troba en simetria C3, mentre que la capa formada pels dominis C-terminals presenta una simetria pròpia a C6. El diàmetre del canal interior central és d'uns 25 Å, tant al llarg de la capa N-terminal com C-terminal. Aquesta és la principal diferència de l'estructura presentada en aquest treball respecte a les estructures hexamèriques de DnaB resoltes anteriorment, doncs en aquelles el diàmetre del canal a través de la capa N-terminal era entre 20 i 25 Å més estret que en

l'hexàmer d'Aq_DnaB[1-439]. Aquest estretament bé marcat pel trencament d'una de les dues superfícies d'interacció entre monòmers N-terminal observades a les estructures prèvies, fet que provoca un augment de la flexibilitat del subdomini N-terminal que es trobava implicat en aquestes interaccions. Igualment, l'estretament implica que només el DNA monocadena pot passar a través del canal central de l'hexàmer, mentre que a les estructures descrites prèviament hi podia passar tant ssDNA com dsDNA. Per tant, possiblement l'estructura de DnaB determinada en aquest treball es troba més propera a la conformació funcional de la proteïna durant la realització de la seva activitat helicasa, mentre que les anteriors correspondrien a la forma inactiva o a la conformació de la proteïna capaç de translocar-se sobre DNA de doble cadena.

INTRODUCCIÓ

Per a una major claredat, la introducció està organitzada en tres grans blocs titulats, respectivament: helicases i translocases, la proteïna DnaB i funcions de DnaB a la cèl·lula bacteriana.

Al primer bloc es descriu l'extens i divers grup format per les helicases i les translocases, al qual pertany DnaB, la proteïna objecte d'estudi d'aquesta tesi. Es detalla la rellevància biològica de les reaccions que catalitzen, la classificació en superfamílies, característiques comunes a tots els grups i forma oligomèrica funcional. Igualment, es descriuen els mecanismes d'hidròlisi de NTP, translocació sobre àcids nucleics i separació de cadena proposats per a les diferents superfamílies, posant especial èmfasi en la superfamília SF4 de les helicases hexamèriques, de la qual DnaB és el membre prototípic.

El segon bloc tracta específicament de la proteïna DnaB, la primera helicasa hexamèrica caracteritzada i la més extensament estudiada, particularment a nivell de la descripció bioquímica de la seva activitat enzimàtica. Igualment, es detalla la seva organització en dominis, així com la seva forma funcional i estructura quaternària, la qual ha estat extensament estudiada al llarg del temps emprant diverses tècniques estructurals, incloent la microscòpia electrònica, la ressonància magnètica nuclear i la cristal·lografia de raigs X.

Finalment, al tercer bloc es llisten les activitats biològiques que DnaB realitza dins la cèl·lula bacteriana, on la seva principal funció és actuar d'helicasa replicativa desenrotllant el DNA de doble cadena al front de la forqueta de replicació. Es descriu la implicació de DnaB a les etapes d'iniciació i d'elongació de la replicació del genoma bacterià.

I. HELICASES I TRANSLOCASES

La primera helicasa es va descriure fa més de 30 anys al bacteri *Escherichia coli* (Abdel-Monem *et al.* 1976). Des de llavors, s'han identificat i caracteritzat centenars de proteïnes amb activitat helicasa i ha quedat palès que es tracta d'un grup molt divers de proteïnes ubiqües presents tant a procariotes, arquees, eucariotes com a virus. Aquesta ubiqüitat posa en rellevància el paper clau de les helicases per a la viabilitat cel·lular, doncs formen part essencial dels complexos proteics que catalitzen les reaccions del metabolisme dels àcids nucleics en aquests organismes (Lohman 1993; Matson *et al.* 1994; Lohman *et al.* 1996; Patel *et al.* 2000; Singleton *et al.* 2007). De fet, hi ha múltiples malalties hereditàries humanes, la majoria relacionades amb càncer, retard mental i envelliment prematur, que tenen el seu origen en mutacions a les diferents helicases humanes (Watt *et al.* 1996; van Brabant *et al.* 2000; Mohaghegh *et al.* 2001; Hickson 2003).

Les helicases es classifiquen com a proteïnes motores d'àcids nucleics degut a la seva habilitat única per convertir l'energia química de la hidròlisi de nucleòtids trifosfat (NTPs) en treball mecànic de moviment direccional i separació de les cadenes dels àcids nucleics. L'activitat de translocació unidireccional de les helicases condueix processos com la recombinació del DNA, la remodelació del nucleosoma, l'empaquetament dels genomes, la restricció del DNA i la dissociació de proteïnes unides als àcids nucleics; mentre que l'activitat de separació de parells de bases d'àcids nucleics és crítica per a la replicació, la transcripció i la reparació dels genomes (Donmez *et al.* 2008).

1. Classificació

1.1 En funció de la seqüència aminoacídica

Originàriament, les helicases varen ser classificades en cinc grans grups, tres superfamílies i dues famílies més petites, en base a l'anàlisi de la seva estructura primària: les superfamílies 1, 2 i 3 (SF1, SF2 i SF3, respectivament) i la família de les proteïnes relacionades amb DnaB i la de les proteïnes relacionades amb Rho. Aquesta classificació identificava motius de seqüència curts específics per a cada família, tant en similitud de seqüència com en nombre i organització d'aquests motius. SF1 i SF2, els dos grups més grans i estretament relacionats, tenien 7 motius definitoris. SF3 només 3 i els grups de DnaB i Rho 5 i 2, respectivament (Gorbalenya *et al.* 1993).

Encara que aquesta classificació s'ha mantingut fins a l'actualitat, presenta alguns problemes, doncs es va fer abans de caracteritzar tant bioquímicament com estructuralment la majoria d'aquests grups d'helicases. Inicialment, es creia que els diferents motius tenien una significació funcional i estaven directament implicats en les funcions pròpies de les helicases, com ara la unió a les estructures en forma de forqueta o el desenrotllament dels àcids nucleics. Posteriorment, es va comprovar que en realitat els diferents motius estan agrupats dins l'estructura tridimensional de la proteïna per formar una butxaca característica d'hidròlisi de NTP i transmissió de l'energia. Per tant, el concepte de que es tracta de motius signatura per helicases és massa restrictiu i amb el temps s'ha evidenciat que els motius característics originàriament d'helicases en realitat comprenen un grup més ampli de proteïnes ATPases que interactuen amb els àcids nucleics (Hall *et al.* 1999; Caruthers *et al.* 2002). Així doncs, moltes proteïnes classificades primerament com a helicases basant-se en estudis de comparació de seqüència en realitat no catalitzen la separació de les cadenes als àcids nucleics, sinó que només es transloquen unidireccionalment al llarg d'aquests. Per això, actualment s'ha proposat anomenar al grup proteic "grup d'helicases i translocases", alhora que s'han definit noves famílies i s'han subdividit les preexistents per adaptar-les als coneixements bioquímics i estructurals actuals (Singleton *et al.* 2007; Berger 2008).

Les translocases es mouen al llarg d'un substrat de DNA o RNA per desplaçar l'àcid nucleic a una nova localització o per alliberar-lo de les proteïnes que s'hi troben unides. Les helicases també es mouen al llarg dels àcids nucleics, però addicionalment separen els dúplexs en cadenes senzilles. Segons aquestes definicions, actualment es distingeixen tres famílies de translocases i sis famílies d'helicases, amb una possible setena família d'helicases putatives, doncs encara cal verificar experimentalment la seva habilitat per separar dúplexs. A la taula I.1 es mostra aquesta classificació, alhora que s'inclouen alguns exemples de membres i funcions de cadascuna de les superfamílies d'helicases/translocases.

| | Superfamília | Membres | Funcions |
|--------------|--------------|---|--|
| HELICASES | SF1 | PcrA, Rep, UvrD, RecBCD, Dda; Rm3, Pif1, Dna2, Srs2, Upf1 | desenrotllament i reparació del DNA; desplaçament proteïnes unides a DNA; degradació d'encebadors i mRNA aberrant |
| | SF2 | RecQ, RecG, UvrB; proteïnes DExD/H (eIF4A, Prp2, Ski2, VASA, Dpbs, NS3); proteïnes Snf2/SWI (Snf2, ISWI, Rad54, Hel308) | remodelació del RNA i la cromatina; desplaçament de proteïnes d'unió a RNA i DNA; desenrotllament de DNA i RNA; translocació DNA i RNA; biogènesi ribosomes; fusió i migració de <i>Holliday junctions</i> |
| | SF3 | E1 del virus del papil·loma, LTag de SV40, Rep40 del virus AAV | desenrotllament / replicació del DNA o RNA; reconeixement i fusió de l'origen de replicació |
| | SF4 | DnaB, RepA; gp4 del fag T7, G40P del fag SSP1; TWINKLE | desenrotllament / replicació del DNA |
| | SF5 | Rho | terminació de la transcripció; desenrotllament heterodúplexs RNA/DNA; translocació de RNA |
| | SF6 | Proteïnes MCM eucariòtiques i d'arquees | desenrotllament / replicació del DNA; fusió origen de replicació; remodelació cromatina |
| | SF7? | Tip48/49, reptin/pontin | remodelació cromatina; maduració snRNA; reparació DNA; assemblatge de la telomerasa |
| TRANSLOCASES | HerA/FtsK | FtsK, SpoIIIE, TrwB, TraD; terminases de caudovirus i herpesvirus, gp16 del podovirus φ29, P4 del fag φ12 | Partició del cromosoma / conjugació; compactació del ssRNA / motors d'empaquetament viral |
| | RuvB | RuvB | Recombinació homòloga: processament de <i>Holliday junctions</i> |
| | McrB | McrB | Enzims de restricció del tipus IV |

Taula I.1: Grups d'helicases i translocases, exemples i funcions. En blau es representen els membres bacterians, en lila els arquees, en marró els vírics i en verd els eucariòtics.

Les SF1 i SF2, relacionades estructural i filogenèticament, són les famílies d'helicases més nombroses i tenen representants a tots els dominis de la vida, on realitzen funcions molt diverses (Lohman *et al.* 2008). Són especialment versàtils els membres de la SF2, gràcies a que contenen múltiples dominis accessoris que els proporcionen noves funcionalitats, alhora que els permeten funcionar com a part de grans complexos multiproteics (Pyle 2008).

Les helicases de la SF3 pertanyen a virus petits de DNA i RNA, on són dominis de proteïnes més grans anomenades iniciadors o proteïnes d'unió a l'origen. Al fusionar l'activitat de reconeixement dels orígens de replicació amb l'activitat helicasa en un sol polipèptid, aquests virus són capaços de saltar-se els mecanismes de regulació dels seus hostes i replicar lliurement els seus genomes (Hickman *et al.* 2005; Duderstadt *et al.* 2008).

La SF4 originàriament incloïa només les helicases replicatives de bacteris i bacteriòfags, però per homologia s'ha afegit a aquest grup la recentment caracteritzada helicasa replicativa mitocondrial, anomenada TWINKLE (Spelbrink *et al.* 2001; Korhonen *et al.* 2003). DnaB, la proteïna objecte d'estudi d'aquest treball, és el membre arquetípic d'aquest grup.

La SF5 només té representants bacterians, homòlegs a la proteïna de terminació de la transcripció Rho.

La SF6, definida recentment (Singleton *et al.* 2007), comprèn les proteïnes del complex de manteniment del minicromosoma (MCM) eucariòtiques i d'arquees, les quals són les helicases replicatives a aquests dos dominis de la vida (Bochman *et al.* 2008; Costa *et al.* 2008).

Finalment, la SF7 d'helicases putatives inclou proteïnes eucariòtiques relacionades topològicament amb la translocasa RuvB (Berger 2008), les quals tenen papers claus en processos tant diversos com la remodelació de la cromatina i, tal i com s'ha de descrit recentment, l'assemblatge de l'holoenzim telomerasa (Gallant 2007; Venteicher *et al.* 2008).

1.2 Característiques comunes a helicases/translocases

Totes les helicases i translocases contenen un domini d'unió a NTP que pot presentar bé un plegament tipus RecA o un plegament tipus AAA+ (sigles de l'anglès *ATPases Associated with several cellular Activities*). Ambdós tipus de plegament són molt similars, doncs deriven d'un domini $\alpha\beta\alpha$ ancestral comú anomenat plegament ASCE (sigles de l'anglès *additional strand, conserved E*), el qual està inclòs en l'extensa superfamília de les NTPases tipus llaç P (*P-loop*). Una de les característiques definitòries de les proteïnes d'unió a nucleòtid tipus *P-loop* és la presència de dues seqüències signatura anomenades Walker A i Walker B, les quals són importants per la unió i la hidròlisi del nucleòtid. Efectivament, totes les famílies d'helicases i translocases contenen aquests dos motius de seqüència, els quals, al pertànyer a proteïnes tipus ASCE, es situen a l'àpex de tres fulles β paral·leles adjacents dins del plegament $\alpha\beta\alpha$ central. Addicionalment, també com a característica ASCE, presenten un residu d'arginina conservat, anomenat dit d'arginina, que juga un paper important en l'acoblament energètic i té una localització marcadament variable entre els diferents grups. Finalment, la diferència entre el plegament RecA i el AAA+ rau en la presència de cadenes β accessòries al *core* ASCE central en el cas del plegament RecA, les quals són absents en AAA+, que en canvi té un domini helicoïdal C-terminal característic anomenat tapa (figura I.1).

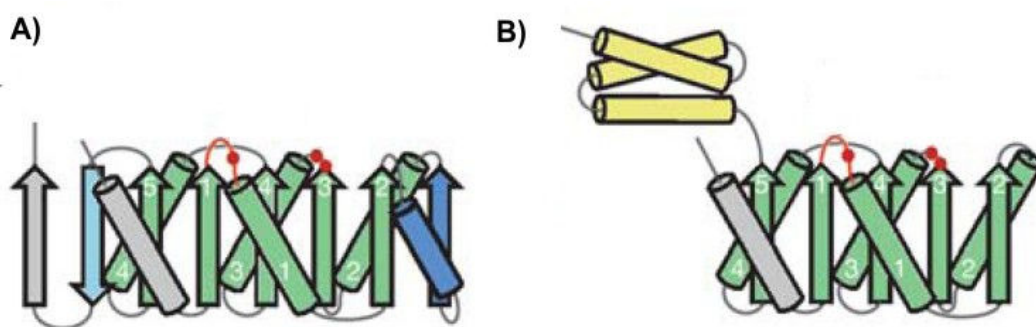


Figura I.1: Comparació dels plegaments tipus RecA i AAA+. Diagrames de topologia d'un plegament tipus RecA (A) i tipus AAA+ (B). El plegament ASCE ancestral es representa en verd; les insercions característiques del plegament RecA en blau; en groc el domini helicoïdal C-terminal característic del plegament AAA+. En vermell es mostra la localització dels motius d'unió a NTP Walker A (un punt vermell) i B (dos punts vermells). Figura adaptada d'Erzberger i Berger, 2006.

La butxaca nucleotídica de les helicases/translocases està invariablement formada per dos dominis tipus RecA o AAA+, que poden procedir de la mateixa cadena polipeptídica (plegament RecA en tàndem) o de diferents monòmers en el cas de les proteïnes que funcionen com a oligòmers. Es tracta, doncs, d'una butxaca catalítica bipartida on el residu d'arginina conservat sempre és proporcionat per un segon mòdul RecA/AAA+, diferent del que proporciona els motius Walker A i B. El dit d'arginina ajuda a la hidròlisi del nucleòtid i transmet l'estat d'unió a nucleòtid d'un mòdul RecA/AAA+ a l'altre, possibilitant la conversió de l'energia química en mecànica gràcies a l'acoblament de la unió i la hidròlisi de NTP a canvis conformacionals (Caruthers *et al.* 2002), alhora que és essencial per la cooperativitat entre els diferents monòmers a les helicases hexamèriques (Crampton *et al.* 2006). Aquests canvis conformacionals, induïts pel NTP, als dominis RecA/AAA+ possibiliten l'activitat de translocació gràcies al reposicionament dels llocs d'unió a àcids nucleics de la superfície de la proteïna (Hopfner *et al.* 2007).

La taula I.2, a l'apartat següent (1.3), mostra el tipus de plegament que presenta cada superfamília d'helicases/translocases. En el cas de les famílies pertanyents a AAA+, s'indica el subgrup o clade al qual pertany el plegament característic d'aquella família. Cal tenir en compte que la superfamília AAA+ comprèn un grup molt divers de proteïnes ATPasa que utilitzen l'energia de l'ATP per remodelar substrats macromoleculars tan diversos com ara complexes multiproteics, membranes cel·lulars, àcids nucleics o microtúbuls. Aquesta diversitat funcional ha provocat l'evolució característica i diferenciada dels diferents grups, els quals consegüentment presenten múltiples variacions del motiu AAA+ original. Aquestes variacions permeten definir fins a 7 clades diferents (Erzberger *et al.* 2006a).

1.3 Classificacions mecàniques

Tot i que la classificació en superfamílies sol ser prevalent, principalment per coherència amb la literatura preexistent, existeixen altres maneres de classificar les helicases/translocases: en funció del substrat sobre el qual actuen (DNA, RNA o heterodúplex RNA/DNA; així com si són àcids nucleics de cadena doble o senzilla), la

seva polaritat de translocació sobre l'àcid nucleic (5'-3' o 3'-5') i també la seva forma oligomèrica. A la taula I.2 es detalla a quin tipus pertany cada família.

| | Família | Plegament | Estat oligomèric | Substrat | Polaritat |
|---------------------|------------------|--------------------------------|--------------------------------------|---------------------------------|---------------------------------|
| HELICASES | SF1 | RecA doble en tàndem | Monòmer (dímer/multímer) | ssDNA i ssRNA | 3'-5' (SF-1A); 5'-3' (SF-1B) |
| | SF2 | RecA doble en tàndem | Monòmer (dímer/multímer) | ssDNA i ssRNA; dsDNA i dsRNA | 3'-5' (SF-2A); 5'-3' (SF-2B) |
| | SF3 | AAA+ (clade propi) | Hexàmer (dodecàmer?) | ssDNA i ssRNA | 3'-5' |
| | SF4 | RecA | Hexàmer (altres?) | ssDNA (dsDNA?) | 5'-3' |
| | SF5 | RecA | Hexàmer | heterodúplex RNA/DNA | 5'-3' |
| | SF6 | AAA+ (clade PS-II) | Homo / hetero hexàmer (altres?) | ssDNA (dsDNA?) | 3'-5' |
| | SF7? | AAA+ (nou clade?) | Doble hexàmer* | DNA | 5'-3' |
| TRANSLOCASES | HerA/FtsK | RecA | Hexàmer (pentàmer) | dsDNA; ssDNA; ssRNA | |
| | RuvB | AAA+ (clade HCLR) | Hexàmer (doble hexàmer* amb RuvA) | dsDNA | |
| | McrB | AAA+ (clade de l'insert H2) | Heptàmer | dsDNA | |

Taula I.2: Topologia de plegament, estat oligomèric i característiques mecàniques de les famílies d'helicases/translocases. (*) Cadascuna de les dues proteïnes que treballen plegades forma un hexàmer, de manera que el complex funcional pren la forma d'un doble anell hexamèric. Veure text per aclarir els aspectes en parèntesi o interrogant.

Cap superfamília conté exclusivament helicases de RNA, sinó que aquestes es troben repartides a SF1, SF2 i SF3. La majoria d'helicases de RNA pertanyen a SF2, on es troben agrupades a les extenses subfamílies DExD/H. SF1 conté un sol grup petit d'helicases de RNA, les proteïnes eucariòtiques Upf1. Finalment, a SF3 hi ha helicases de RNA d'origen viral (Kadare *et al.* 1997; Jankowsky *et al.* 2007). El substrat de SF5 és l'heterodúplex RNA/DNA que es forma durant la transcripció. La resta de superfamílies comprenen helicases de DNA.

Respecte a la capacitat de translocació sobre àcids nucleics de doble cadena, SF2 conté varies subfamílies d'helicases poc processives, per exemple la subfamília Snf2/SWI, les quals funcionen majoritàriament com a translocases d'àcids nucleics de doble cadena (Mackintosh *et al.* 2006; Singleton *et al.* 2007). Pel que fa a les superfamílies 4 i 6, tot i que la seva tasca principal és exercir d'helicases replicatives, s'ha demostrat que alguns dels seus membres, com ara la gp4 del fag T7, la DnaB bacteriana i l'heterohexàmer Mcm4,6,7 eucariòtic, poden translocar-se *in vitro* sobre dsDNA, encara que no està clara la rellevància d'aquest fenomen *in vivo* (Kaplan *et al.* 2002; Kaplan *et al.* 2003).

És particularment útil classificar les helicases en funció de la seva estructura quaternària, doncs això influencia la seva manera d'interaccionar amb els àcids nucleics. Així podem diferenciar les helicases monomèriques (o no hexamèriques), que comprenen les superfamílies SF1 i SF2, i les helicases hexamèriques, que comprenen la resta de famílies. L'helicasa objecte d'estudi d'aquesta tesi pertany a les helicases hexamèriques, concretament a la família SF4, que precisament es va definir en relació a aquesta proteïna. Per tant, tot i que es descriuran també breument les helicases monomèriques, aquesta introducció es centrarà en la descripció de les característiques bioquímiques, funcionals i estructurals de les helicases hexamèriques i, sempre i quan sigui possible, centrant-nos en la família 4, corresponent a les helicases semblants a DnaB.

2. Helicases no hexamèriques: SF1 i SF2

2.1 Forma funcional

L'estat oligomèric responsable de la forma funcional dels membres de SF1 i SF2 ha originat força controvèrsia i continua sent objecte d'estudi. Actualment, es considera que, encara que totes les activitats dels enzims SF1 i SF2 (helicasa, translocasa, ATPasa i desplaçament de proteïnes unides als àcids nucleics) poden ser atribuïdes a la forma monomèrica, a la majoria dels casos aquestes activitats estan influenciades per l'estat d'assemblatge de l'enzim (Lohman *et al.* 2008).

Estudis *in vitro* han demostrat que els enzims de SF1 i SF2 en estat monomèric són helicases pobres. De fet, les formes monomèriques d'alguns membres de SF1 (per exemple Rep, UvrD i PrcA) són incapaçs de funcionar com helicases *in vitro*. En canvi, aquests monòmers de SF1 i SF2 són translocases ràpides i processives d'àcids nucleics de cadena senzilla *in vitro*. En tots els casos, l'autoassemblatge de múltiples monòmers sobre l'àcid nucleic de cadena senzilla que flanqueja la zona dúplex que encara s'ha de separar activa i/o estimula específicament l'activitat helicasa d'aquestes proteïnes (Maluf *et al.* 2003; Byrd *et al.* 2005; Tackett *et al.* 2005; Patel *et al.* 2006). Per tant, aquests enzims tenen una activitat helicasa latent que requereix activació mitjançant la cooperació de varis monòmers.

Les activitats helicasa i translocasa a les superfamílies SF1 i SF2 són, doncs, funcions separables conduïdes per diferents estat oligomèrics de la proteïna. De fet, es considera que l'oligomerització de l'enzim, juntament amb possibles interaccions amb proteïnes accessòries, és vital en la regulació específica de la translocació enfront de l'activitat helicasa per tal d'evitar danys al DNA produïts per activitats helicases descontrolades.

2.2 Mecanismes de translocació del ssDNA

2.2.1 Models pas a pas (*stepping*)

Els mecanismes pas a pas, que tradicionalment s'han anomenat mecanismes “*inchworm*” (en anglès, eruga geomètrica) per les helicases monomèriques, necessiten l'existència d'almenys dos llocs d'unió a àcid nucleic independents dins la unitat que es transloca, els quals uneixen i alliberen l'àcid nucleic en resposta a l'estat d'unió a NTP de l'enzim. L'helicasa sempre es manté unida a l'àcid nucleic mitjançant un dels seus llocs d'unió (Patel *et al.* 2006).

En el model pas a pas, la translocació es dona a través de moviments successius entre els dos llocs d'unió a àcid nucleic a mesura que el NTP és unit i hidrolitzat. Aquest dos llocs d'unió es troben a la superfície dels dos dominis RecA, un a cada domini, que conformen el lloc d'unió a NTP a la seva interfície. Inicialment, el primer lloc interacciona fortament amb l'àcid nucleic i el segon només dèbilment. Subsegüentment, el lloc unit feblement es dissocia de l'àcid nucleic, es mou cap endavant i estableix noves interaccions amb bases riu amunt de la cadena senzilla. Una vegada aquest lloc inicialment feble s'ha desplaçat i s'ha unit fortament a una regió riu amunt de l'àcid nucleic, el lloc d'unió originalment fort esdevé feble, es dissocia de l'àcid nucleic i es mou endavant, cap al nou lloc fort (Yarranton *et al.* 1979). Els canvis en l'afinitat per l'àcid nucleic als dos llocs d'unió són conduïts per la unió del NTP al solc d'unió entre dominis, cosa que provoca que el solc es tanqui i es reposicionin els llocs d'unió a àcid nucleic (Velankar *et al.* 1999). La figura I.2 mostra una representació esquemàtica d'aquest model pas a pas.

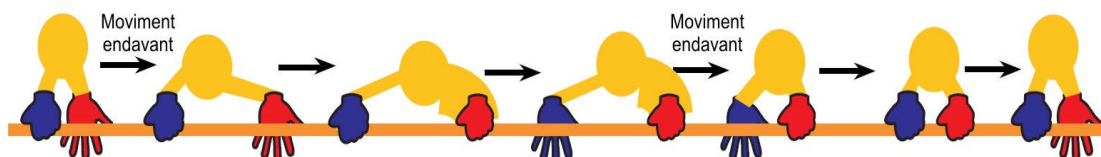


Figura I.2: Esquema d'un mecanisme de translocació pas a pas. Un monòmer d'helicasa amb un lloc d'unió fort (mà tancada) i un lloc d'unió feble (mà oberta) a l'àcid nucleic de cadena senzilla, pateix etapes de moviment cap endavant i canvis en l'afinitat per a l'àcid nucleic (transicions de l'estat fort o feble) a mesura que es mou al llarg de la cadena senzilla. Aquest canvis estan associats a l'estat d'unió a nucleòtid de l'enzim.

Originalment, aquest mecanisme pas a pas va ser proposat en base als resultats estructurals per helicases SF1 i SF2 monomèriques (Kim *et al.* 1998; Velankar *et al.* 1999; Soutanas *et al.* 2000b), en contraposició al mecanisme rodant (*rolling*) proposat per helicases SF1 i SF2 que actuessin com a dímers o oligòmers majors (Wong *et al.* 1992). En el model rodant cadascuna de les dues subunitats d'helicasa alternen la seva unió a àcids nucleics de cadena senzilla i dúplex a mesura que canvien el seu estat d'unió a NTP, doncs una mateixa unitat no pot unir simultàniament cadena senzilla i doble. A diferència del model *inchworm*, on les subunitats mantenen la seva posició relativa al llarg del DNA, les subunitats en el model rodant s'alternen per ser la subunitat líder. Aquest mecanisme es descartà degut a les estructures de monòmers de SF1 i SF2 units simultàniament a ssDNA i dsDNA. Actualment, en concordança amb estudis bioquímics, estructurals i mecànics de molècula sola (*single molecule*), es proposa un mecanisme pas a pas dimèric (*dimeric inchworm*) per a la translocació i el desenrotllament del dsDNA per part dels dímers de UvrD i Rep *in vitro* (Ha *et al.* 2002; Maluf *et al.* 2003). En aquest cas, cada subunitat d'helicasa contindria un lloc d'unió a àcid nucleic sota el control de la butxaca nucleotídica de la pròpia unitat, però que estaria coordinada amb l'activitat de la butxaca de NTP de la segona subunitat.

2.2.2 Model Brownià

Aquest model només requereix un lloc d'unió a àcids nucleics a l'enzim, però implica l'existència de dos estats conformacionals de la translocasa: un estat d'alta afinitat per a l'àcid nucleic de cadena senzilla i un altre d'afinitat dèbil. Ambdós conformacions estan modulades per l'estat d'unió a NTP. La unió del NTP afebleix la interacció de la proteïna amb l'àcid nucleic, cosa que permet un període breu de moviment Brownià abans que la hidròlisi del NTP provoqui una reunió a l'àcid nucleic esbiaixada vers a un moviment cap endavant (Levin *et al.* 2005). És el model actualment proposat per a la translocació i el desenrotllament per part del domini NS3 del virus HCV.

A la figura I.3 es representa esquemàticament un moviment de translocació de tipus brownià.



Figura I.3: Esquema d'un mecanisme de translocació de motor brownià. L'helicasa pateix canvis en la seva afinitat per als àcids nucleics de cadena senzilla. En l'estat d'unió forta amb l'àcid nucleic de cadena senzilla (esquerra), l'helicasa no pot moure's. En l'estat d'unió feble (centre), l'helicasa es mou en ambdós sentits. La curta durada de l'estat d'unió feble, degut al ràpid canvi en l'estat d'unió a NTP, manté l'helicasa propera al punt de partida. Quan torna a l'estat d'unió forta (dreta) algunes molècules d'helicasa s'han desplaçat cap endavant i algunes, que s'havien desplaçat cap enrere, retornen a la seva posició original. La repetició d'aquest passos comporta un desplaçament net cap endavant.

2.3 Mecanismes de desenrotllament d'àcids nucleics de cadena doble

Les helicases poden desenrotllar els àcids nucleics dúplexs gràcies a que acoblen la seva activitat de translocació sobre els àcids nucleics de cadena senzilla a l'activitat de separació de bases. La translocació al llarg de l'àcid nucleic es pot donar per qualsevol dels mecanismes descrits a l'apartat anterior. A continuació es descriuen els mecanismes propis de l'activitat de desenrotllament.

2.3.1 Diferenciació mecanismes actius i passius

Els mecanismes de desenrotllament del DNA es poden classificar en actius o passius, en funció de si l'helicasa participa en el procés de desenrotllament o només estableix l'àcid nucleic de cadena senzilla resultant (Lohman 1993; Lohman *et al.* 1996; von Hippel *et al.* 2001). Aquesta classificació és aplicable a tots els grups d'helicases, tant les no hexamèriques com les hexamèriques.

En un mecanisme passiu, l'helicasa no interacciona amb la zona de doble cadena, sinó que utilitza la seva activitat de translocació direccional al llarg dels àcids nucleics de cadena senzilla per moure's cap a la regió monocadena adjacent a la zona aparellada i estabilitzar-la. La cadena senzilla s'origina gràcies a les fluctuacions

tèrmiques que provoquen l'obertura transitòria i el posterior tancament dels parells de bases a la zona d'unió entre la regió aparellada i la desaparellada de l'àcid nucleic. Aquest fenomen s'anomena *breathing* (respiració en anglès) i es dona espontàniament en una escala temporal de microsegons (Nonin *et al.* 1995). Com la probabilitat de fusió simultània de varis parells de bases és molt baixa, es considera que els mecanismes passius només permeten desenrotllar un o pocs parells de bases a cada moviment de l'helicasa.

En un mecanisme actiu, l'helicasa interacciona directament amb la regió de doble cadena i facilita la desestabilització del dúplex. Per tant, un model d'aquest tipus implicaria l'existència de llocs d'unió a àcids nucleics de cadena senzilla i doble dins de l'helicasa. Alternativament, en comptes d'interaccionar amb el dúplex, l'helicasa podria interaccionar simultàniament amb les dues cadenes senzilles desenrotllades i, mitjançant torsió, desestabilitzar el dúplex adjacent. Es considera que els mecanismes actius serien propis d'helicases ràpides, que es mouen a una taxa superior a la d'obertura transitòria de les regions en forqueta (Betterton *et al.* 2005).

Cal tenir en compte que la classificació en actiu o passiu distingeix entre dos casos extrems que no reflecteixen el continu de comportaments que probablement es donen a la natura. Així doncs, poden existir diferents graus de desenrotllament actiu dependent del grau en que l'helicasa desplaça l'equilibri cap a l'obertura de la regió aparellada, cosa que al seu torn depèn del mecanisme d'acoblament de l'activitat NTPasa usat per l'helicasa i la natura de la seva interacció amb l'estructura en forqueta.

2.3.2 Models per a SF1 i SF2

Els mecanismes mitjançant els quals les helicases de SF1 i SF2 desenrotllen processivament els àcids nucleics són menys coneguts que els mecanismes de translocació. El fet que la translocació sigui necessària però no suficient per una activitat helicasa processiva, suggereix que aquestes superfamílies empenen mecanismes de desenrotllament actiu. Això és consistent amb les evidències experimentals de que Rep, UvrD, PcrA i NS3 utilitzen mecanismes actius per al desenrotllament dels dúplex (Wong *et al.* 1992; Soultanas *et al.* 2000a; Fischer *et al.* 2004; Cheng *et al.* 2007).

Inicialment, es va proposar un model actiu de desenrotllament tipus *inchworm* basant-se en les estructures cristal·lines de monòmers d'helicases SF1 (PcrA i UvrD) i SF2 (Hel308) unides a substrats en forqueta, les quals mostraven dos subdominis del monòmer (2B i 1B) contactant la regió de doble cadena (Velankar *et al.* 1999; Soultanas *et al.* 2000a; Lee *et al.* 2006; Buttner *et al.* 2007). En aquest model, la unió de l'ATP i el tancament concomitant del solc entre els subdominis d'unió a nucleòtid (1A i 2A) està acoblat a moviments dels subdominis 2B i 1B, els quals interaccionen amb el dúplex i el desestabilitzen. El model assumeix que les formes monomèriques dels enzims funcionen com a helicases actives. Ara bé, l'observació que el subdomini 2B és autoinhibitori per a l'activitat helicasa del monòmer *in vitro*, i que aquesta requereix l'activació per dimerització o interaccions amb proteïnes accessòries, suggereix que les estructures cristal·lines dels monòmers de SF1 i SF2 podrien ser complexes inactius que requereixen l'activació mitjançant dimerització i/o interaccions amb proteïnes auxiliars. Conseqüentment, actualment aquest model està en dubte i s'ha substituït pel model *inchworm* dimèric (Lohman *et al.* 2008).

En un mecanisme *inchworm* dimèric, proposat originàriament per a UvrD, la subunitat líder funciona com a helicasa, interaccionant directament amb la zona dúplex, mentre que la subunitat del darrera funciona com a translocasa de cadena senzilla (Maluf *et al.* 2003).

Finalment, també s'han proposat models pas a pas de cooperació funcional (*cooperative inchworm*) per explicar el reforçament de l'activitat helicasa gràcies a la multimerització de varis monòmers de SF1 o SF2 sobre el substrat a desenrotllar. Aquestes helicases no formen oligòmers estables i no mostren cooperativitat en la hidròlisi del NTP ni en la unió a l'àcid nucleic. En aquests models, la presència de múltiples helicases bé inhibiria moviments cap enrere no productius dels enzims o bé, si l'helicasa líder es dissocia de l'àcid nucleic, podria ser reemplaçada per l'helicasa del darrera (Byrd *et al.* 2005; Tackett *et al.* 2005). De totes maneres, els determinants d'aquesta cooperativitat no estan clars, doncs la mutació de la superfície de contacte entre monòmers no implica una davallada important de l'activitat helicasa *in vitro*, tot i que sí s'observa disminució de viabilitat *in vivo* (Mackintosh *et al.* 2006).

3. Helicases hexamèriques: superfamílies SF3, SF4, SF5, SF6 i SF7.

3.1 Forma funcional

Les helicases pertanyents a les superfamílies SF3, SF4, SF5, SF6 i SF7 actuen com a homohexàmers en forma d'anell, tal i com s'ha demostrat mitjançant estudis bioquímics, de microscòpia electrònica i cristal·logràfics dels membres de les diferents superfamílies (Patel *et al.* 2000; Matias *et al.* 2006; Gallant 2007; Singleton *et al.* 2007; Costa *et al.* 2008). De fet, la forma monomèrica d'aquestes helicases no és activa com NTPasa i tampoc poc catalitzar el desenrotllament dels àcids nucleics de doble cadena, per tant, l'hexamerització és essencial per a la funció. Aquesta estructura en forma d'anell està estabilitzada per la unió a NTP i/o a ions metàl·lics, així com pel substrat d'àcid nucleic (Patel *et al.* 2000).

A la figura I.4 es mostra l'estructura quaternària d'un representat de cadascuna de les cinc superfamílies d'helicases hexamèriques.

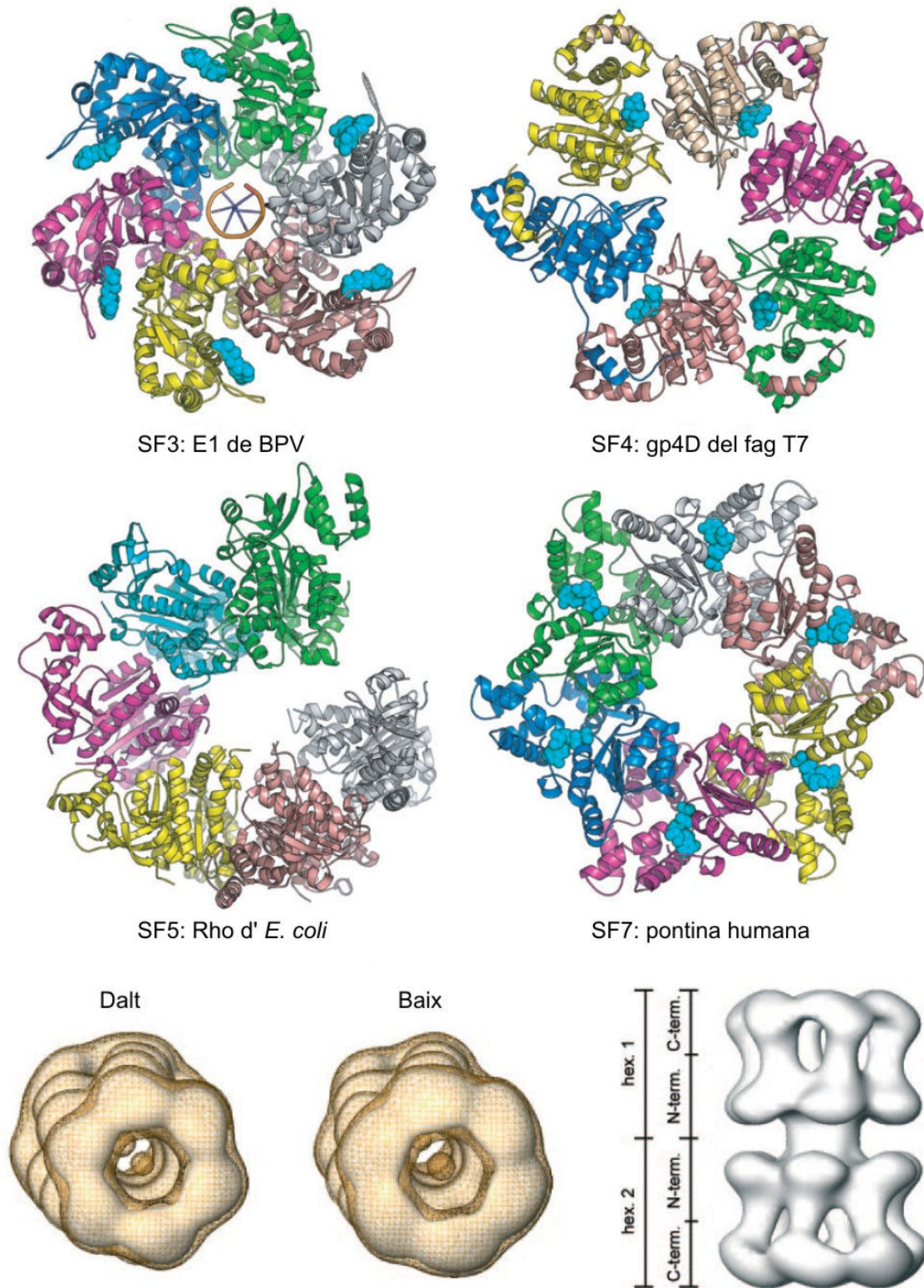


Figura I.4: Organització quaternària de les famílies d'helicases hexamèriques. De SF3 es representa la E1 del BPV en complex amb ssDNA (codi PDB: 2GXA). De SF4, la helicasa del fag T7 (codi PDB: 1EOJ). De SF5, la Rho d'*E. coli* (codi PDB: 1PV4). Per a SF6 no hi ha cap estructura cristal·lina i es mostra la reconstrucció de microscòpia electrònica del doble hexàmer de MCM de l'arquea *Methanobacterium thermoautotrophicum* (Gomez-Llorente *et al.* 2005). Per a SF7, es mostra l'hexàmer cristal·logràfic de la pontina humana (codi PDB: 2C9O).

Les estructures en forma d'anell són comunes als enzims del metabolisme dels àcids nucleics (Hingorani *et al.* 1998). Per exemple, la polimerasa replicativa assoleix l'alta processivitat observada *in vivo* gràcies a trobar-se associada al seu substrat de DNA físicament mitjançant la pinça lliscant, la qual té forma d'anell tant a virus, bacteris com eucariotes (Johnson *et al.* 2005). Igualment, totes les helicases que treballen desenrotllant el DNA al front de la forqueta de replicació presenten una estructura en forma d'anell. Això els permet romandre unides al substrat durant més temps i poder desenrotllar centenars de kilobases del genoma abans de dissociar-se (Pomerantz *et al.* 2007). Altres processos que també requereixen una alta processivitat, com ara l'empaquetament del genoma viral, també estan catalitzats per proteïnes en forma d'anell, per exemple la translocasa hexamèrica P4 del fag $\phi 12$ (Kainov *et al.* 2006). Per tant, l'evolució de les helicases hexamèriques cap a un oligòmer en forma d'anell els confereix la capacitat d'augmentar de manera molt eficient la seva processivitat, característica clau en els processos que catalitzen, per exemple en la replicació, la reparació i la recombinació dels genomes. Addicionalment, l'arranjament en anell permet l'acoblament dels cicles NTPasa entre totes les subunitats de l'hexàmer, augmentant així l'eficiència d'aquests cicles catalítics per promoure la translocació de l'helicasa.

Ara bé, al tractar-se d'anells tancats això suposa un problema alhora d'interaccionar amb el seu substrat, de manera que es necessita un procés de càrrega al motlle d'àcid nucleic. Les helicases hexamèriques poden emprar proteïnes accessòries que el permetin carregar-se al DNA/RNA o bé autoassemblar-se al voltant de l'àcid nucleic sense emprar cofactors addicionals (Davey *et al.* 2003). Per exemple, DnaB de SF4 empra la proteïna DnaC per carregar-se a l'origen de replicació (veure apartat 1.3.1 de la introducció), mentre que Rho de SF5 s'autoassembla mitjançant un mecanisme d'obertura d'anell (Skordalakes *et al.* 2003; Skordalakes *et al.* 2006).

3.1.2 Variacions respecte a un homohexàmer

Tot i que aquestes cinc superfamílies d'helicases invariablement oligomeritzen formant un anell, existeixen algunes excepcions i/o variacions a l'associació en homohexàmers a les superfamílies SF3, SF4 i SF6, tal i com es descriu a continuació.

Els membres de SF3 formen dobles hexàmers sobre els orígens de replicació (Fouts *et al.* 1999; Valle *et al.* 2000), estat que assoleixen mitjançant la transició seqüencial de monòmers/dímers a dobles trímers i finalment a dobles hexàmers que es dona a mesura que passen de l'activitat inicial d'unió i fusió de l'origen a l'activitat d'helicasa replicativa (Schuck *et al.* 2005). Es creu que la forma de doble hexàmer en disposició cap a cap, amb els dos hexàmers interaccionant a través dels seus dominis d'unió a l'origen, és la que funciona durant la replicació del genoma víric (Hickman *et al.* 2005). Ara bé, les dades estructurals recents no corroboren aquesta disposició, doncs la polaritat de la cadena de ssDNA unida al canal central de l'helicasa E1 del virus del papil·loma boví (BPV) és incompatible amb aquest arregament ja que implica que el ssDNA entra a l'hexàmer per la zona corresponent al domini d'unió a l'origen, que conseqüentment estaria disposat proper a la zona del dúplex. Per tant, aquestes dades suggereixen que, un cop assemblats a l'origen, els dos hexàmers es mourien passant un al costat de l'altre i quedarien associats a través dels seus dominis ATPasa, deixant els dominis d'unió a DNA a la part exterior del dodecàmer, o bé funcionarien separatament (Enemark *et al.* 2006).

A SF4, l'helicasa replicativa del fag T7, a part de la seva forma hexamèrica predominant, també pot presentar-se com a anell heptamèric. La rellevància biològica de l'heptàmer no està clara i s'ha suggerit que podria tractar-se de la forma de l'helicasa capaç de translocar-se al llarg de dsDNA o tenir algun paper durant l'assemblatge de l'anell al voltant del DNA (Toth *et al.* 2003).

A SF6, el complex MCM eucariòtic conté 6 proteïnes homòlogues que heterooligomeritzen per formar un hexàmer. Encara que no hi ha dades estructurals directes, es creu que aquest pren la forma d'un anell. A arquees, el MCM conté un sola proteïna que, segons els estudis de microscòpia electrònica, presenta un gran polimorfisme estructural, doncs pot associar-se en hexàmers, anells dobles tant hexamèrics com heptamèrics i fins i tot en filaments helicoïdals. Malgrat aquesta diversitat, es creu que la forma rellevant biològicament és l'hexàmer o el doble hexàmer (Costa *et al.* 2008).

3.2 Mecanismes d'hidròlisi d'ATP

Una característica comuna a les helicases hexamèriques és que posseeixen sis llocs d'unió a NTP amb una forta cooperativitat negativa entre ells (Patel *et al.* 2000). L'excepció és l'antigen T gran de SV40, que mostra poca cooperativitat i és capaç d'unir 6 molècules de nucleòtid amb afinitat similar (Huang *et al.* 1998). També és comú a totes les helicases hexamèriques que el lloc d'unió a nucleòtid es situï a la interfície entre dos monòmers, concretament entre dos dominis amb plegament RecA/AAA+ procedents de subunitats adjacents. Això implica que la unió del NTP i la seva subsegüent hidròlisi està influenciada per la orientació relativa entre els protòmers que conformen l'hexàmer. Conseqüentment, els models d'unió i hidròlisi de NTP proposats per a les helicases hexamèriques van des de mecanismes concertats, on les sis subunitats actuen simultàniament i mantenen la simetria al llarg de tot el procés enzimàtic, a mecanismes asimètrics seqüencials, en els quals els protòmers adjacents es troben en estats diferents.

A continuació es presenten els models d'hidròlisi de NTP considerats actualment per les helicases hexamèriques. Cal tenir en compte que més d'un d'ells pot ser correcte, depenent de la família, i també que una mateixa proteïna pot presentar mecanismes d'hidròlisi de nucleòtid diferents en presència o absència del seu substrat d'àcid nucleic. A la figura I.5 es representen esquemàticament els tres mecanismes explicats al text.

3.2.1 Model dels tres llocs seqüencials (basat en l'ATPasa F1)

Aquest model va ser proposat originàriament degut a la percepció de similituds entre les helicases SF4 i SF5 i l'ATPasa F1 (Hingorani *et al.* 1997; Stitt *et al.* 1998). Postula l'existència de tres llocs potencials d'hidròlisi de NTP a l'hexàmer situats a les interfícies entre subunitats alternades. Aquests llocs poden trobar-se en un dels tres estats interconvertibles següents: unit a NTP (estat T), unit a NDP (estat D) o buit (estat E). A cada moment, cadascun dels tres llocs d'unió a nucleòtid es troba en un estat diferent. Quan un NTP es hidrolitzat al lloc T, el lloc D allibera NDP i fosfat inorgànic, mentre que el lloc E uneix un NTP entrant. Per tant, es produeixen canvis d'estat simultanis a tots tres llocs d'unió a nucleòtid. Les reaccions a cada lloc depenen de les

reaccions que es donen als llocs adjacents, fet que assegura la catàlisi seqüencial als tres llocs. Aquest cicle es repeteix de manera que tots tres llocs catalítics passen pels tres estats d'unió a nucleòtid, retornant a l'estat inicial al cap de tres cicles.

Els primers estudis bioquímics semblaven donar suport a aquest model de tres llocs seqüencials per les helicases hexamèriques de SF4 i SF5 (Patel *et al.* 2000). Per exemple, els estudis en l'helicasa DnaB suggerien la presència de tres llocs d'alta afinitat a NTP per hexàmer, amb els llocs restants de més baixa afinitat (Biswas *et al.* 1986; Bujalowski *et al.* 1993). Ara bé, aquest model presenta el problema de que, a diferència de l'ATPasa F1 que té cadenes diferenciades catalítiques (α) i no catalítiques (β), les helicases SF4/SF5 només tenen un tipus de subunitat. Conseqüentment, tots sis llocs són potencialment competents per l'intercanvi de NTP i no està clar com s'establirien els llocs no catalítics alternants. De fet, a la llum de les noves dades mecanístiques determinades amb tècniques més precises, aquest model ja no es considerat ni per les helicases gp4 de T7 ni per Rho, per les quals va ser proposat, i ha estat substituït per un model de “tots els llocs seqüencialment” (Liao *et al.* 2005; Adelman *et al.* 2006).

3.2.2 Model de tots els llocs seqüencialment

En aquest model, totes les subunitats de l'hexàmer són catalítiques i cicleu a través dels estats T, D i E. La hidròlisi de NTP està coordinada entre subunitats adjacents, amb tres subunitats contigües en els estats T, D i E en un determinat moment (Singleton *et al.* 2000). Una variació d'aquest model proposa l'existència de blocs d'estats d'unió a nucleòtid, amb múltiples subunitats adjacents en el mateix estat (Enemark *et al.* 2006). Comú a aquests dos models és la noció de que totes les subunitats de la proteïna són catalítiques i contacten l'àcid nucleic en algun moment durant el cicle de translocació. Això implica que el sistema hauria de ser extremadament sensible a un “enverinament” per subunitats mutants defectives bé en la hidròlisi de NTP o en la unió a DNA. Treballs recents amb la gp4 de T7 han demostrat que és així, doncs l'activitat NTPasa dependent de ssDNA és gairebé abolida completament per una petita quantitat de subunitats no catalítiques. A més, aquest treball també demostra que,

en absència de DNA, el sistema esdevé no acoblat i els monòmers són capaços d'hidrolitzar NTP de manera independent l'un de l'altre (Crampton *et al.* 2006).

Evidències estructurals donen suport a aquest model per les helicases de SF3 (Enemark *et al.* 2006) i les de SF4 i SF5 (Singleton *et al.* 2000). Tal i com s'ha comentat anteriorment, dades bioquímiques recents combinades amb mètodes computacionals també donen suport a aquest model per a gp4 i Rho (Liao *et al.* 2005; Adelman *et al.* 2006).

3.2.3 Model concertat

Aquest model, també anomenat de tancament d'iris, proposa que totes sis subunitats es troben en el mateix estat i que s'intercanvien entre els estats T, D i E simultàniament (Gai *et al.* 2004).

El model ha estat suggerit en base als estudis estructurals de l'antigen T gran de SV40 de la família SF3 i també podria ser aplicable a les translocases amb plegament tipus AAA⁺ (Singleton *et al.* 2007). Hi ha algunes evidències de que aquestes proteïnes mostren característiques d'unió a nucleòtid que difereixen de les dels enzims de SF4 i SF5 (Huang *et al.* 1998), les quals serien consistents amb un mecanisme concertat, però difícils de reconciliar amb les estructures dels complexos de les helicases gp4 (SF4) o E1 amb ssDNA (SF3) (Singleton *et al.* 2007). Alternativament, s'ha proposat que les helicases de SF3 emprarien específicament un mecanisme d'hidròlisi de NTP concertat durant la fusió de l'origen i que després canviarien a un mecanisme seqüencial per al desenrotllament replicatiu (Hopfner *et al.* 2007).

A la figura I.5 es representen esquemàticament els tres mecanismes d'hidròlisi de NTP que s'han explicat, així com una variació del mecanisme de tots els llocs seqüencialment.

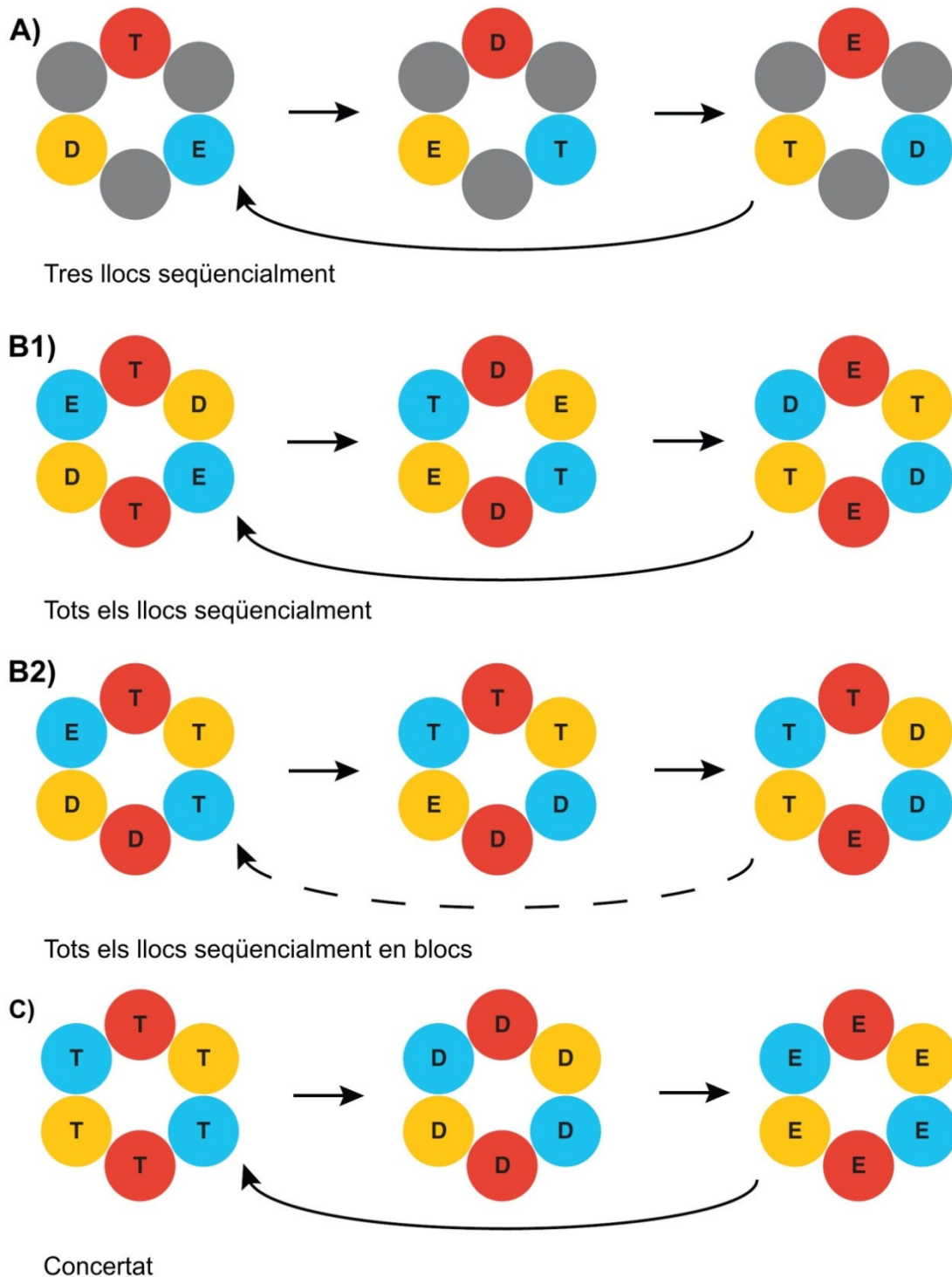


Figura I.5: Models d'hidròlisi i alliberament de NTP per a les helicases hexamèriques. Es representen els tres mecanismes explicats al text. Els colors indiquen diferents subunitats de l'hexàmer, no diferents estats nucleotídics. Les subunitats no catalítiques del model (A) són grises. Els diferents estats d'unió a nucleòtid estan etiquetats T (unit a NTP), D (unit a NDP) i E (buit). S'inclouen dues variants per al model de tots els llocs seqüencialment: (B1) una versió amb simetria dos proposada per a gp4 no unida a DNA, i (B2) una versió en blocs basada en l'estructura de E1 unida a ssDNA. En els models que només tenen tres estats dintre del cicle d'hidròlisi de NTP, una fletxa sòlida uneix l'estat final amb l'inicial. La fletxa discontinua del model (B2) indica que hi ha més de tres passos.

3.3 Acoblament de l'activitat NTPasa a la translocació

A tots els models anteriors, la unió a nucleòtid i la seva subsegüent hidròlisi estan acoblades a la translocació de la proteïna al llarg de l'àcid nucleic.

Els models de translocació al llarg dels àcids nucleics de cadena senzilla han estat inferits de les progressives dades estructurals obtingudes per les helicases hexamèriques. Una característica comuna a totes les estructures quaternàries d'aquestes helicases és que els llaços d'unió als àcids nucleics es projecten cap al canal central de l'anell. Addicionalment, les estructures de les helicases gp4D i E1 unides a nucleòtid suggereixen que aquests llaços rastregen l'espiral de l'esquelet de DNA i que els seus moviments estan conduïts per la unió a nucleòtid i la seva subsegüent hidròlisi. Cal destacar que a tots els estudis estructurals, les posicions dels llaços d'unió al DNA no presenten un moviment semblant a una frontissa respecte a la resta del domini, sinó que els dominis RecA/AAA+ sencers es mouen uns respecte dels altres i dels dominis d'oligomerització. Ara bé, la naturalesa exacte dels canvis mecànics que condueixen el moviment dels llaços i modulen l'afinitat de la unió al DNA no han estat determinats encara, alhora que es probable que presentin lleugeres variacions a les diferents famílies d'helicases hexamèriques. A continuació es presenten els models de translocació més acceptats actualment.

3.3.1 Model de transferència entre subunitats

La primera estructura cristal·lina d'un anell va ser del domini helicasa de la proteïna del gen 4 (gp4D) del bacteriòfag T7, la qual pertany a SF4 (Singleton *et al.* 2000). Cristal·litzà en forma d'hexàmer amb nucleòtids units a quatre de les sis interfícies entre subunitats i els llaços d'unió a DNA projectats cap a l'interior de l'anell. L'hexàmer mostra desviacions respecte a una simetria sis, cosa que porta com a conseqüència que les interfícies entre subunitats, els llaços d'unió a DNA i els dits d'arginina no es trobin en conformacions equivalents a les diferents subunitats. L'anàlisi dels llocs d'unió a nucleòtid de les posicions no equivalents revela diferents ocupàncies nucleotídiques a cadascun dels tres tipus de butxaca nucleotídica, que es van designar com estat d'unió competent a NTP (T), d'unió competent a NDP (D) i buit (E).

Aquests estats estan correlacionats amb la posició vertical del llaços d'unió al DNA quan es visualitzen perpendiculars al canal central, l'eix putatiu de translocació del DNA. Aquesta asimetria de l'anell proporciona una explicació per la cooperativitat negativa observada, alhora que suggereix un mecanisme per a l'acoblament de la hidròlisi de nucleòtids amb la translocació del ssDNA. Es va inferir un mecanisme de translocació rotacional conduïda per hidròlisis de NTP seqüencials, durant el qual els llocs d'unió a NTP permuten cíclicament entre l'estat T, D i E a mesura que els seus llaços d'unió al DNA associats es mouen de la part superior del canal (l'entrada) a la part inferior (sortida). El ssDNA seria transferit d'una subunitat unida a NTP a la subunitat veïna, que proporciona el dit d'arginina. A la figura I.6 es representa esquemàticament aquest model. El mecanisme proposat estava d'acord amb estudis cinètics previs que suggerien una ruta de reacció rotacional no concertada i tenia varies analogies notables al model proposat per a la F1-ATPasa.

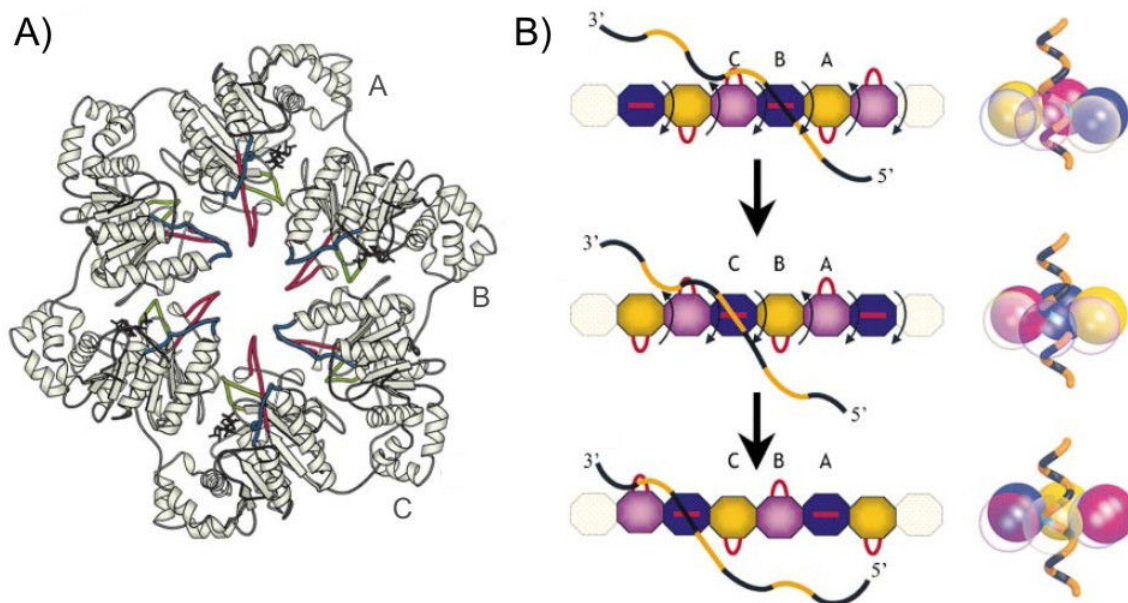


Figura I.6: Model de transferència entre subunitats per a la translocació de ssDNA proposat per l'helicasa de T7. A) Localització dels llaços d'unió a DNA dins l'hexàmer de gp4. **B)** Representació esquemàtica del mecanisme de translocació en un l'hexàmer obert formant una línia. Les subunitats en blau es troben en l'estat T, les grogues en el D i les liles en l'estat E. Els llaços d'unió a ssDNA es representen en vermell i el ssDNA en groc i negre per il·lustrar el seu pas al llarg de l'anell, amb cada banda de color representant una unitat de translocació. A l'esquerra es mostra la vista lateral de la translocació, amb les tres subunitats frontals de l'hexàmer representades en transparència per claredat. Figura adaptada de Singleton *et al.*, 2000.

3.3.2 Model concertat (tot o res)

Subsegüentment, es va determinar l'estructura hexamèrica de l'helicasa de SF3 antígen gran (Tag) de SV40 en tres estats nucleotídics diferents: (Tag-ATP)₆, (Tag-ADP)₆ i (Tag-buit)₆ (Li *et al.* 2003; Gai *et al.* 2004). Contràriament a l'estructura de l'helicasa de T7, aquestes estructures eren altament simètriques, especialment (Tag-ATP)₆ i (Tag-buit)₆. L'observació d'aquests tres estats nucleotídics independents i la mancança aparent d'espècies nucleotídiques mixtes portaren a la hipòtesi de que la molècula adoptava exclusivament configuracions de tot o res als sis llocs d'unió a ATP. Aquestes conformacions eren mantingudes col·lectivament al llarg del cicle NTPasa: hidròlisi concertada a totes 6 interfícies entre subunitats, seguida d'un alliberament d'ADP concertat, seguit d'una unió concertada de molècules d'ATP a cadascuna de les interfícies per completar així el cicle. En aquestes estructures, així com a gp4 de T7, les posicions dels llaços d'unió a DNA de l'interior de l'anell hexamèric estaven correlacionades amb l'estat nucleotídic, amb (Tag-ATP)₆ col·locant els llaços a la part superior del canal i (Tag-buit)₆ a la part inferior. Al contrari que pel model T7gp4, es proposava que el DNA entrava al complex a la cara associada amb les configuracions buides en comptes de per la cara associada amb la unió a ATP.

3.3.3 Model d'escorta coordinada

Finalment, s'obtingué la primera estructura d'una helicasa hexamèrica unida amb ssDNA: l'helicasa E1 del virus del papil·loma boví (sense el domini N-terminal d'unió a l'origen), pertanyent a SF3 (Enemark *et al.* 2006). A l'estructura, el domini d'oligomerització forma un collar amb simetria sis que estabilitza l'hexàmer, amb els dominis AAA⁺ desplaçats respecte a una simetria radial perfecte. Sis bases de l'oligonucleòtid de cadena senzilla són unides al canal central de l'hexàmer mitjançant interaccions estretes entre l'esquelet fosfodiester i sis llaços d'unió a ssDNA que disminueixen progressivament d'alçada al llarg de l'anell. Conseqüentment, les interaccions entre aquests llaços i el ssDNA generen una disposició semblant a una escala de cargol de mà dreta, la qual reforça l'asimetria radial de l'anell.

Encara que cadascuna de les sis butxaques nucleotídiques de l'hexàmer estan ocupades per ADP, no són equivalents, sinó que hi ha tres coordinacions nucleotídiques diferents, les quals s'interpreten com d'unió a ATP (T), unió a ADP (D) i buida (E). Aquesta classificació està parcialment derivada de la proximitat de les dues subunitats que conformen la butxaca nucleotídica bipartida. Al tipus T, els residus catalítics procedents d'ambdues subunitats veïnes es posicionen de manera compacta al voltant del nucleòtid unit. Al tipus D, els residus procedents de les subunitats adjacents estan més allunyats però encara interactuen. Al tipus E, no existeix cap mena d'interacció. Si ens moguem al voltant de l'anell, hi ha dos o tres llocs del tipus ATP, dos del tipus ADP i un o dos llocs buits. Al contrari que per a la F1-ATPasa i el model proposat per gp4, hi ha múltiples llocs de cada tipus presents a l'hexàmer i estan agrupats seqüencialment al voltant de l'anell (veure figura I.7). Paral·lelament, tal i com ja s'havia observat per gp4D i Tag, la posició vertical dels llaços d'unió al ssDNA de l'interior del canal es correlaciona amb l'estat de la butxaca nucleotídica associada. El llocs T tenen els llaços a la part de dalt de l'espiral (és a dir, a l'extrem 5' de la cadena de DNA), els llocs D els tenen a nivells entremitjos i els llocs E col·loquen els llaços a la part de baix.

En base a aquestes observacions, es proposa un mecanisme d'escorta coordinada per la translocació unidireccional del ssDNA dependent d'ATP. Els llaços d'unió interaccionen amb l'esquelet del DNA monocadenari a mesura que aquest passa pel canal central, de manera que el DNA és unit simultàniament per múltiples subunitats des dels llaços T a l'extrem superior de l'hexàmer fins als llaços E a l'inferior. Durant la progressió del cycle catalític, els llaços es mouen cadascun un pas cap avall, estirant el DNA amb ells. La hidròlisi de l'ATP i el moviment dels llaços està coordinada entre les subunitats adjacents gràcies a una sèrie d'interaccions mantingudes entre residus bàsics i àcids dels llaços d'unió a DNA procedents de monòmers adjacents. Aquestes interaccions s'anomenen "formadores d'escala", doncs estableixen l'arranjament en escala de cargol que formen els llaços d'unió a àcid nucleic. A mesura que el llaç assoleix el fons del canal, allibera el DNA monocadenari i torna altre cop a la part superior per unir una nova molècula d'ATP i tornar a començar el cycle. La característica diferencial d'aquest model és que cada llaç d'unió al DNA manté un contacte continu amb un nucleòtid del ssDNA i tot l'arranjament en escala de cargol migra col·lectivament cap a baix al produir-se la hidròlisi de l'ATP i l'alliberament de

l'ADP. Quan un llaç arriba al fons de l'escala, allibera el ssDNA associat així com les interaccions formadores d'escala amb el llaç d'unió a DNA adjacent. Subsegüentment, una molècula d'ATP s'uneix a la interfície lliure i, al unir el proper nucleòtid del ssDNA, el llaç migra cap a la posició de dalt de l'escala i s'acobla al següent llaç mitjançant les interaccions formadores d'escala. Una característica important d'aquest mecanisme és que la posició de cada llaç d'unió al DNA està governada no només per la configuració associada amb el seu lloc d'unió a ATP, sinó també per les posicions dels llaços d'unió al DNA de les subunitats adjacents gràcies a les interaccions formadores d'escala que s'estableixen entre ells.

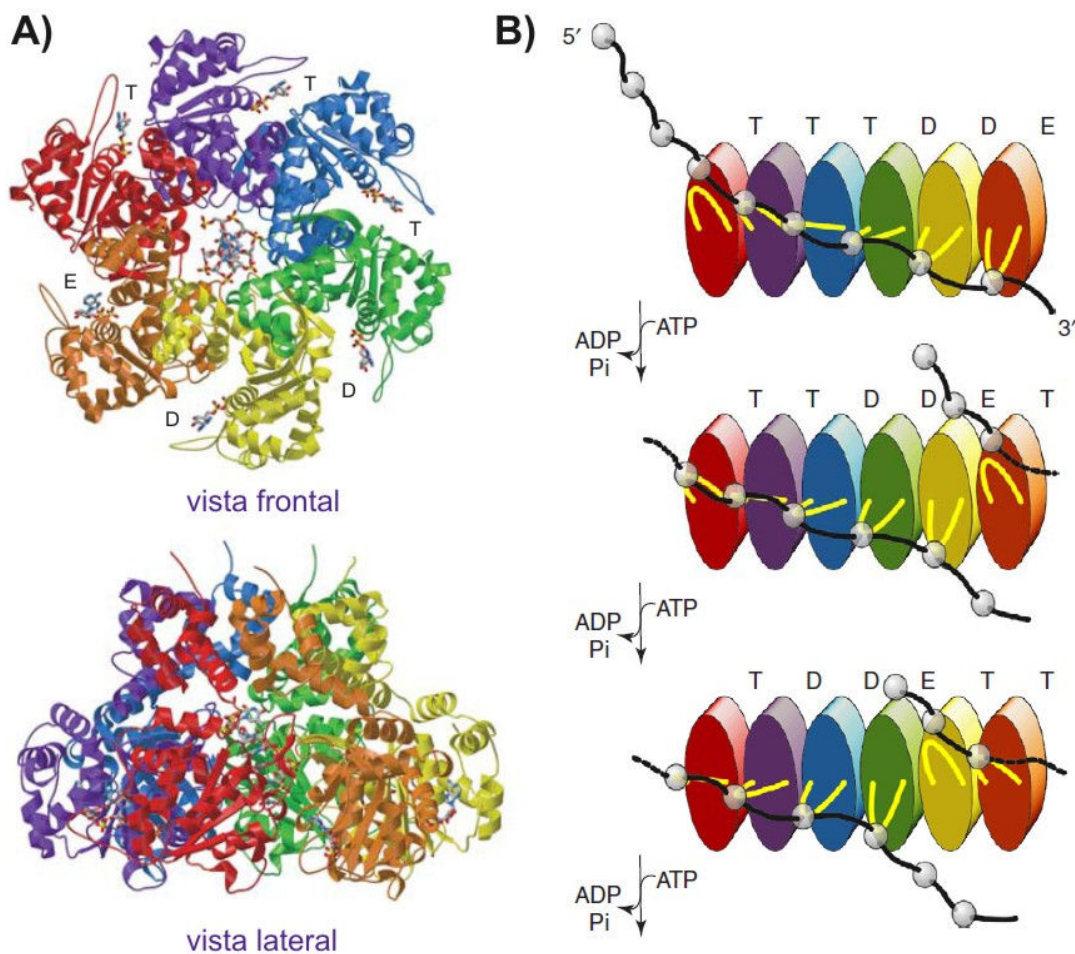


Figura 1.7: Mecanisme d'escorta coordinada per a la translocació de ssDNA proposat per l'helicasa E1 de BPV. A) Model cristal·logràfic de l'hexàmer d'E1 unit a un 13-mer de ssDNA i a ADP. La vista frontal és des l'extrem 5' del ssDNA. Cada monòmer es representa en un color diferent i s'indica l'estat de la butxaca d'unió a nucleòtid a cada interfície entre subunitats. Les molècules d'ADP i el ssDNA es representen en *sticks*. Codi PDB: 2GXA. **B)** Representació esquemàtica de l'hexàmer E1 obert formant una línia. Les subunitats vermella i taronja estan en contacte en l'hexàmer tancat. Cada llaç d'unió a DNA (en groc) contacta un nucleòtid (gris). El panell superior equival a la estructura de (A), amb el llaç del monòmer vermell en la posició més alta de l'escala de cargol. La translocació del ssDNA, en passos d'un sol nucleòtid, està

conduïda pel moviment coordinat de tots sis llaços en resposta a la unió, hidròlisi i alliberament net d'una molècula d'ATP. A cadascun dels cicles, cada subunitats transloca un nucleòtid i cada interfície intersubunitat hidrolitza una molècula d'ATP i allibera un molècula d'ADP. Un cicle complert doncs, transloca sis nucleòtids de ssDNA, hidrolitza 6 ATP i allibera 6 ADP. Figura adaptada de Ha, 2007.

3.3.4 Consideracions finals

Les parelles de residus implicades en les interaccions formadores d'escala que s'estableixen entre residus bàsics i àcids als llaços d'unió a ssDNA d'E1 estan conservades en els llaços d'unió a àcid nucleic de les helicases de SF4 i, de fet, també interactuen estructuralment en el cas de gp4 de T7 (Singleton *et al.* 2000). Això suggereix que els membres de SF4 podrien emprar un mecanisme d'escorta coordinada com el descrit per l'E1 de SF3.

Tot i que Tag, l'altre membre de SF3 caracteritzat estructuralment, s'ha proposat que actua mitjançant un mecanisme d'hidròlisi i translocació concertat, la semblança estructural i funcional d'aquesta proteïna amb E1 suggereix que ambdues podrien emprar el mateix mecanisme i que la simetria observada a Tag es trencaria al interaccionar amb l'àcid nucleic de cadena senzilla. Alternativament, podria ser que Tag (i la resta de membres de SF3) empressin un mecanisme concertat per fondre l'origen i després canviessin a un mecanisme d'escorta coordinada (Enemark *et al.* 2008).

Finalment, una característica interessant comuna a tots aquests models és que la geometria dels llaços i la seva interacció amb l'àcid nucleic fa que es necessiti molt poca rotació del DNA en relació a la proteïna, reduint els possibles problemes de superenrotllament durant la replicació.

3.4 Mecanismes de desenrotllament del dúplex

Igual que en el cas de les helicases monomèriques, a les helicases hexamèriques el desenrotllament del substrat està acoblat a la translocació. A continuació es descriu el model més àmpliament acceptat, juntament amb les altres propostes existents a la

literatura. Finalment, es discuteix sobre si les helicases en forma d'anell empenen mecanismes actius o passius pel desenrotllament.

3.4.1 Model d'exclusió estèrica

Les evidències bioquímiques i estructurals de que l'àcid nucleic s'uneix al canal central de l'anell hexamèric i que les helicases interactuen més fortament amb una de les monocadenes de l'àcid nucleic, fan que actualment el mecanisme més consensuat de desenrotllament del dúplex impliqui un model d'exclusió estèrica, originalment anomenat model de cunya mecànica (*wedge*). En aquest model, una cadena de l'àcid nucleic a desenrotllar passa a través de l'anell hexamèric i l'altra cadena queda completament fora d'aquest, de manera que la translocació unidireccional de l'helicasa al llarg d'una de les cadenes senzilles mentre s'exclou l'altra separa físicament el dúplex. En aquest model l'helicasa no interaccionaria ni amb la cadena exclosa ni amb la regió de doble cadena.

La majoria, potser totes, les helicases hexamèriques operen per mitjà d'aquest model, basant-nos en la seva capacitat per passar per sobre d'un substituent voluminós quan és present en una de les cadenes (l'exclosa) però no a l'altra (sobre la que es transloquen). Per als membres de SF4, evidències bioquímiques (Bujalowski *et al.* 1995; Hacker *et al.* 1997; Jezewska *et al.* 1998a; Kaplan 2000) i de microscòpia electrònica (Yu *et al.* 1996a) suggereixen que una sola cadena de DNA passa a través del canal amb la segona cadena desplaçada fora de l'anell. Igualment, aquest model també és consistent amb l'estructura de l'helicasa de SF3 E1 en complex amb DNA de cadena senzilla.

Inicialment, s'havia proposat una variant d'aquest model d'exclusió estèrica, el model torsional, en el qual l'helicasa interaccionaria tant amb la cadena que es transloca a través del seu canal interior com amb la cadena exclosa. Empraria aquestes interaccions per rotar una cadena respecte l'altre, desestabilitzant així el dúplex. També s'havia proposat un model de desestabilització del dúplex, en el qual l'helicasa interaccionaria directament amb aquesta regió (Patel *et al.* 2000). Ambdós models s'han

descartat perquè no hi ha evidències experimentals de que les helicases hexamèriques interaccionin amb la cadena exclosa o amb el dúplex.

3.4.2 Altres models

Per altres membres de SF3, així com per les restants famílies amb plegament AAA+ (SF6 i SF7), s'han proposats models en que l'enzim encercla les dues cadenes del DNA. Un problema evident de l'encerclament del DNA de doble cadena és que no hi ha cap manera evident d'impedir el reanellament immediat de les cadenes a mida que aquestes passen a través de l'hexàmer. Això és podria solucionar postulant que altres proteïnes accessòries, com ara altres components de la forqueta de replicació, s'uneixen a les cadenes senzilles separades i impedeixen el reanellament.

Per exemple, per Tag s'han proposat ambdós tipus de models, tant d'exclusió estèrica amb translocació de ssDNA com d'encerclament i translocació de dsDNA. Els models que impliquen DNA de doble cadena es basen en que la unitat funcional com a helicasa replicativa és un doble hexàmer encarat cap a cap, cadascun del quals bombejaria dsDNA cap al seu interior. La fusió del dúplex es donaria bé dins del canal mitjançant distorsions electroestàtiques o gràcies a la força generada al empènyer el DNA dúplex contra ell mateix. El ssDNA resultant sortiria a través de la zona central del dodecàmer gràcies a canals formats entre els dominis d'unió a DNA dels dos hexàmers associats o entre els dominis d'unió a DNA i els AAA+ de cada hexàmer (Hickman *et al.* 2005). Aquest model es basava en les imatges de microscòpia electrònica dels dobles hexàmers formats per Tag sobre *oriC*, doncs s'observava una massa que sobresortia de la zona central del doble hexàmer i que s'havia atribuït al ssDNA (Valle *et al.* 2000). Ara bé, les estructures cristal·lines de Tag mostren que el canal interior presenta diàmetres inferiors a 7 Å en l'estat unit a ATP i d'uns 15 Å sense nucleòtid, insuficients per permetre el pas del dsDNA que té aprox. 20Å de diàmetre, per tant, només hi cabria ssDNA (Li *et al.* 2003; Gai *et al.* 2004). Aquest fet, juntament amb que l'helicasa E1, també pertanyent a SF3, uneixi ssDNA al seu interior, ha portat a que actualment es consideri que les helicases de SF3 separen els dúplexs d'àcid nucleic mitjançant un mecanisme d'exclusió estèrica (Singleton *et al.* 2007). En tot cas, similarment al que ja s'ha comentat pels mecanismes d'hidròlisi d'ATP i translocació,

el bombeig de dsDNA i la formació de dobles hexàmers estaria associat específicament a l'etapa de fusió de l'origen d'aquestes proteïnes, que després canviarien de mecanisme.

Per a les helicases de SF6, la situació és similar doncs, en base a la seva capacitat de translocar-se tant sobre DNA de doble cadena com de cadena senzilla, s'han proposat tant mecanismes d'exclusió estèrica com similars al descrit anteriorment per Tag de bombeig de dsDNA per part d'un doble hexàmer (Takahashi *et al.* 2005). En el cas de separar les bases del DNA mitjançant l'exclusió d'una de les cadenes, s'ha proposat que aquesta proteïna tindria un comportament equivalent a l'observat per DnaB, la qual és capaç de translocar-se sobre dsDNA sense desenrotllar-lo, però canvia a una activitat pròpiament helicasa quan troba un substrat amb una cua de cadena senzilla adequada (Kaplan *et al.* 2003).

Finalment, un model d'encerclament del dsDNA és consistent amb la funció de la translocasa RuvB, la qual actua com a una bomba de dsDNA, amb el dúplex separat subsegüentment per la proteïna associada RuvA.

4.4.3 Desenrotllament actiu o passiu?

En el model més simple, la força generada per la translocació direccional al llarg del ssDNA separa les cadenes complementàries sense necessitar un mecanisme específic de fusió dels parells de bases. Els primers càlculs mecanístics semblaven indicar que les helicases replicatives desenrotllaven els àcids nucleics a través d'un mecanisme passiu. Així doncs, es va proposar que atrapar la cadena senzilla originada per les fluctuacions tèrmiques de la forqueta ja era suficient per generar la ràpida taxa d'activitat helicasa, concretament associada a la replicació genòmica, observada *in vivo* (Delagoutte *et al.* 2002). Igualment, es proposà que Rho, el factor de terminació de la transcripció, també actuaria a través d'un mecanisme passiu (von Hippel *et al.* 2001). Similarment, en base a la important reducció observada entre la taxa de translocació i desenrotllament de gp4 de T7 (nou vegades inferior), també s'inferí que gp4 empraria un mecanisme passiu per separar les bases del dsDNA (Jeong *et al.* 2004).

L'avenç de les tècniques de molècula senzilla (*single molecule*) ha permès realitzar uns càlculs molt més acurats i sofisticats de l'activitat tant de translocació com de separació de cadena de les helicases (Rasnik *et al.* 2006). Actualment, dues helicases hexamèriques han estat caracteritzades mitjançant aquesta tècnica: gp4 i gp41 dels fags T7 i T4, respectivament. Encara que ambdues pertanyen a la mateixa superfamília, SF4, i funcionen com a helicases replicatives, s'ha deduït que gp4 de T7 emprava un mecanisme de desenrotllament actiu (Johnson *et al.* 2007), mentre gp41 utilitza un mecanisme passiu (Lionnet *et al.* 2007).

Segons les dades de *single-molecule*, gp4 de T7 desenrotlla el dsDNA gràcies a un mecanisme actiu, en el qual desestabilitza la forqueta bé per repulsió electroestàtica (degut a la càrrega superficial predominantment negativa de l'exterior de l'hexàmer) i/o per torcedura severa del ssDNA entrant degut a l'exclusió estèrica. La taxa de desenrotllament depèn de la composició de la cadena a separar, amb taxes més baixes per a zones riques en AG i més altes per a zones riques en AT. A més, es demostrà que al aplicar una força externa en el sentit de desenrotllament de la forqueta, l'helicasa augmentava significativament la seva processivitat i velocitat de desenrotllament fins a propiar-se als valors observats *in vivo*. Aquests resultats indicaven que gp4 emprava un mecanisme parcialment actiu per desenrotllar el DNA, que podia ser ajudat *in vivo* per l'associació amb altres components del replisoma. Concretament, la polimerasa que s'associa a la cadena líder, mentre l'helicasa ho fa a la retardada, podria generar la força complementària que augmentaria la velocitat de l'enzim, fet que explicaria l'activació de l'helicasa per part de la polimerasa *in vivo* (Hamdan *et al.* 2007). Aquesta cooperació té una gran rellevància biològica, doncs implica que l'helicasa només és totalment activa en el context del replisoma, evitant així desenrotllament incontrolat a zones no permeses del genoma. Paral·lelament, un estudi mecanoquímic de gp4 també dona suport a un mecanisme helicasa parcialment actiu (Donmez *et al.* 2007).

Per contra, estudis similars realitzats amb l'helicasa gp41 del fag T4 indiquen que aquesta emprava un mecanisme passiu (Lionnet *et al.* 2007). També s'observa un augment de la processivitat i la velocitat de desenrotllament al aplicar una força externa, en concordança amb la major activitat de la proteïna quan s'associa als altres enzims que actuen a la forqueta de replicació.

El fet que dues helicases tan semblants estructural i funcionalment semblin emprar mecanismes diferents, exemplifica la dificultat d'arribar actualment a una conclusió general respecte a la natura del desenrotllament emprat per les helicases hexamèriques. Caldrà esperar a noves dades i millores en les tècniques per mesurar l'activitat d'aquestes helicases.

II. LA PROTEÏNA DNAB

1. Caracterització històrica

El gen *dnaB* va ser caracteritzat a partir de l'aïllament de mutants de la bactèria *Escherichia coli* sensibles a temperatura, els quals eren incapaços de replicar el seu genoma a temperatures no permeses. El fet que a aquests mutants es produís una parada immediata de la síntesi de DNA va portar a considerar el producte gènic de *dnaB* essencial per a la replicació del cromosoma bacterià, concretament per la propagació de la forqueta de replicació (Wechsler *et al.* 1971).

La posterior purificació de la proteïna DnaB va permetre començar a caracteritzar-la: es va determinar la seva mida funcional aproximada (uns 250 kDa, amb monòmers de 55 kDa), així com la seva activitat ATPasa (Wickner *et al.* 1974; Ueda *et al.* 1978). Posteriors estudis van concloure que aquesta activitat ATPasa era altament dependent de la presència d'àcids nucleics, tant de cadena senzilla com doble i requeria Mg^{2+} . Igualment, es va caracteritzar la seva capacitat d'unió a ssDNA, la qual també necessitava la presència de Mg^{2+} i estava estabilitzada per la unió a ATP (Arai *et al.* 1981ba; Arai *et al.* 1981ab).

En quan a la possible funció de DnaB durant la replicació, inicialment, basant-se en la seva interacció amb DnaG (la primasa) en presència d'ATP, es va hipotitzar que DnaB estaria implicada en l'estimulació de la síntesi d'encebadors per part de la primasa, segurament gràcies a proporcionar-li un substrat de DNA adequat (Ueda *et al.* 1978). En estudis de la replicació del genoma del fag $\phi X174$, sistema models doncs hi actuen els factors de la cèl·lula hoste necessaris per la iniciació de la síntesi a la cadena retardada, es va determinar que només els enzims DnaB, DnaG i la polimerasa III eren necessaris per a la replicació i que, en absència de polimerasa, DnaB i DnaG podien catalitzar la síntesi d'encebadors (Arai *et al.* 1979). Aquests resultats semblaven donar suport a la idea de que DnaB estaria implicada en el sistema d'encebatge catalitzat per DnaG.

Finalment, es va caracteritzar la funció de DnaB com a helicasa de DNA, activitat que realitza impulsada per la hidròlisi d'ATP amb polaritat de 5' a 3' i que és estimulada per la presència de la primasa i les proteïnes SSB (LeBowitz *et al.* 1986). També es va comprovar que DnaB necessita una forqueta de replicació preformada per a poder desenrotllar el DNA dúplex. Amb aquests resultats, DnaB es va convertir en la primera helicasa hexamèrica caracteritzada. A més, la descripció d'aquesta nova activitat de desenrotllament del dsDNA, juntament amb les observacions prèvies de que DnaB era un component essencial de la maquinària replicativa a *E. coli* i que realitzava la seva funció en col·laboració amb DnaG, va portar a considerar-la la principal helicasa replicativa bacteriana. Com a tal, actuaria a la cadena retardada de la forqueta de replicació desenrotllant el dúplex de DNA, alhora que serviria de lloc de reclutament mòbil per la primasa (LeBowitz *et al.* 1986).

Aquesta hipòtesi es va confirmar al comprovar-se que DnaB és la única helicasa necessària per reconstituir la replicació del DNA des d'un origen de replicació cromosomal (*oriC*) *in vitro*, procés en el qual actua conjuntament amb la primasa i l'holoenzim polimerasa III, els altres dos components essencials (Baker *et al.* 1987). Aquests estudis de replicació bacteriana *in vitro* van determinar que DnaB era essencial durant l'etapa d'iniciació de la replicació, durant la qual era reclutada a l'origen per DnaA (proteïna iniciadora) i DnaC (carregador d'helicasa), així com per la posterior etapa d'elongació. Igualment, es va constatar que l'activitat helicasa de DnaB es reforçava en presència de DnaG i de l'holoenzim polimerasa III.

2. Organització en dominis

Primerament, els dominis de DnaB es van definir mitjançant proteòlisi limitada de la proteïna sencera d'*E. coli*. La proteòlisi alliberava en primer lloc un oligopèptid de 14 residus de la part N-terminal de la proteïna, per generar un fragment de 50 kDa que es va designar com a fragment I. La ulterior proteòlisi del fragment I generava dos nous fragments proteolítics, els fragment II i III. El fragment II tenia 33 kDa i corresponia a la regió C-terminal de la proteïna, mentre que el III tenia 12 kDa i corresponia a la regió N-terminal. Es va concloure que el monòmer de DnaB estava format per dues regions, un domini N-terminal més petit i compacte i un domini C-terminal que representava dos tercers parts de la proteïna, separats per una regió d'enllaç flexible d'uns 45 residus, que es perdien durant l'experiment de proteòlisi (Nakayama *et al.* 1984).

El subsegüent anàlisi de les activitats específiques de cadascun d'aquests fragments proteolítics demostraren que el fragment III, és a dir el domini N-terminal de la proteïna, no era necessari per a l'activitat d'unió a NTP, NTPasa i d'unió a DNA del fragment II C-terminal, però en canvi era essencial per a l'activitat helicasa, la qual estava totalment inhibida en absència del domini N-terminal. Aquesta regió N-terminal, que no presenta cap activitat enzimàtica específica, també semblava ser important pel manteniment de l'estructura hexamèrica, doncs els dominis C-terminals aïllats eren menys estables com hexàmer. Finalment, el fragment I tenia una activitat NTPasa, d'unió a DNA i helicasa equivalent a la de la proteïna sencera, indicant que els 14 primers residus de la proteïna no eren essencials per l'activitat d'aquesta (Nakayama *et al.* 1984; Biswas *et al.* 1994).

Els límits entre els dominis de DnaB definits per proteòlisi es confirmaren mitjançant estudis de ressonància magnètica nuclear del domini N-terminal, els quals determinaren que la regió compresa entre els residus 24 i 136 és una unitat de plegament independent i es troba estructuralment ben definida, amb un plegament d'hèlixs α . També es comprovà que els residus 1-23 i 137-161 estan en conformació *random coil* i són mòbils a escala de nanosegons. Així doncs, aquests estudis permeteren definir més concretament les fronteres del domini N-terminal i comprovar

l'existència d'una regió flexible d'enllaç que actuaria de frontissa entre el domini N i C-terminal (Miles *et al.* 1997).

3. Forma funcional

Com ja indicaven els primers estudis sobre l'helicasa DnaB purificada (Ueda *et al.* 1978; Arai *et al.* 1981a), la caracterització bioquímica rigorosa de l'estat d'oligomerització funcional de DnaB va confirmar que en solució és un hexàmer estable d'uns 310 kDa, específicament estabilitzat per múltiples cations de magnesi (Mg^{2+}). Igualment, la comparació de les taxes de sedimentació de l'hexàmer i el monòmer de DnaB en experiments d'ultracentrifugació analítica suggeria que l'helicasa formava una associació amb simetria cíclica, on els contactes entre protòmers estarien limitats a només dues subunitats veïnes (Bujalowski *et al.* 1994).

Aquesta simetria cíclica de l'oligòmer es va confirmar amb les primeres imatges de microscòpia electrònica de DnaB, obtingudes mitjançant tinció negativa, les quals mostraven un hexàmer en forma d'anell amb clara simetria rotacional tres. Aquestes imatges, preses en presència d'ATP, presentaven sis grans pics de densitat que suggerien que l'helicasa estava disposada com a trímer de dímers, els quals deixaven un canal central buit al llarg de tota la molècula d'uns 4 nm de diàmetre. El model tridimensional generat a partir d'aquestes imatges a una resolució de 2.7 nm permetia la diferenciació dintre de cada subunitat de l'hexàmer d'una regió més petita, atribuïda al domini N-terminal, i una regió més gran, atribuïda al domini C-terminal. Igualment, atribuïa una certa polaritat estructural en l'associació oligomèrica de DnaB (San Martin *et al.* 1995).

Posteriors estudis de microscòpia electrònica, també mitjançant tinció negativa, de l'helicasa DnaB a un valor de pH de 7.2 i en presència de varis factors nucleotídics (ADP, ATP i anàlegs no hidrolitzables d'ATP) mostraren que, a més dels anells hexamèrics de simetria rotacional tres (C3) trobats anteriorment, també s'observaven anells hexamèrics amb simetria rotacional sis (C6), a més d'anells en un estat intermedi.

En tots els casos, es necessitava la presència de nucleòtids per observar anells. Igualment, a totes les condicions estudiades hi havia un equilibri gairebé simètric entre les dues conformacions, la circular i la triangular. Aquestes dades també semblaven indicar que la unió d'ADP provocava una reducció d'uns 10 Å en el diàmetre global de l'anell comparat amb el valor en presència d'AMP-PNP, tot i que donada la resolució de les imatges no era conclouent (Yu *et al.* 1996bb).

La subsegüent reconstrucció per criomicroscòpia electrònica a una resolució de 35 Å de l'hexàmer de DnaB a pH 7.6 i en presència d'ADP i Mg²⁺ (San Martin *et al.* 1998), el primer model obtingut per a una helicasa hexamèrica mitjançant criomicroscòpia, confirmà les característiques estructurals que ja s'havien descrit al model tridimensional generat a partir de preparacions de partícules tenyides per tinció negativa. Així doncs, l'oligòmer de DnaB també presentava sis màxims de densitat organitzats en un arranjament exclusivament triangular, cosa que suggeria, com ja s'havia apuntat, l'existència d'un trímer de dímers. Ara bé, a diferència del model anterior (San Martin *et al.* 1995), els dímers consecutius que formaven l'hexàmer són asimètrics, amb un de cada dos presentant un volum major que els altres. Igualment es confirmà l'existència d'un canal central d'uns 3 nm de diàmetre, amb un diàmetre total de la partícula de 12,5 nm i una alçada de 5,7 nm, superior a la determinada anterior, de 3,8 nm, tal i com és esperable de les diferents tècniques emprades. Per últim, les dades de criomicroscòpia evidencien amb major claredat que les dues cares de l'hexàmer són diferents, mostrant una transició entre simetria 6 a tres a mesura que ens moguem al llarg de la partícula. Aquesta polaritat ja havia estat observada per altres helicases hexamèriques com ara gp4 de T7 i SV40. A la figura I.8c es mostra aquesta reconstrucció tridimensional de DnaB.

Una posterior reconstrucció, també per criomicroscòpia, del complex format per les proteïnes DnaB i DnaC (l'helicasa i la seva proteïna carregadora a l'origen) d'*E. coli* a una resolució de 26 Å va permetre assignar la cara de l'hexàmer de DnaB que presenta simetria 3 als dominis N-terminal i la cara que presenta simetria 6 als dominis C-terminals (Barcena *et al.* 2001).

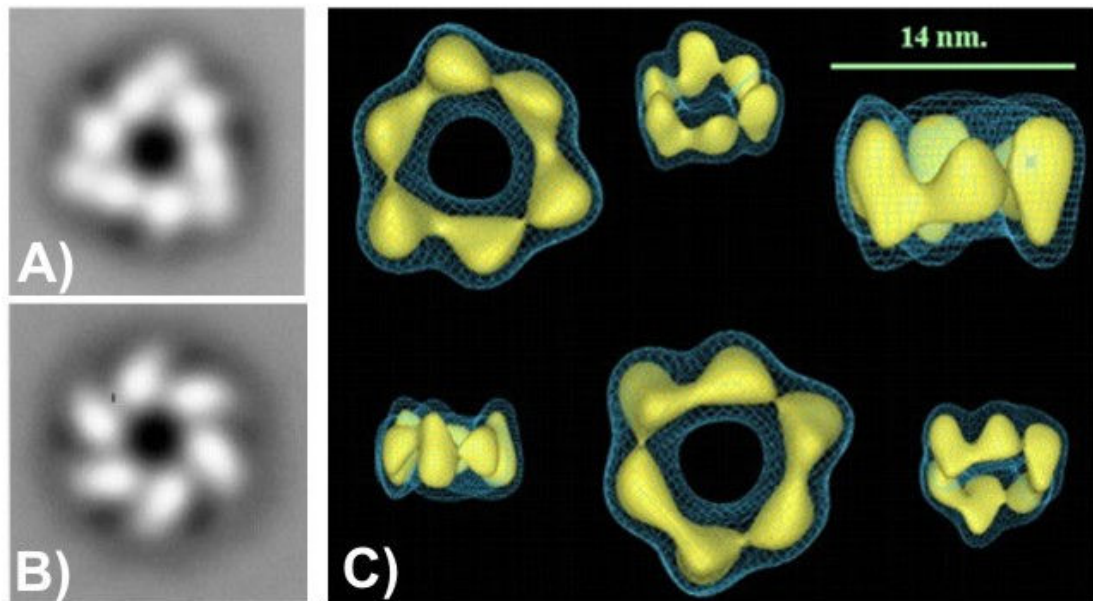


Figura 1.8: Anàlisi de la DnaB d'*E. coli* per microscòpia electrònica. Es mostren imatges de microscòpia electrònica obtingudes mitjançant tinció negativa en les quals s'observa la simetria C3 (A) i C6 (B) que poden presentar els hexàmers de DnaB (Donate *et al.* 2000). També es mostra la reconstrucció tridimensional a una resolució de 35 Å (C) generada a partir d'imatges de criomicroscòpia electrònica (San Martín *et al.* 1998).

En quant a l'existència de dues simetries per a l'hexàmer de DnaB, un anàlisi acurat per microscòpia electrònica de mostres obtingudes per tinció negativa de DnaB en presència de diferents nucleòtids i a diferents pHs de treball, va permetre establir que la variabilitat conformacional observada depèn exclusivament del pH de treball i no del nucleòtid. A pH de 7.6 o superior, només es detecta la forma que presenta simetria C3, mentre que a pH més baixos, inferiors o iguals a 7.2, coexisteixen la forma C3 i C6 en proporcions gairebé equivalents. Aquest canvis en l'estat quaternari de DnaB són totalment reversibles entre els valors de pH 7.2 i 7.6. Cal remarcar que els pHs inferiors a 6,5 tenen un efecte dramàtic sobre la integritat estructural dels oligòmers de DnaB, doncs per sota d'aquest valor la majoria d'hexàmers es dissocien i deixen de ser viables. Per contra, el pH no influeix la simetria del complex entre DnaB i DnaC, que sempre presenta simetria C3 (Donate *et al.* 2000). El fet que DnaB en complex amb DnaC sigui inactiva tant com a NTPasa com a helicasa, suggeriria que la forma amb simetria C3 representava l'estat inactiu de l'hexàmer, mentre que segurament la transició entre els estats C3 i C6 era necessària per la funció de DnaB.

3.1 Interaccions amb altres proteïnes

Com a principal helicasa replicativa bacteriana, DnaB no actua de forma aïllada a la cèl·lula, sinó que ho fa en associació amb les altres proteïnes que constitueixen la maquinària replicativa dels genomes bacterians, anomenada replisoma. Inicialment, es va suggerir que la regió N-terminal de DnaB seria la responsable de les interaccions entre l'helicasa i els altres membres del replisoma (Nakayama *et al.* 1984). Actualment s'ha descrit que, efectivament, les regions d'interacció amb DnaA (proteïna iniciadora de la replicació) i DnaG es troben fins del domini N-terminal de la proteïna. Ara bé, la interacció amb DnaC és realitzada a través del domini C-terminal de DnaB (Barcena *et al.* 2001), encara que la presència del domini N-terminal sembla influenciar aquesta associació (Biswas *et al.* 1999). DnaB també interacciona amb la subunitat τ de l'holoenzim polimerasa III.

La vital importància d'aquestes interaccions per regular l'activitat tant de l'helicasa com dels seus companys d'unió dins del context de la replicació del DNA es descriuen a l'apartat III d'aquesta introducció.

4. Caracterització bioquímica

4.1 Llocs d'unió a nucleòtid

L'hexàmer de DnaB té 6 llocs d'unió a nucleòtid que presenten una forta cooperativitat negativa entre ells, la qual augmenta amb la temperatura, amb tres llocs d'alta afinitat i tres llocs de baixa afinitat (Biswas *et al.* 1986; Bujalowski *et al.* 1993).

S'ha demostrat que DnaB pot unir tant ADP com ATP així com varis anàlegs d'aquests nucleòtids com ara TNP-ATP, TNP-ADP, TNP-AMP, MANT-ADP, ϵ ADP i ATP- γ -S. L'afinitat per l'ADP i els seus anàlegs és major que per als anàlegs d'ATP (Biswas *et al.* 1986; Bujalowski *et al.* 1993)

4.2 Interacció amb ssDNA

En un complex estacionari, en presència d'un anàleg no hidrolitzable d'ATP, un hexàmer de DnaB encercla unes 20 bases de ssDNA, independentment de la seqüència nucleotídica d'aquest àcid nucleic. Estudis termodinàmics de la unió de DnaB a diferents oligòmers de ssDNA mostren que l'enzim té un sol lloc d'unió fort per al ssDNA, mentre que l'anàlisi per fotoentrecreuament químic indica que només hi ha una subunitat, com a molt dues en cooperació, en contacte amb el ssDNA. Aquests tres resultats, considerats conjuntament, portaren a excloure la possibilitat de que l'àcid nucleic s'emboïllés al voltant de l'hexàmer durant el procés de translocació i/o desenrotllament, doncs en aquest cas el nombre de nucleòtids i subunitats de l'hexàmer que interaccionarien seria molt superior (Bujalowski *et al.* 1995; Jezewska *et al.* 1996). Ara bé, aquestes dades no permetien discriminar entre la possibilitat que el DNA de la cadena senzilla passes per l'exterior de l'hexàmer, interaccionant amb un sol monòmer, o ho fes per l'interior del canal central d'aquest.

Les evidències bioquímiques de que la cadena de ssDNA passa pel canal central de l'hexàmer s'obtingueren finalment mitjançant estudis de transferència d'energia de fluorescència de la unió de DnaB a oligòmers de diverses llargàries (Jezewska *et al.* 1998aa). Es determinà que els dominis N-terminals de l'anell hexamèric estaven disposats propers a l'extrem 5' de la cadena de DNA unida a DnaB. També es determinà que aquesta disposició és la mateixa que presenta l'hexàmer de DnaB unit al braç 5' d'un substrat en forma de forqueta. Així doncs, a una forqueta de replicació, l'hexàmer de DnaB es disposaria encerclant la cadena 5', amb els dominis N-terminal encarats cap a la regió de cadena senzilla i els dominis C-terminals encarats cap a la regió dúplex (Jezewska *et al.* 1998bb). Igualment, es determinà que en un complex estacionari de DnaB unida a un substrat en forma de forqueta, l'helicasa no envaeix la part dúplex de la forqueta més enllà dels dos o tres primers parells de bases. Així doncs, aquestes dades permeteren eliminar la possibilitat de que el ssDNA passes per l'exterior de l'hexàmer i es considerà provat que ho feia pel canal interior (Jezewska *et al.* 1998aa), tal i com ja s'havia suggerit per l'helicasa de T7, pertanyent a la mateixa família que DnaB, en base a estudis de microscòpia electrònica (Egelman *et al.* 1995). Aquests estudis també permeteren determinar que el lloc d'unió a ssDNA està format

per dues regions, cadascuna de les quals compren 10 nucleòtids (Jezewska *et al.* 1998aa; Jezewska *et al.* 1998bb).

4.3 Interacció amb substrats en forma de forqueta

DnaB s'uneix, en una reacció independent de la hidròlisi d'ATP, preferentment al braç 5' d'un substrat en forma de forqueta de replicació, amb una afinitat intrínseca sis vegades superior a la presentada pel braç 3'. La cadena senzilla 3' de la forqueta no forma cap complex estable amb l'hexàmer de DnaB que es troba associat a la monocadena 5', de manera que la cadena 3' roman en una conformació accessible a la unió d'un segon hexàmer de DnaB. A més, la part dúplex de la forqueta tampoc estableix interaccions significatives amb l'hexàmer de DnaB (Jezewska *et al.* 1997). Aquestes interaccions asimètriques de l'enzim amb un DNA en forqueta són consistents amb la polaritat de 5' a 3' de l'activitat helicasa de DnaB (LeBowitz *et al.* 1986).

En cas que es produeixi la unió d'un segon hexàmer de DnaB al braç 3' lliure de la forqueta, aquest segon hexàmer presenta una disposició oposada a la que pren el del braç 5', indicant que existeix una sola possible orientació de DnaB respecte a la polaritat de l'esquelet del ssDNA (Jezewska *et al.* 1998bb).

4.4 Activitat de desenrotllament de dúplex

Estudis d'activitat helicasa amb la DnaB de *Thermus aquaticus* indiquen que, per a un desenrotllament òptim, es necessita una cua de 6 a 10 bases a la cadena de ssDNA a 5' i una més llarga de 20 a 30 bases a la cadena 3' (Kaplan *et al.* 1999). Aquesta dependència de la llargada de la regió de cadena senzilla de la forqueta és diferent, pràcticament inversa, a la presentada per la gp4 de T7, en la qual uns 35 nucleòtids a la cua 5' i uns 10-15 a la cua 3' estimulen el desenrotllament (Hacker *et al.* 1997). Aquesta discrepància provocà el plantejament de la possibilitat de que la polaritat de la DnaB de *T. aquaticus* fos l'inversa a la descrita per gp4 i per la DnaB d'*E. coli* fins aquell moment. Aquesta hipòtesi es descartà al comprovar-se que revertir

la polaritat de la cua 3' o modificar-la químicament tenia poc efecte sobre el desenrotllament del DNA per part de la DnaB de *T. aquaticus*, mentre que revertir la polaritat o modificar químicament la cua 5' impedia el desenrotllament. Conseqüentment, aquests resultats demostraven que l'helicasa es translocava al llarg de la cadena 5' de la forqueta, excloent la cadena 3' del seu canal central (Kaplan 2000). Per tant, actualment es considera que DnaB desenrotlla el DNA dúplex gràcies a un mecanisme d'exclusió estèrica.

Gràcies a posteriors estudis d'activitat helicasa, es demostrà que la DnaB de *T. aquaticus* no és capaç de continuar desenrotllant substrats de DNA que presentin talls (*nicks*) a la cadena 5' o 3'. Aquesta impossibilitat tindria gran rellevància biològica, doncs impediria que l'activitat helicasa continués indiscriminadament *in vivo* en situacions en que la progressió del replisoma està aturada a la cadena líder o retardada degut a la incapacitat de l'holoenzim polimerasa III per superar forats a la cadena que està replicant. D'aquesta manera, es possibilitaria l'entrada de la maquinària de reparació de lesions del DNA i la posterior reiniciació de la replicació un cop arreglat el problema (Kaplan 2000).

4.4.1 Anàlisi quantitatiu de la reacció de desenrotllament catalitzada per la DnaB d'*E. coli*.

La mida del pas cinètic de la reacció de desenrotllament catalitzada per DnaB és de 1.4 pb. Donat que aquesta mesura pot estar lleugerament sobreestimada degut a que el lloc d'unió a ssDNA de l'hexàmer es troba dins del canal central i a possibles diferències entre les taxes de catàlisi i de translocació, es considera que DnaB desenrotlla 1 sol parell de bases a cada pas catalític (Galletto *et al.* 2004aa).

La taxa intrínseca de desenrotllament de l'hexàmer de DnaB a 25° C és ràpida, d'uns 290 pb per segon (Galletto *et al.* 2004aa). L'extrapolació d'aquest valor a la temperatura *in vivo*, és a dir 37° C, dóna una taxa aproximada de 708 pb/s, valor que concorda amb la velocitat estimada de desplaçament de la forqueta de replicació *in vivo* a les cèl·lules d'*E. coli* en creixement ràpid (LeBowitz *et al.* 1986; Baker *et al.* 1987). Per tant, l'activitat de desenrotllament intrínseca de DnaB explicaria les altes taxes de

desenrotllament observades tant *in vivo* com *in vitro*, sense que fos necessari un mecanisme addicional d'activació de l'activitat helicasa.

Ara bé, l'eficiència real de desenrotllament per part de DnaB està condicionada per la seva processivitat, és a dir per la probabilitat de que l'helicasa realitzi el proper cicle catalític en comptes de dissociar-se del seu substrat. Aquest valor és de 0.89 per a l'hexàmer de DnaB, cosa que indica que la processivitat d'aquesta reacció només és moderada (Galletto *et al.* 2004aa). En conseqüència, tot i tenir una taxa de desenrotllament ràpida, DnaB no és intrínscament capaç de desenrotllar grans distàncies de DNA de doble cadena i necessita l'ajuda de proteïnes accessòries per a dur a terme eficientment la seva funció com a helicasa durant la replicació. Efectivament, la processivitat de l'enzim *in vivo* augmenta gràcies a que les interaccions que DnaB estableix amb els altres membres del replisoma augmenten el temps de vida mitja del complex entre l'helicasa i el seu substrat de DNA.

També es caracteritzà l'efecte sobre la reacció de desenrotllament catalitzada per DnaB de la llargada del braç 3' del substrat de DNA en forqueta, així com de la composició de la regió dúplex a desenrotllar.

Es determinà que la llargada de la cadena senzilla 3' de la forqueta controla tant la taxa de desenrotllament com la processivitat de l'helicasa DnaB, doncs al augmentar la llargada d'aquest braç augmenta la taxa de desenrotllament i disminueix la taxa de dissociació de DnaB del DNA. Concretament, per a una llargària de 15 bases, la processivitat de l'helicasa disminueix a valors de 0.35, convertint-se així en un enzim distributiu. El fet que afecti tant a la velocitat com a la processivitat de desenrotllament i que existeixi una llargada òptima per a aquesta estimulació, corresponent a unes 20-25 bases, indica que aquest efecte es produeix gràcies a una interacció directe però transitòria entre l'helicasa i la cadena senzilla 3', interacció que es produeix a una distància específica de la regió dúplex. Igualment, el fet que s'observi un grau d'estimulació similar quan un segon hexàmer de DnaB està unit al braç en 3' indica que l'estimulació dependent del braç 3' és un procés no específic, que a l'interior de la cèl·lula pot ser substituït per la presència d'altres proteïnes unides a la cadena 5' exclosa

del canal central de l'hexàmer, les quals anàlogament interaccionarien i estimularien l'activitat de DnaB (Galletto *et al.* 2004bb).

Finalment, la taxa de desenrotllament del dsDNA per part de DnaB decreix linealment a mesura que augmenta el contingut en parells GC del dúplex a desenrotllar. Paral·lelament, una major estabilitat de la regió dúplex també afecta negativament, encara que en menor mesura, la processivitat de l'helicasa (Galletto *et al.* 2004bb). Recentment, s'ha descrit la mateixa relació entre la composició del dsDNA i la taxa de desenrotllament en estudis cinètics de molècula senzilla de l'helicasa gp4 de T7 (Johnson *et al.* 2007), suggerint que es tracta d'una característica comuna per als membres de SF4.

En base als resultats cinètics descrits en aquest apartat, així com als resultats bioquímics descrits als apartats anteriors, es proposa un mecanisme de desenrotllament del DNA de doble cadena catalitzat per l'hexàmer de DnaB mitjançant un mecanisme d'exclusió estèrica. A la figura I.9 es representa esquemàticament aquest mecanisme.

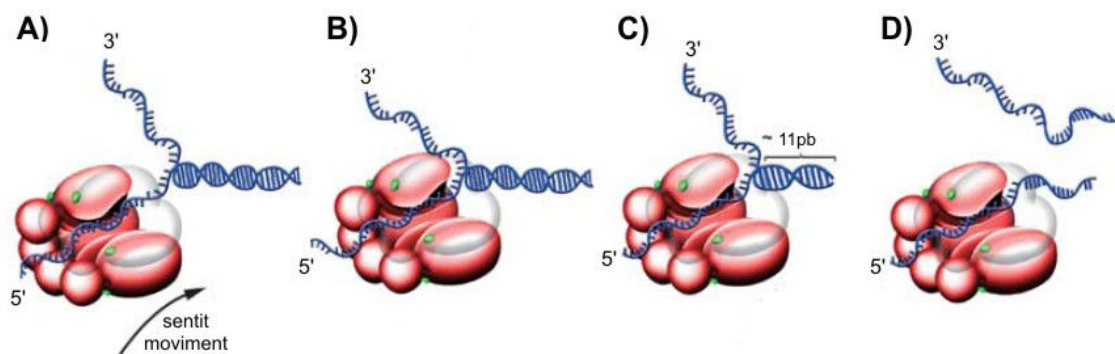


Figura I.9: Model de desenrotllament de dsDNA basat en estudis cinètics de DnaB.

A) L'helicasa forma un complex estacionari amb el braç 5' d'un substrat en forqueta, sense interaccionar amb el braç 3' ni amb la regió dúplex. **B)** Transició de l'helicasa cap a un estat productiu al iniciar-se la hidròlisi d'ATP. En l'estat productiu, l'helicasa té un lloc d'unió transitori al ssDNA del braç 3', el qual controla la taxa de desenrotllament i la processivitat de la reacció. **C)** DnaB desenrotllar processivament el dsDNA, una base per a cada pas catalític, interactuant transitòriament amb el braç 3' a cada pas. La reacció de desenrotllament catalítica continua fins a un punt en que la llargada del fragment dúplex que queda per desenrotllar es massa petita per preservar la seva integritat dins d'un substrat en forqueta, aquest valor per DnaB a 20°C és d'uns 11 pb. **D)** El ssDNA deixa el complex sense participació catalítica de l'helicasa. Figura adaptada de Galletto *et al.*, 2004a.

4.5 Translocació al llarg de dsDNA

Els experiments de desenrotllament de múltiples substrats en forqueta modificats químicament per part de la DnaB de *T. aquaticus* permeteren observar que DnaB és capaç de translocar-se al llarg de DNA de doble cadena sense desenrotllar-lo i que és la longitud de la cua 3' allò que determina estèricament si una o dues cadenes de DNA passen a través del canal central de l'hexàmer. Si la cua 3' del substrat en forqueta és massa curta, o inexistent, quan l'hexàmer de DnaB que s'ha carregat a la cadena en 5' arriba a la zona dúplex, ambdues cadenes de DNA passen a través del canal central de DnaB, a mesura que l'enzim continua movent-se en sentit 5' a 3'. Aquesta translocació al llarg de dsDNA no produeix separació de cadena del dúplex i, per tant, en aquesta situació DnaB actuaria com a translocasa de dsDNA (Kaplan 2000). Posteriors estudis amb la DnaB d'*E. coli* confirmaren que aquesta també funciona com a translocasa de dsDNA en absència d'una cua 3' (Kaplan *et al.* 2002).

Estudis amb la DnaB d' *E. coli* mostraren que la translocació de l'hexàmer al llarg del DNA de doble cadena és un procés actiu, doncs genera energia suficient per desplaçar proteïnes fortament unides al DNA. A més, DnaB acobla aquesta força a la conducció de la migració de branca en una *Holliday junction* sintètica amb braços heteròlegs. Durant aquest procés, DnaB encercla les dues cadenes de DNA però s'uneix només a una d'elles. La translocació activa en el sentit 5' a 3' al llarg de la cadena a la qual s'uneix provoca la migració de la branca, gràcies a que la segona cadena també es estirada cap al canal central de l'hexàmer (Kaplan *et al.* 2002). Aquesta nova capacitat de DnaB indueix a especular que l'helicasa podria estar directament involucrada en processos de recombinació *in vivo*. Per exemple, les dues principals rutes de reparació per recombinació d'una forqueta encallada degut a una lesió al DNA tenen com a substrats intermediaris de DNA una estructura en *Holliday junction*. Per tant, a part de les proteïnes específiques per a aquesta tasca, com ara RuvAB i RecG, DnaB també podria catalitzar aquesta reacció. Això explicaria alguns fenotips de mutants de DnaB, els quals mostren propensió a fenòmens de recombinació il·legítima (Kaplan *et al.* 2002; Bujalowski 2003).

5. Estructures atòmiques de DnaB

Prèviament a la iniciació d'aquesta tesi doctoral, únicament s'havia descrit l'estructura atòmica del domini N-terminal de la DnaB d'*E. coli* (Fass *et al.* 1999; Weigelt *et al.* 1999).

A la part final d'aquest treball, però, es va publicar l'estructura del monòmer de la DnaB de *Thermus aquaticus*, així com l'hexàmer de la DnaB de *Bacillus stearothermophilus* i de la proteïna GP40, l'homòleg de DnaB del fag SPP1 de *Bacillus sp* (Bailey *et al.* 2007b; Bailey *et al.* 2007a; Wang *et al.* 2008).

A continuació es descriuen breument aquestes estructures.

5.1 Domini N-terminal de la DnaB d'*E. coli*

L'estructura del domini N-terminal de la DnaB d'*E. coli* es va resoldre per cristal·lografia de raigs X, concretament el fragment comprès entre els residus 15 i 128, (Fass *et al.* 1999) i per ressonància magnètica nuclear, en aquest cas el fragment corresponent als residus del 24 al 136 (Weigelt *et al.* 1999).

L'estructura cristal·lina del domini N-terminal de DnaB a 2.3 Å revelava un plegament globular i totalment helicoïdal formant per sis hèlixs α , cinc de les quals ($\alpha 2$ - $\alpha 3$, $\alpha 4$, $\alpha 5$ i $\alpha 6$) s'emboïllaven al voltant d'una hèlix central ($\alpha 1$). El cristall pertanyia al grup espacial $P2_1$ i presentava quatre molècules per unitat asimètrica disposades en dos dímers idèntics. Les interfícies de contacte entre dímers estaven formades per dues hèlixs antiparal·leles, $\alpha 4$ i la $\alpha 6$, de manera que es generava un llaç de 4 hèlixs a la zona d'interacció (Fass *et al.* 1999). La figura I.10 mostra el dímer cristal·logràfic del domini N-terminal de DnaB.

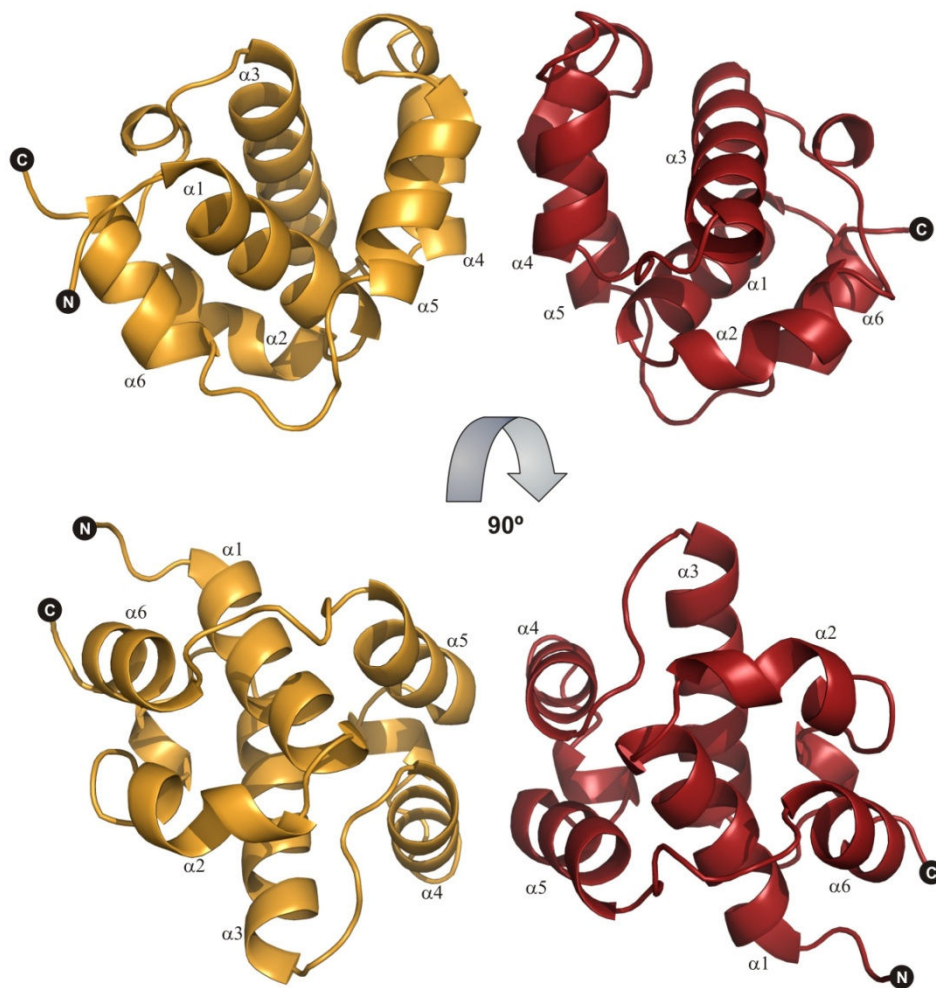


Figura 1.10: Dímer cristal·logràfic del domini N-terminal de la DnaB d'*E. coli*. Cada monòmer es representa en un color diferent. Es mostren dues vistes de la superfície d'interacció entre monòmers: perpendicular a l'eix de simetria dos (dalt) i al llarg d'aquest eix (baix). Codi PDB: 1B79.

Paral·lelament, es va resoldre per ressonància magnètica nuclear l'estructura en solució d'un fragment lleugerament més llarg del domini N-terminal també de la DnaB d'*E. coli*, el qual presentava el mateix el plegament descrit anteriorment. En aquest cas, es va observar que en solució hi havia un equilibri entre monòmer i dímer, i es proposà una superfície d'interacció entre dos dominis N-terminals molt similar a la descrita anteriorment per l'estructura resolta per cristal·lografia de raigs X, doncs implicava els mateixos residus (Weigelt *et al.* 1999).

En funció d'aquestes dues estructures, es va hipotitzar que els dímers observats pels dominis N-terminals de DnaB serien rellevants biològicament i els responsables de

la transició entre l'estat de simetria C3 i C6 observat per microscòpia electrònica (Fass *et al.* 1999; Weigelt *et al.* 1999).

5.2 Monòmer de la DnaB de *Thermus aquaticus*

El monòmer de DnaB de *Thermus aquaticus*, resolt a una resolució de 2.9 Å (Bailey *et al.* 2007a), mostra el plegament tipus RecA (fulla β central de 9 cadenes paral·leles, flanquejada per 3 hèlixs α a cada banda) esperat pel domini C-terminal de la proteïna i el plegament globular en hèlixs α ja observat pel domini N-terminal. A la figura I.11 es mostra l'estructura d'aquest monòmer.

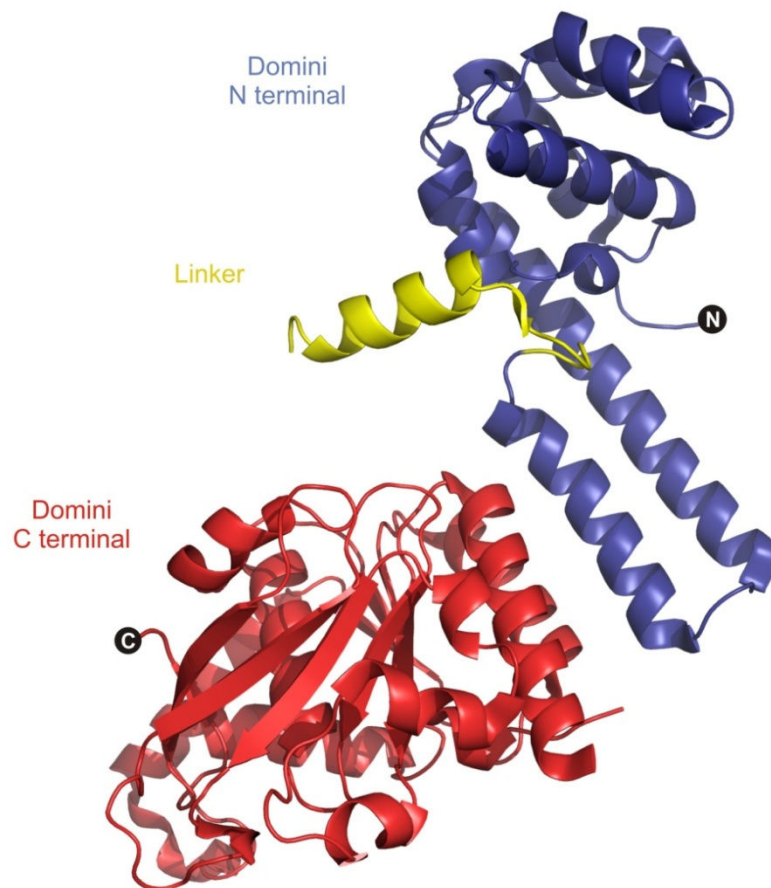


Figura I.11: Representació en cintes del monòmer de la DnaB de *T. aquaticus*. Cada domini es representa en un color diferent: en blau el domini N-terminal, el groc la regió d'enllaç i en vermell el domini C-terminal. Codi PDB: 2Q6T.

Ara bé, aquesta estructura permet redefinir les fronteres entre els dominis de DnaB, doncs s'observa que en realitat el domini N-terminal és 40 residus més llarg del que s'havia descrit anteriorment basant-se en els fragments proteolítics de l'helicasa. A l'estructura del monòmer, l'hèlix $\alpha 6$ s'allarga i interacciona amb una hèlix addicional ($\alpha 7$), formant una estructura en α -hairpin que constitueix un nou subdomini estès clarament diferenciat del subdomini N-terminal globular prèviament caracteritzat. A la figura I.12 es mostra aquesta nova definició del llindar del domini N-terminal de DnaB, comparant-lo amb el descrit anteriorment.

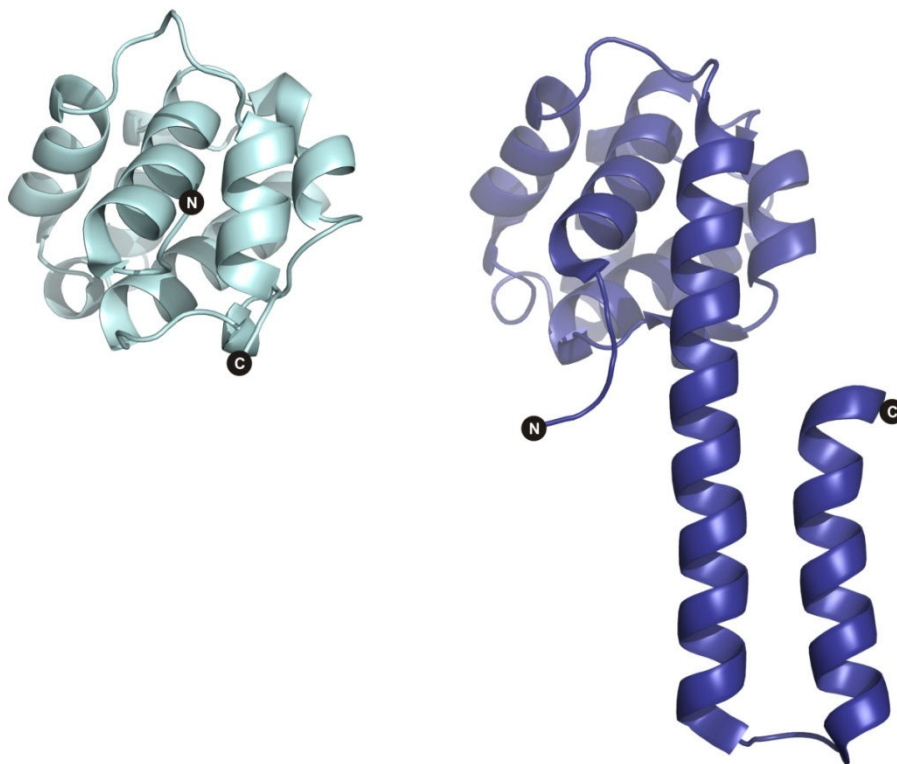


Figura I.12: Nova definició del domini N-terminal de DnaB. A l'esquerra, en blau clar, es representa el domini N-terminal de la DnaB d'*E. coli*, que havia estat definit en funció de la digestió proteolítica de la proteïna sencera (codi PDB: 1B79). A la dreta, en blau fosc, es representa el domini N-terminal de la DnaB de *T. aquaticus* en el context d'un monòmer de DnaB sencer (codi PDB: 2Q6T).

La DnaB de *T. aquaticus* cristal·litzà en el grup espacial $P3_121$, amb quatre monòmers per unitat asimètrica. Cadascun d'aquests monòmers presenta el domini N-terminal en una orientació diferent respecte al domini C-terminal, fet que indica que la regió d'enllaç permet una gran varietat de gir entre ambdós dominis i que, en absència

de les restriccions estèriques imposades per l'organització quaternària en hexàmer, el monòmer de DnaB té una estructura altament modular i flexible.

Paral·lelament, al cristall es poden observar dos dímers independents que s'estableixen mitjançant la interacció de dos dominis N-terminals de la proteïna. La superfície d'interacció d'aquests dímers és diferent de la descrita anteriorment (Fass *et al.* 1999) i implica el subdomini en α -hairpin que s'ha descrit de nou en aquesta estructura. A la figura I.13 es mostra aquesta interacció entre dominis N-terminals.

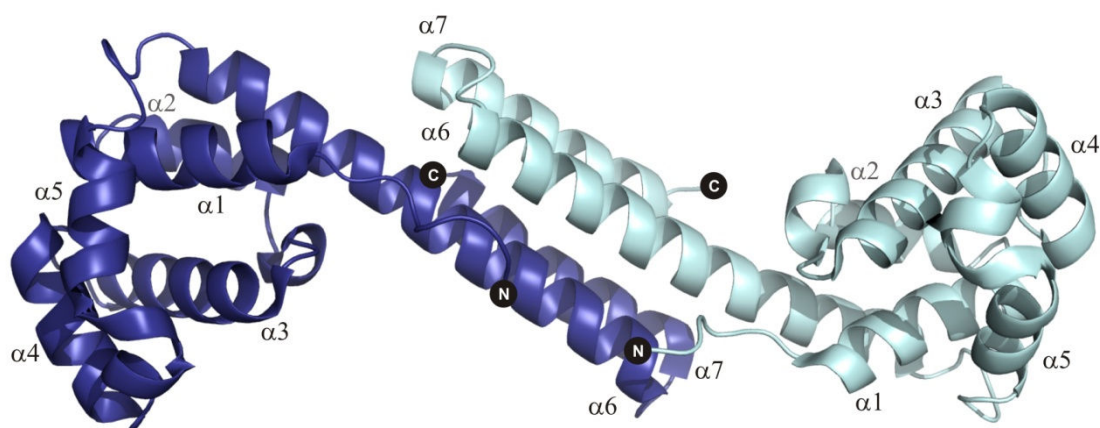


Figura I.13: Representació en cintes del dímer format entre dos dominis N-terminals de la DnaB de *T. aquaticus* mitjançant els seus subdominis α -hairpin. Cada monòmer es representa en un color diferent. Codi PDB: 2Q6T

5.3 Hexàmer de la DnaB de *Bacillus stearothermophilus* i de G40P del fag SPP1 de *Bacillus sp*

Ambdues estructures mostren un hexàmer format per dues capes diferenciades. Els dominis N-terminals formen un collar rígid amb una simetria 3 sobre la capa formada pels dominis C-terminals, que tenen una disposició més irregular i heterogènia (Bailey *et al.* 2007b; Wang *et al.* 2008). La disposició dels diferents monòmers de C-terminal varia àmpliament tant a les diferents formes cristal·lines com entre els dos homòlegs estudiats, suggerint una gran flexibilitat d'aquesta capa. Per contra, la capa N-terminal sempre té una simetria 3 rígida.

Sorprenentment, l'amplada de la capa N-terminal és superior a la C-terminal, i això provocà una assignació incorrecte dels dominis en les reconstruccions tridimensionals de microscòpia electrònica de la DnaB d'*E. coli* (Yang *et al.* 2002) i de GP40 (Nunez-Ramirez *et al.* 2006). A més, les dimensions globals tampoc coincideixen. Ara bé, tant la forma com el volum de les estructures hexamèriques de DnaB sí que són compatibles amb les projeccions en dos dimensions usades per a aquestes reconstruccions. Per tant, sembla probable que les diferències entre les dades cristal·logràfiques i les reconstruccions de criomicroscòpia siguin degudes a la compressió pròpia de la tècnica o a les dificultats metodològiques per generar models de microscòpia tridimensionals de molècules tant flexibles com DnaB.

5.3.1 Hexàmer de DnaB de *Bacillus stearothermophilus*

De la DnaB de *B. stearothermophilus* es van resoldre quatre formes cristal·lines diferents, dues de la proteïna sola (formes B1 i B2) i dues de DnaB en complex amb el domini d'interacció amb l'helicasa (HBD) de la proteïna DnaG (formes BH1 i BH2) (Bailey *et al.* 2007b). L'estructura amb una millor resolució, 2.9 Å, correspon a la forma BH1, que cristal·litzà en el grup espacial P321 amb un dímer de DnaB unit a una molècula de HBD per unitat asimètrica. Per simetria cristal·logràfica es genera l'hexàmer de DnaB en complex amb 3 HBD de DnaG, els quals interaccionen exclusivament amb el domini N-terminal de l'helicasa. L'estructura de DnaB sola amb millor resolució, 3.7 Å, és la B1, pertanyent al grup espacial P2₁2₁2 i amb un hexàmer per unitat asimètrica. Les quatre formes cristal·lines difereixen només en la orientació relativa dels seus dominis C-terminals. A la figura I.14 es mostra les vistes laterals i frontals de la forma B1 de la DnaB de *B. stearothermophilus*.

Els dominis N-terminals sempre adopten una disposició en trímer de dímers. Els monòmers N-terminals adjacents adopten dues conformacions diferents: bé col·loquen el subdomini α -hairpin sobre el seu propi domini C-terminal o sobre el domini C-terminal de la subunitat veïna. Aquestes dues conformacions determinen la formació d'una disposició en trímer de dímers, en els quals es poden diferenciar dos superfícies d'interacció. La primera, més extensa i segurament la més rellevant, s'estableix entre dos subdominis α -hairpin; mentre que la segona, que involucra menys residus, es forma

a través dels subdominis globular. El collar N-terminal podria proporcionar un lloc d'unió a ssDNA addicional, doncs el seu interior presenta tres regions diferenciats de potencial electrostàtic positiu, separades per regions de potencial electrostàtic negatiu. Aquestes regions positives contenen residus conservats a totes les DnaB bacterianes.

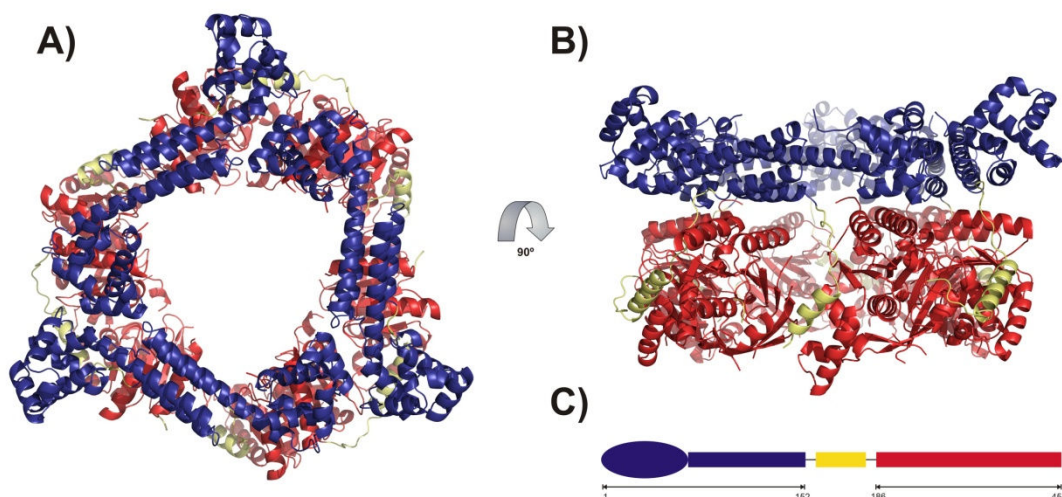


Figura I.14: Representació en cintes de l'hexàmer de la DnaB de *B. Stearothermophilus*. Es representa la forma cristal·lina B1 (codi PDB: 2RED) en vista frontal (A) i lateral (B). C) esquema de l'organització en dominis: N-terminal (blau), *linker* (groc) i C-terminal (vermell).

Els dominis C-terminals de les diferents formes cristal·lines de la DnaB de *B. Stearothermophilus* adopten una gran varietat d'orientacions al llarg de l'anell, encara que sempre interaccionen amb l'hèlix de la zona d'enllaç del monòmer adjacent a la perifèria de l'hexàmer i orienten els llaços d'unió a DNA putatius cap a l'interior del canal central. A la forma B1, els dominis C-terminals formen un anell irregular amb cap mena de simetria rotacional. La capa C-terminal de la resta de formes cristal·lines adopta una simetria tres aproximada. Els dominis C-terminals es mantenen units entre ells principalment gràcies a les seves interaccions amb la zona del *linker* i no per interaccions entre dominis C-terminals adjacents. Ara bé, la principal força estabilitzadora de la capa C-terminal són les fortes interaccions que s'estableixen entre els dímers de la capa N-terminal.

El diàmetre exterior global del hexàmer és de 115 Å i l'alçada de 75 Å. El diàmetre del canal central a través dels dominis N-terminals és d'uns 50 Å, mentre que

les diferents orientacions dels anells de dominis C-terminals donen com a resultat diàmetres que varien entre 25 i 50 Å. Per tant, en tots els casos es tracte de diàmetres suficientment grans per acomodar DNA dúplex.

5.3.2 Hexàmer de la G40P del fag SPP1 de *Bacillus sp*

G40P cristal·litzà en el grup espacial $P2_12_12_1$, amb un hexàmer complet a la unitat asimètrica. L'estructura a 3.9 Å mostra la capa N-terminal amb simetria $C3$ i la capa C-terminal propera a la simetria $C6$ (figura I.15). El diàmetre interior de la capa N-terminal de l'hexàmer de G40P és de ≈ 42 Å i el de la capa C-terminal de ≈ 17 Å (Wang *et al.* 2008).

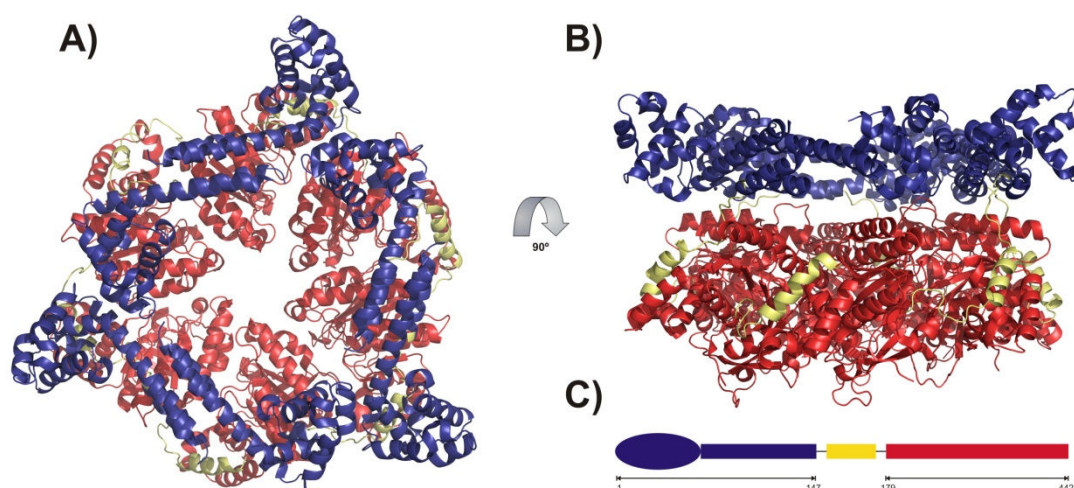


Figura I.15: Representació en cintes de l'hexàmer de G40P del bacteriòfag SPP1.

Es representa la vista frontal (A) i la lateral (B). C) Esquema de l'organització en dominis: N-terminal (blau), *linker* (groc) i C-terminal (vermell). Codi PDB: 6RT1

Tal i com s'ha descrit per la DnaB de *B. stearothermophilus*, es poden diferenciar dos dímers del domini N-terminal, aquells que interaccionen a través del subdomini α -hairpin i els que ho fan a través del subdomini globular. A la figura I.16 es mostren aquests dos tipus d'interaccions. Igualment, la regió del *linker* d'una subunitat interacciona amb el domini C-terminal adjacent, presentant diferents angles que demostren la seva capacitat per a actuar facilitant els canvis conformacionals a l'hexàmer.

Al mateix treball, també es va resoldre el monòmer del domini C-terminal de G40P a 2.3 Å i es varen realitzar estudis de funcionalitat amb diferents mutants de delecció del domini N-terminal. Es determinà que la delecció del subdomini globular N-terminal redueix l'activitat helicasa i que la supressió addicional del subdomini α -hairpin la elimina completament. Aquests mutants de delecció conserven gran part de l'activitat ATPasa, cosa que suggereix que els canvis en la integritat estructural de la capa N-terminal de simetria 3 afecten més a la funció helicasa que a l'activitat ATPasa.

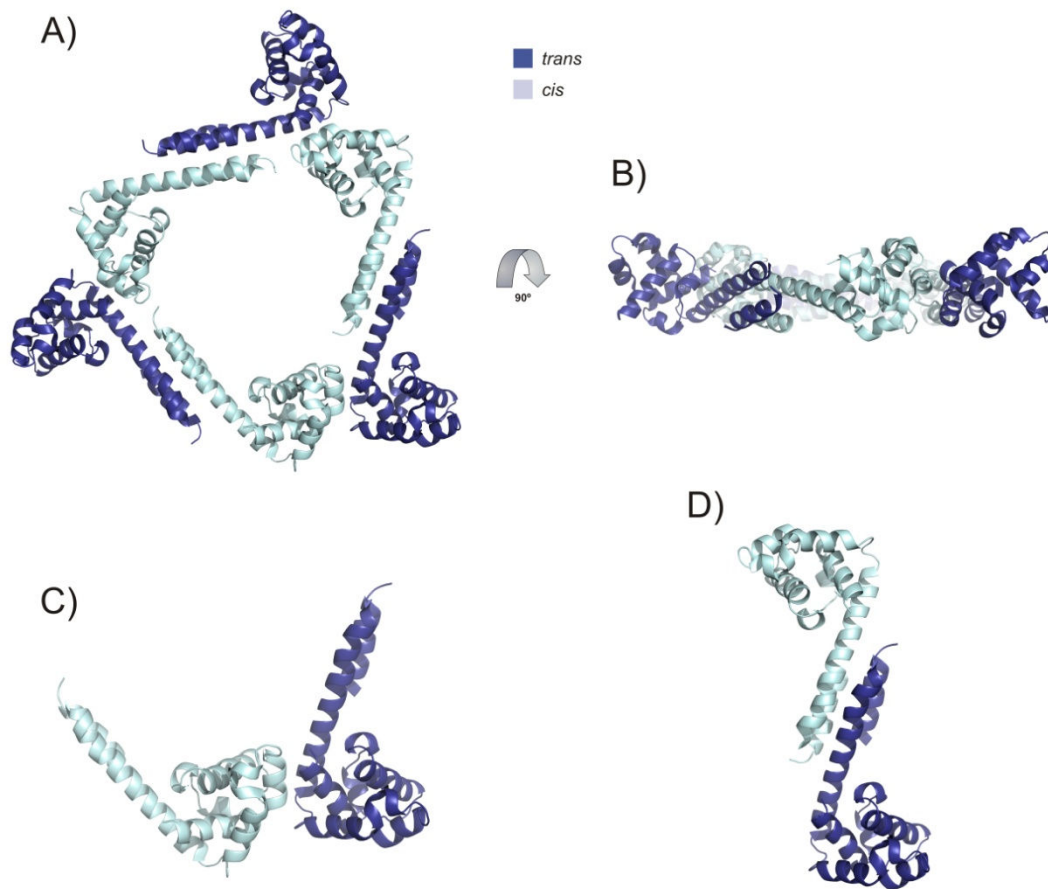


Figura I.16: Representació en cintes de la capa N-terminal de G40P i els dos tipus de dímers que es formen. Els monòmers alternats es representen en blau clar i blau fosc. Codi PDB: 6RT1. **A)** Vista frontal del trímer de dímers. **B)** Vista lateral de la capa N-terminal. **C)** Dímer format mitjançant interacció entre els subdominis globulars, que amaga una àrea superficial de 922 \AA^2 amb 19 residus formant enllaços entre ells. **D)** Dímer format mitjançant els subdominis α -hairpin. Aquesta interfície compren una àrea de 1781 \AA^2 i inclou interaccions extensives entre 32 residus.

III . FUNCIONS DE DNAB A LA CÈL·LULA BACTERIANA

La replicació del DNA genòmic és un procés clau en la divisió i la propagació cel·lular. DnaB participa en l'etapa d'iniciació i d'elongació de la replicació bacteriana, així com també en la reiniciació de la replicació després de la reparació d'errors al DNA. Juga un paper rellevant no només gràcies a la seva activitat principal de desenrotllament de la doble hèlix de DNA, sinó també com a element organitzador de la maquinària replicativa gràcies a les múltiples interaccions que efectua amb altres proteïnes replicatives, regulant-ne la funció.

A continuació es descriuen les etapes d'iniciació i elongació de la replicació del DNA a bacteris, posant especial èmfasi en la funció de DnaB.

1. Iniciació de la replicació bacteriana

L'objectiu de l'etapa d'iniciació és garantir un accés controlat al DNA dúplex per originar una regió de DNA desenrotllada que permeti l'assemblatge de la maquinària de replicació que propagarà la forqueta de replicació. Es tracta d'un procés altament regulat, doncs cal que es produeixi durant la fase correcta del cicle cel·lular, només una sola vegada per cicle i que no es generin mutacions al genoma. Tal i com ja proposava originàriament el model del replicó (Jacob *et al.* 1963), el control de la iniciació està regulat bàsicament per dos elements: el replicador i l'iniciador.

El replicador és el lloc del cromosoma on comença i és regulada la replicació del DNA. L'iniciador és el factor proteic que inicia aquest procés gràcies a la seva interacció específica amb el replicador. Així doncs, les proteïnes iniciadores són les encarregades de reconèixer l'origen de replicació i activar l'etapa d'iniciació en el moment precís del cicle cel·lular. Tot i que el seu mode d'acció i regulació varia als diferents dominis de la vida, totes tenen característiques arquitectòniques comunes que posen en evidència la seva rellevància per a la viabilitat cel·lular. Totes les proteïnes iniciadores comparteixen un plegament d'unió a nucleòtid d'adenina del tipus AAA+ associat a un domini d'unió a DNA que ajuda a aquestes proteïnes puguin localitzar-se específicament a l'origen de replicació (Duderstadt *et al.* 2008).

Als bacteris l'iniciador és la proteïna DnaA i el replicador és l'origen cromosòmic *oriC*, el qual funciona com a lloc per a l'assemblatge de la maquinària replicativa cel·lular (Kaguni 2006; Mott *et al.* 2007).

1.1 La proteïna iniciadora: DnaA

DnaA té quatre dominis, que han estat descrits mitjançant tècniques bioquímiques, homologia de seqüència i estructuralment (Kaguni 2006; Mott *et al.* 2007; Duderstadt *et al.* 2008). El domini I, situat a la part N-terminal de la proteïna, conté el lloc d'unió amb la proteïna DnaB (l'helicasa replicativa bacteriana) i també està

involucrat en facilitar la interacció entre diferents molècules de DnaA per formar oligòmers. El domini II és una regió d'enllaç flexible, no vital per a la viabilitat cel·lular i poc conservada entre les diferents espècies bacterianes, on presenta llargàries àmpliament variables. El domini III conté el mòdul conservat AAA+, on es localitza l'activitat ATPasa de la proteïna així com el determinant d'oligomerització primari de DnaA. L'estat d'unió a nucleòtid d'aquest domini regula la capacitat de DnaA per formar autoassemblatges. Dins del domini III es poden diferenciar dos subdominis (IIIa i IIIb) anomenats, respectivament, base i tapa. Finalment, a la regió C-terminal de la proteïna es localitza el domini IV, responsable de la unió a DNA a través del reconeixement de les seqüències d'unió específiques de DnaA. Presenta un plegament del tipus hèlix-gir-hèlix (*helix-turn-helix*). Aquest domini, que està molt conservat, és el responsable de reclutar DnaA específicament a l'origen.

Actualment, es desconeix l'estructura tridimensional de la proteïna sencera, però sí que s'han determinat mitjançant cristal·lografia de raigs X el domini IV de la DnaA d'*E. coli* unit a la seva seqüència consens de la caixa DnaA (Fujikawa *et al.* 2003); i els dominis III i IV de la DnaA d'*Aquifex aeolicus* units a ADP (Erzberger *et al.* 2002) i a un anàleg no hidrolitzable d'ATP, l'AMP-PCP (Erzberger *et al.* 2006b). Addicionalment, l'estructura en solució del domini I ha estat resolta per ressonància magnètica nuclear (Abe *et al.* 2007).

El constructe que conté els dominis III i IV de la DnaA d'*A. aeolicus* cristal·litza com a monòmer quan està unit a ADP (Erzberger *et al.* 2002), però la unió d'anàlegs d'ATP promou l'empaquetament de molècules de DnaA adjacents a través d'una extensa superfície d'oligomerització. Així doncs, l'ATP fa d'interruptor molecular que reposiciona els motius específics del domini AAA+ per permetre una oligomerització productiva de DnaA. A la figura I.17 s'esquematitzen els canvis conformacionals que permeten l'autoensamblatge de la proteïna iniciadora.

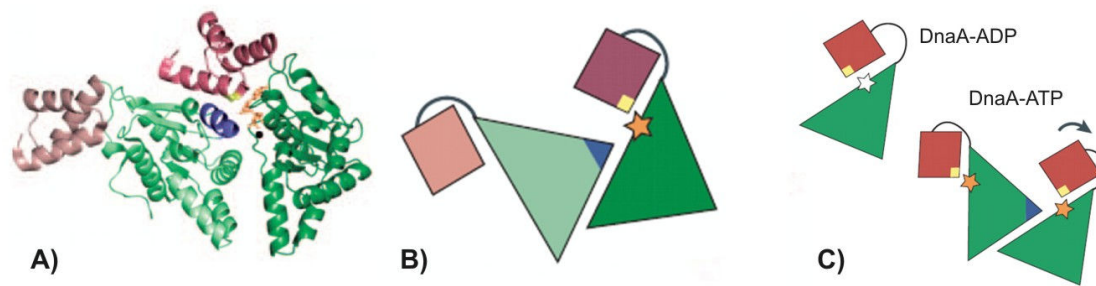


Figura I.17: Paper de la unió d'ATP en l'oligomerització de DnaA. **A)** Representació en cintes del dímer de DnaA en presència d'ATP. En verd es mostren els subdominis AAA⁺ i en lila els subdominis tapa α . Les regions que conformen la butxaca nucleotídica bipartida es marquen en rosa a una subunitat i en blau a l'altra. **B)** Representació esquemàtica del dímer mostrat a (A). **C)** Esquema de la transició entre els estats de DnaA units a ADP i a ATP. El canvi conformacional provocat per la presència del fosfat γ de l'ATP exposa una regió del centre actiu que permet la unió d'una segona molècula de DnaA-ATP, la qual, al seu torn, contribueix a la butxaca d'unió a NTP de la primera subunitat. En l'estat DnaA-ADP, la tapa α bloqueja estèricament aquest contacte intersubunitat, consistent amb les dades bioquímiques que mostren que l'ADP no afavoreix l'oligomerització de DnaA. Figura adaptada de Mott *et al.*, 2007.

DnaA-ATP oligomeritza formant un filament helicoïdal estès de mà dreta, no en forma d'anell tancat com es habitual dins la superfamília AAA+. El determinant estructural d'aquest ordenament característic és una inserció única present al domini III de DnaA, que classifica l'iniciador dins d'un clade independent dintre de la superfamília AAA+ (Koonin 1992). Aquesta inserció forma una cunya d'hèlix α que es projecta del core AAA+ de l'oligòmer i actua a la interfície entre subunitats, orientant els mòduls AAA+ consecutius en un arranament no planar (Erzberger *et al.* 2006b).

1.2 L'origen de replicació bacterià: *oriC*

El cromosoma bacterià típicament conté un sol origen de replicació anomenat *oriC*. Les seqüències de l'origen bacterianes varien entre espècies en la llargada total d'*oriC*, així com en la disposició i el nombre total de llocs d'unió a DnaA. Tot i aquesta variabilitat, tots els orígens bacterians contenen llocs d'unió específics per a DnaA i la regió DUE, que es fon per unió cooperativa de molècules de DnaA.

L'origen bacterià més ben caracteritzat és el d'*E. coli*, doncs ha servit com a model durant anys per estudiar l'etapa d'iniciació de la replicació. A la figura I.18 es mostra un esquema d'aquest origen. L'*oriC* d'*E. coli* compren una regió d'uns 250 pb de llargària que conté varies caixes DnaA, que són llocs d'unió específics de seqüència de 9 parells de bases per a la proteïna iniciadora DnaA. S'anomenen de R1 a R5. Addicionalment, DnaA també s'uneix a una altra classe seqüències, anomenades llocs I, que difereixen subtilment de l'element consens de la caixa DnaA i es troben intercalats entre les caixes DnaA a *oriC*. Adjacent a aquest seguit de caixes DnaA i llocs I, hi ha un element de desenrotllament del DNA ric en AT anomenat DUE (sigles de l'anglès *DNA unwinding elements*) format per tres repeticions de 13 pb. Aquesta regió conté una tercera classe de seqüències d'unió a DnaA, les caixes DnaA-ATP, que són similars als llocs I en el fet que són units selectivament per DnaA només en presència d'ATP. Els posicionament relatiu dels diferents llocs d'unió de DnaA entre ells i respecte al DUE és molt important per a una correcta activació de l'origen així com per un *timing* adequat del procés d'iniciació de la replicació.

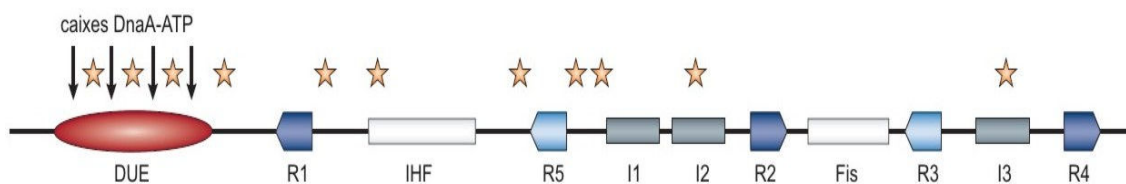


Figura I.18: Organització de l'*oriC* d'*E. coli*. La regió del DUE es mostra en vermell. Les caixes DnaA d'alta afinitat (R1, R2 i R4) es representen en blau fosc i les de baixa afinitat (R3 i R5) en blau clar. L'orientació de les caixes DnaA correspon a la cadena superior de DNA. També es mostra la localització dels llocs I (gris), les caixes DnaA-ATP (fletxes) i els llocs d'unió de les proteïnes IHF i Fis (blanc). Les estrelles grogues representen les dianes de metilació presents a l'origen.

Tal i com s'ha apuntat anteriorment, DnaA presenta una afinitat diferencial per als diferents tipus de seqüències d'unió presents a *oriC*. Roman unida a les caixes DnaA R1, R2 i R4 durant la major part del cicle cel·lular, tant en la seva forma unida a ADP com a ATP. Per contra, presenta una afinitat intermèdia pel llocs R5, R3, I1, I2 i I3 i encara més baixa pels llocs DnaA-ATP.

A *oriC* també hi ha llocs d'unió per a altres proteïnes implicades en la iniciació i la seva regulació. Els factors arquitectònics del nucleoid IHF (*integration host factor*) i Fis (*factor for inversion of stimulation*) tenen llocs d'unió a *oriC*. Ambdues proteïnes regulen la interacció de DnaA amb els seus llocs d'unió d'afinitat baixa i moderada durant l'etapa d'iniciació. IHF afavoreix la unió de DnaA a aquests llocs, cosa que ajuda a promoure la formació del complex oligomèric de DnaA i, consegüentment, al desenrotllament del DUE. Per contra, Fis impedeix la unió de IHF i disminueix la unió de DnaA als seus llocs més febles, cosa que impedeix el seu assemblatge i, consegüentment, la fusió del DUE.

1.3 Procés d'iniciació

Les interaccions de DnaA amb els seus diferents llocs d'unió a *oriC* defineixen la progressió de l'etapa d'iniciació de la replicació del DNA. La interacció d'alta afinitat amb les caixes de DnaA fortes localitza DnaA a l'origen i les interaccions dependents d'ATP i cooperatives amb els llocs de més baixa afinitat permeten que l'iniciador formi un gran complex oligomèric nucleoproteic, visible per microscòpia electrònica. Conjuntament amb el DNA superenrotllat negativament i els factors arquitectònics de nucleoid, l'assemblatge de DnaA directament facilita la fusió del DNA dins del DUE, generant les cadenes senzilles necessàries pel carregament de l'helicasa i el posterior assemblatge de la maquinària replisomal.

Entre les diferents etapes d'iniciació, Fis roman unida a *oriC* juntament amb molècules de DnaA a les caixes R1, R2 i R4. La concentració cel·lular de DnaA roman pràcticament constant al llarg de tot el cicle cel·lular. En canvi, la concentració de DnaA unida a ATP augmenta molt just abans de la iniciació de la replicació.

Durant la iniciació, Fis és desplaçada d'*oriC*, de manera que allibera la resta de llocs d'unió de DnaA. Això també possibilita la unió de IHF, que augmenta la força de la interacció de DnaA amb els seus llocs de baixa afinitat, especialment pels llocs I (anomenats així degut a la seva resposta dependent de IHF). Quan l'origen es satura de DnaA-ATP, es produeix la separació de cadena del DUE.

L'arquitectura del complex nucleoproteic format per DnaA i *oriC* totalment assemblat i el mecanisme mitjançant el qual es fon la regió del DUE encara no estan completament caracteritzats. A continuació es descriu la hipòtesi de treball actual, que es basa en les dades bioquímiques i estructurals de que es disposa fins al moment (Mott *et al.* 2007; Duderstadt *et al.* 2008).

Tot i que l'estructura del filament helicoïdal format per DnaA unida a ATP es va obtenir en absència de DNA, estudis previs indiquen que el DNA s'embolcalla al voltant de la cara externa del complex nucleoproteic (Funnell *et al.* 1987). Si és així, l'organització en mà dreta del filament de DnaA-ATP imposaria una clara direccionalitat a l'embolcallament del DNA al voltant seu. D'aquesta manera s'estabilitzaria un superenrotllament de DNA positiu al voltant d'un *core* proteic central. El *footprinting* topològic de la DnaA d'*A. aeolicus* ha mostrat que es generen superenrotllaments de DNA positius en presència d'anàlegs d'ATP, però no d'ADP, donant suport a aquesta hipòtesi. En un sistema topològic tancat, com es el cas del nucleoid bacterià, s'esperaria que sorgís una torsió negativa compensatòria en resposta a aquest embolcallament toroïdal positiu. Aquest estrés superhelicoïdal negatiu, en comptes de ser difús al llarg de tota la regió de l'*oriC*, estaria restringit i focalitzat a la zona del DUE adjacent, facilitant així el desenrotllament d'aquest element de DNA ric en parells AT i, per tant, ja intrínsicament inestable.

A la figura I.19 es representen esquemàticament les funcions que realitza la proteïna DnaA en l'etapa d'iniciació bacteriana, així com en la progressió d'aquesta fins a la formació d'un complex replicatiu obert, apte per a la progressió cap a l'etapa d'elongació de la replicació.

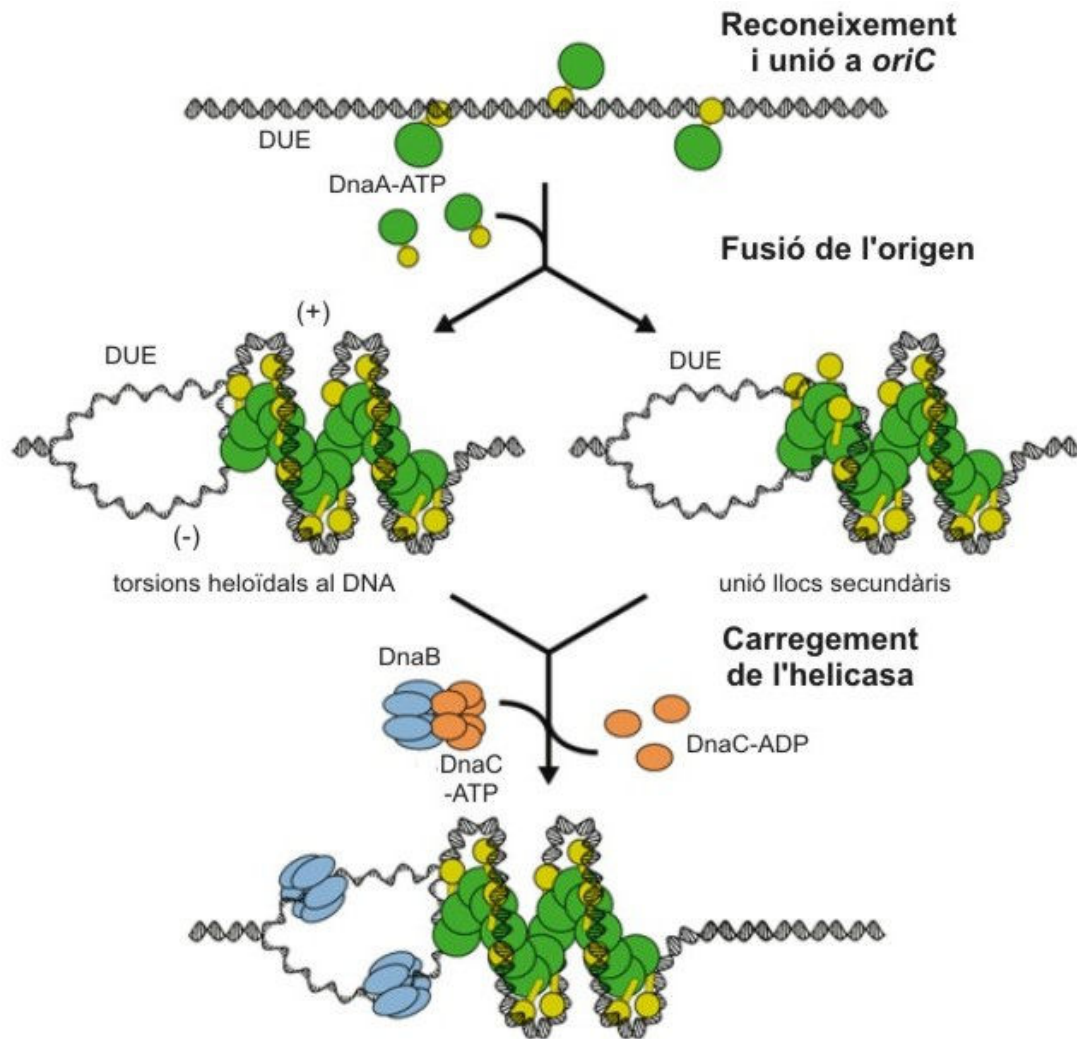


Figura I.19: Model de la iniciació de la replicació en bacteris. Els monòmers de DnaA es representen amb el domini d'unió a DNA en groc i els dominis I i III en verd. **Dalt:** Els monòmers de DnaA s'associen a *oriC* a través d'un reconeixement específic entre el seu domini IV i els llocs d'unió a DnaA presents a l'origen. **Centre:** En presència d'ATP, DnaA oligomeritza formant un filament helicoïdal al voltant del qual s'embolica el DNA. Això provoca la fusió del DUE gràcies a la formació de retorciments negatius compensatoris a l'exterior de l'assemblatge i/o a la unió de DnaA al ssDNA a través dels seus llocs d'unió secundaris. **Sota:** Un cop fos l'origen, es produeix el carregament de dues helicases DnaB (blau) a través de l'acció coordinada de les proteïnes DnaA i DnaC (taronja). Figura adaptada de Mott *et al.*, 2007.

Després de la seva fusió inicial, el DUE es manté desunit gràcies a la unió de DnaA a les caixes DnaA-ATP monocadena, una interacció que requereix la forma unida a ATP de l'iniciador. Es coneix molt poc sobre la situació d'aquest lloc d'unió a ssDNA dins la proteïna, però basant-se en els resultats estructurals actuals tant de DnaA com d'altres membres de la superfamília AAA+, s'hipotitza que el canal central del filament de DnaA-ATP constituïria una regió d'unió a DNA secundària, la qual podria ser

utilitzada bé per conduir a la fusió del DUE i/o per establir el DNA desenrotllat com a preparació pels següents passos de l'assemblatge del primosoma (Erzberger *et al.* 2006b; Mott *et al.* 2007; Duderstadt *et al.* 2008). Anàlisis mutacionals dels residus que formen el canal central del filament de DnaA corroboren aquesta hipòtesi, doncs no poden promoure la fusió del DUE *in vivo* i els estudis *in vitro* indiquen que, tot i ser capaços de formar autoassemblatges, són deficientes en la unió a ssDNA (Ozaki *et al.* 2008).

Després del desenrotllament del DUE, es produeix el carregament de dues molècules de l'helicasa replicativa DnaB a *oriC*, cosa que inicia la cascada d'esdeveniments que provocaran l'assemblatge del replisoma per formar dues forquetes de replicació completament equipades i la conseqüent progressió vers l'etapa d'elongació de la replicació del DNA (Tougu *et al.* 1996b). Tal i com es descriu a la secció següent, en aquesta etapa hi intervé la pròpia DnaA i una segona proteïna, el carregador d'helicasa DnaC (Baker *et al.* 1987; Funnell *et al.* 1987; Sutton *et al.* 1998; Fang *et al.* 1999).

1.3.1 Carregament de DnaB a un *oriC* obert

DnaB és capaç d'autocarregar-se a una cadena senzilla de DNA en una reacció dependent d'ATP, però ho fa de manera poc eficient (Jezewska *et al.* 1996). Així doncs, el carregament eficient de DnaB necessita la participació de la proteïna carregadora de l'helicasa, DnaC (Wahle *et al.* 1989b), per facilitar l'obertura de l'anell (Davey *et al.* 2003). De totes maneres, aquesta activitat no és suficient *in vivo* per carregar l'helicasa a l'origen de replicació, doncs ni DnaB per si sola ni el complex DnaB-DnaC presenten cap mena d'afinitat per ssDNA unit per les proteïnes d'unio a DNA de cadena senzilla (proteïnes SSB de l'anglès *single stranded DNA binding proteins*) (Arai *et al.* 1981b; LeBowitz *et al.* 1986; Wahle *et al.* 1989b), les quals embolcallen el ssDNA originat per la fusió de l'origen. Aquesta barrera de SSB proporciona un mecanisme de control per impedir la formació aberrant de forquetes de replicació a llocs no permesos. La proteïna que permet superar la barrera de SSB a *oriC* és DnaA (Marszalek *et al.* 1994). Per tant, l'entrada de DnaB a l'*oriC* desunit depèn de dos factors proteics addicionals, les

proteïnes DnaA i DnaC (Messer *et al.* 2001; Konieczny 2003). Ambdues pertanyen al mateix subgrup dins la superfamília AAA+ (Koonin 1992) i, tal i com es detalla més endavant, comparteixen moltes característiques funcionals i estructurals. Aquestes semblances són particularment rellevants pel fet que només una part dels bacteris contenen un ortòleg de DnaC identificable. En aquestes espècies DnaA podria realitzar el paper tant de l'iniciador com del carregador d'helicasa (Mott *et al.* 2008).

1.3.1.1 Procés

El primer pas en el procés de carregament de l'helicasa replicativa a un *oriC* fos és l'associació entre DnaB i DnaC. DnaC és una ATPasa de dos dominis: un domini N-terminal petit encarregat de la interacció amb DnaB i un domini ATPasa C-terminal amb el plegament característic AAA+. En presència d'ATP, DnaC forma un anell hexamèric sobre d'una de les cares de l'hexàmer de DnaB (Kobori *et al.* 1982; Nakayama *et al.* 1987; Wahle *et al.* 1989a; San Martin *et al.* 1998; Barcena *et al.* 2001). DnaC ha de trobar-se unida a ATP per a carregar l'helicasa a l'origen, malgrat que en aquesta forma DnaC no pot dissociar-se de l'helicasa i inhibeix la seva activitat de desenrotllament (Arai *et al.* 1981a). Per a l'alliberament i l'activació de DnaB, però no per a la seva deposició a *oriC*, és necessària l'activitat ATPasa de DnaC, la qual està induïda per la presència simultània de ssDNA i DnaB (Davey *et al.* 2002).

Un cop format el complex DnaB₆-DnaC₆, la DnaA unida al DUE desenrotllat l'atreu cap a l'origen. Diferents dades bioquímiques mostren que una interacció directa entre DnaA i DnaB és essencial per a la càrrega i el correcte posicionament de l'helicasa a *oriC* (Marszalek *et al.* 1994) i que aquests depenen no només de la unió de DnaA a les caixes DnaA presents a la seqüència d'*oriC*, sinó també a la unió de DnaA a la regió oberta de l'origen (Messer *et al.* 2001).

Subsegüentment, DnaC carrega DnaB al complex obert (Davey *et al.* 2003). L'assemblatge de DnaB a *oriC* per DnaC estén la regió desenrotllada de 44 a uns 65 nucleòtids, en un procés independent de l'activitat desenrotlladora de l'helicasa (Fang *et al.* 1999). Un cop s'ha produït el carregament, s'activa l'activitat ATPasa de DnaC, que es dissocia de l'helicasa, inactivant així la inhibició que exercia sobre aquesta (Funnell

et al. 1987; Wahle *et al.* 1989b; Allen *et al.* 1991). Per aquest motiu és considerada DnaC un interruptor dual ATP/ADP: la forma unida a ATP promou el carregament de l'helicasa i la forma unida a ADP promou el desenrotllament del DNA (Davey *et al.* 2002). L'activitat d'unió a ssDNA de DnaC està involucrada en el carregament de DnaB (Learn *et al.* 1997).

1.3.1.2 Model de carregament de dues DnaB a una forqueta bidireccional

Tal i com s'ha descrit anteriorment, estudis primerencs ja indicaven una interacció directe entre DnaB i DnaA (Marszalek *et al.* 1994), que es va localitzar al domini N-terminal d'ambdues proteïnes (Marszalek *et al.* 1996; Sutton *et al.* 1998; Seitz *et al.* 2000; Abe *et al.* 2007). S'havia hipotetitzat llargament sobre una possible interacció directe entre DnaA i DnaC (Funnell *et al.* 1987), basant-se en la pertinença d'ambdues a superfamília AAA+, però no ha estat fins recentment que s'ha demostrat que ambdues proteïnes poden interactuar de manera dependent d'ATP (Mott *et al.* 2008).

Paral·lelament, s'ha resolt l'estructura del domini AAA+ de la DnaC d'*Aquifex aeolicus*. Aquesta estructura estableix que el carregador és un paràleg estructural proper de la proteïna iniciadora DnaA, alhora que revela que DnaC pot adoptar un assemblatge helicoidal de mà dreta similar al de DnaA i que aquesta propietat es deguda a una inserció helicoidal α conservada compartida amb l'iniciador. La mutació dels residus que es troben a la interfície entre subunitats distorsiona la funció de DnaC *in vivo* i compromet l'habilitat del carregador per associar-se productivament amb ssDNA *in vitro*, mostrant que DnaC necessita l'autoassemblatge per a la seva funció. Això suggereix que DnaC té funcions més enllà del carregament de l'helicasa al ssDNA. En particular, podria funcionar com a adaptador molecular per reconèixer específicament la forma activada amb ATP de l'iniciador a través d'interaccions AAA+/AAA+ directes entre DnaC i els extrems lliures del complex d'iniciador assemblat (Mott *et al.* 2008).

Aquestes noves dades (la interacció entre DnaA i DnaC i la caracterització estructural de la darrera), permeten formular un model per l'assemblatge a *oriC* de les dues molècules de DnaB necessàries per formar una forqueta de replicació bidireccional.

Durant el carregament de l'helicasa replicativa a *oriC*, els dos hexàmers de DnaB han de posicionar-se en direccions oposades a les dues cadenes del DUE desenrotllat (Fang *et al.* 1999; Kaplan *et al.* 2004). Estudis bioquímics i de microscòpia electrònica indiquen que el complex nucleoproteic de DnaA es forma a un dels costats del DUE i que DnaA s'associa predominantment amb la cadena superior de la regió de l'origen fosa (Funnell *et al.* 1987; Speck *et al.* 2001). Aquesta organització suposa la paradoxa de com un assemblatge d'iniciació asimètric facilita el carregament simètric de dues helicases en aquesta regió.

En substrats de DNA en forqueta artificials que mimetitzen un *oriC* fos, DnaA pot reclutar un sol complex DnaB-DnaC a la cadena de sota de l'origen desunit, suggerint que aquest complex és dipositat gràcies a interaccions directes entre l'iniciador i l'helicasa (Weigel *et al.* 2002). Per contra, la forma en que la segona helicasa es carregada roman poc clara. Donat que DnaA es troba dipositada a la part dreta del DUE, una interacció idèntica a la descrita anteriorment entre l'iniciador i DnaB no permetria una correcta orientació de l'helicasa a la cadena superior del DUE, doncs la polaritat d'aquesta no seria l'adequada.

Es coneix que l'iniciador i el carregador de l'helicasa reconeixen cares diferents de l'hexàmer de DnaB. El domini N-terminal de DnaA s'uneix a la cara N-terminal de l'helicasa (Sutton *et al.* 1998), mentre que el domini N-terminal de DnaC s'uneix a la cara C-terminal de DnaB (Barcena *et al.* 2001; Ludlam *et al.* 2001). Per tant, un contacte directe entre DnaC i l'oligòmer de DnaA permetria que la segona helicasa fos col·locada amb la geometria correcta en aquest segment de DNA gràcies a trobar-se associada amb la proteïna carregadora de l'helicasa (Mott *et al.* 2008). Finalment, el mecanisme mitjançant el qual s'obre l'anell de DnaB per permetre l'entrada del DNA de cadena senzilla no es coneix, encara que la formació d'una estructura oligomèrica en espiral de DnaC a la superfície de l'helicasa podria proporcionar un mitjà per facilitar aquest esdeveniment.

A la figura I.20 es representa esquemàticament el model de carregament a *oriC* de dues molècules de DnaB en l'orientació correcta.

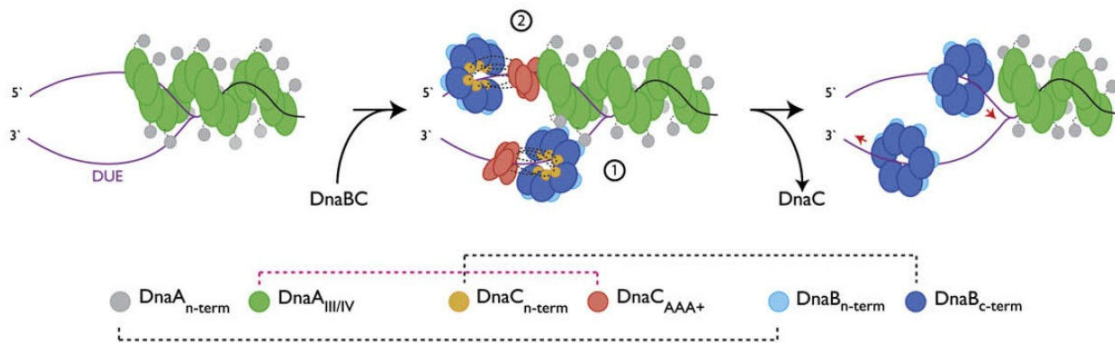


Figura I.20: Model per a un carregament asimètric de dues molècules de DnaB a *oriC*. **Esquerra:** L'assemblatge de DnaA a l'origen de replicació fon el DUE. **Centre:** L'helicasa de la forqueta esquerra és dipositada prop de l'oriC a través de la seva interacció directa amb DnaA (1), mentre que l'helicasa de la forqueta dreta és situada a una certa distància de l'oriC a través de la seva associació amb DnaC i la interacció d'aquesta amb el filament de DnaA (2). **Dreta:** Es desfà el complex DnaB-DnaC i DnaC deixa l'oriC, alliberant la inhibició que exercia sobre DnaB. Ambdós hexàmers d'helicasa migren fins assolir la seva posició correcta dins la forqueta de replicació bidireccional. (Adaptat de Mott *et al.*, 2008).

Aquest model prediu que l'helicasa destinada a encapçalar la forqueta de replicació esquerra és carregada inicialment a la cara dreta de la cadena inferior del DUE, mentre que l'helicasa que generarà la forqueta de replicació dreta és dipositada a la part esquerra de la cadena de sobre del DUE. Efectivament, durant la iniciació es produeix aquest arrencament asimètric, amb les dues helicases lliscant en sentit 5' a 3', passant una de llarg de l'altre després de la deposició per establir la geometria final de la forqueta de replicació bidireccional (Fang *et al.* 1999). Així doncs, aquest model proporciona una explicació física a la configuració de carregament inicial.

1.4 Regulació de la iniciació: evitar sobreiniciacions

Les bactèries usen molts mecanismes regulatoris diferents per assegurar-se que la replicació es dona una sola vegada a cada cicle cel·lular i de manera simultània a tots els orígens presents a la cèl·lula (Mott *et al.* 2007). Tal i com s'ha descrit anteriorment, l'estat nucleotídic de DnaA ajuda a aquesta regulació, però no és l'únic mecanisme emprat. Algunes espècies bacterianes tenen mecanismes de control homòlegs mentre que altres utilitzen rutes regulatòries completament diferents, emfatitzant la importància de controlar la iniciació de la replicació, fins i tot en cèl·lules procariotes.

A *E. coli*, tant *oriC* com DnaA són dianes per al control regulatori. A part de l'activació de DnaA per ATP, altres mecanismes regulatoris inclouen el segrest de l'origen entre cicles d'iniciació per la proteïna SeqA, que s'uneix a seqüències d' *oriC* hemimetilades doncs acaben de ser replicades; i un descens de la concentració intracel·lular de DnaA lliure a través de la unió de l'iniciador a regions genòmiques fora de l'origen que contenen un gran nombre de llocs d'unió a DnaA. Addicionalment, DnaA pot regular la seva pròpia expressió per mitjà de la seva unió a caixes DnaA que flanquegen el gen *dnaA*. Finalment, l'activitat de l'iniciador està subjecte a processos de control externs que inclouen la unió d'una proteïna coneguda com DiaA per assegurar la sincronia de la iniciació, i la ruta d'inactivació regulatòria de DnaA (RIDA: *Regulatory Inactivation of DnaA*).

RIDA és particularment important, doncs hi està implicada la proteïna Hda, membre del *clade* dels iniciadors de la superfamília AAA+ i un paràleg de DnaA. Hda és un homodímer que s'associa amb la pinça de processivitat β d'*E. coli* i subsegüentment catalitza la hidròlisi de nucleòtid per part de la proteïna iniciadora, convertint la DnaA-ATP en DnaA-ADP. Hda inicia aquesta hidròlisi als extrems lliures del complex nucleoproteic que DnaA-ATP forma a *oriC*, provocant el desassemblatge d'aquest complex, bé de manera seqüencial o catastròfica (Duderstadt *et al.* 2008). Es creu que Hda i DnaA interaccionen a través del seu mòdul AAA+, que en aquest cas estaria utilitzat no com un mòdul d'autoassemblatge sinó com una bastida per promoure l'associació i la comunicació entre diferents factors d'iniciació de la replicació. Hda només és activa en associació amb una pinça β que es trobi en DNA dúplex que hagi estat extès per la DNA polimerasa (Su'etsugu *et al.* 2004; Su'etsugu *et al.* 2005). Aquesta associació assegura que la hidròlisi de nucleòtid per part de DnaA es dona al moment adequat durant la formació del replisoma per promoure el desassemblatge del complex de DnaA i prevenir la reiniciació.

2. Elongació de la replicació bacteriana

El carregament de l'helicasa marca el final de l'etapa d'iniciació i el començament de l'etapa d'elongació, doncs dirigeix l'assemblatge del replisoma, la maquinària cel·lular encarregada de la replicació del genoma bacterià.

Per a poder replicar el genoma, calen tres activitats bàsiques: el desenrotllament del DNA dúplex parental per poder accedir a la informació que aquest codifica, la síntesi d'encebadors des d'on iniciar la replicació i la síntesi del nou DNA seguint la complementarietat amb la cadena motlle. El replisoma agrupa físicament aquestes activitats en un sol complex multiproteic que és capaç de replicar el DNA cromosòmic amb gran fidelitat i velocitat gràcies a la coordinació dels diferents actors: l'helicasa DnaB que desenrotlla el DNA dúplex al front de la forqueta de replicació, la primasa DnaG que sintetitza els encebadors de RNA i l'holoenzim polimerasa de DNA III que sintetitza les cadenes filles.

2.1 Components del replisoma

2.1.1. Holoenzim polimerasa de DNA III (Pol III)

L'holoenzim Pol III està compost de més de 10 subunitats proteiques que es poden agrupar en tres subcomplexes funcionals: el *core* polimerasa, la pinça lliscant β i el complex carregador de pinça γ .

El *core* polimerasa

El *core* polimerasa és un heterotrímer format per les subunitats α , ϵ i θ (McHenry 2003). La subunitat α és una polimerasa de DNA pertanyent a la família C, exclusiva d'eubactèries (Braithwaite *et al.* 1993). L'estructura cristal·lina de la subunitat α d'*E. coli* i de *Thermus aquaticus* indiquen que l'organització del seu centre actiu no és homòloga a la de les polimerases replicatives eucariòtiques o d'arquees, sinó que és similar a les polimerases de DNA de la família X, les quals estan involucrades en

l'edició del RNA i la reparació del DNA (Bailey *et al.* 2006; Lamers *et al.* 2006; Wing *et al.* 2008). La subunitat ϵ és una exonucleasa 3'-5' amb activitat correctora que funciona editant el producte de DNA per assegurar una alta fidelitat de còpia. Finalment, la subunitat θ no té cap activitat coneguda llevat d'una lleugera estimulació d' ϵ (Johnson *et al.* 2005).

El *core* $\alpha\epsilon\theta$ és una polimerasa lenta (≈ 20 nucleòtids s^{-1}) i poc processiva, que incorpora només uns 10 nucleòtids abans de dissociar-se del dúplex encebador-motlle. En canvi, quan actua lligat a la pinça lliscant β , la seva processivitat s'incrementa marcadament fins assolir varies kilobases (>50 kb) a un velocitat aproximada de $1\text{ kb } s^{-1}$ (Stukenberg *et al.* 1991).

La pinça lliscant β

Les pinces β confereixen processivitat al replisoma gràcies a que lliguen les polimerases replicatives a les seves respectives cadenes motlle. Aquesta associació s'aconsegueix gràcies a que la pinça β encercla el DNA dúplex immediatament darrera de la Pol III mentre roman unida a la subunitat α del *core* polimerasa (Stukenberg *et al.* 1991). A mesura que la replicació progressa, β llisca al llarg del DNA desplaçant-se conjuntament amb el *core* de la Pol III.

La pinça lliscant és un homodímer de subunitats β en forma d'anell que presenta pseudosimetria hexamèrica degut a que cada protòmer β conté tres dominis globulars que comparteixen una mateixa topologia de plegament. Els protòmers de β es troben en una disposició cap-i-cua, donant lloc a dos plans estructuralment diferents (Kong *et al.* 1992). El *core* Pol III i el carregador de pinça γ s'uneixen a la pinça β a través d'una butxaca hidrofòbica situada a la superfície de la cara C-terminal. Addicionalment, la cara C-terminal de β també està involucrada en interaccions intermoleculares amb proteïnes de recombinació, reparació i control del cicle cel·lular, fet que indica que les pinces lliscants tenen un paper molt ampli dins dels processos metabòlics del DNA (Johnson *et al.* 2005).

Recentment, s'ha determinat experimentalment que la pinça β pot unir el DNA de manera específica i que interactua tant amb DNA de cadena senzilla com de cadena

doble. L'estructura cristal·lina d'una pinça β unida a un DNA encebata mostra que el dsDNA passa a través de l'anell β formant un angle brusc de 22° . Al seu torn, el motlle de cadena senzilla del lloc encebata s'uneix a la butxaca hidrofòbica de la cara C-terminal, el lloc d'unió a proteïnes de la pinça β (Georgescu *et al.* 2008).

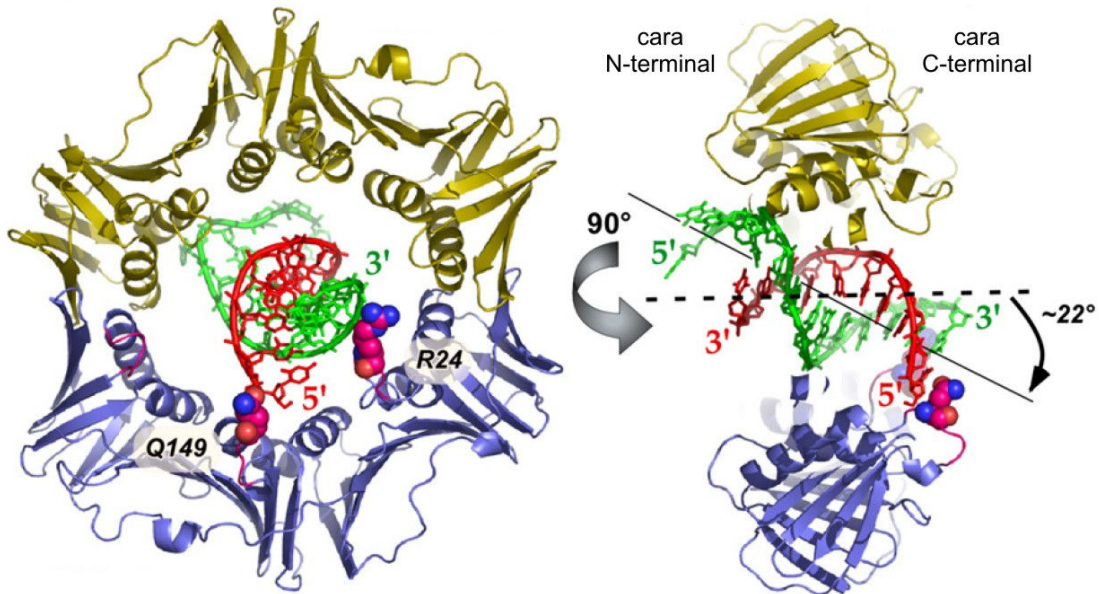


Figura I.21: Estructura de la pinça β unida a DNA. Cada protòmer β es representa en un color diferent, groc i blau. A la part esquerra (vista lateral) es pot observar com el dsDNA unit a l'interior de la pinça β presenta un angle de 22° . Codi PDB: 3D1E.

In vivo, per tal que la pinça β s'associï a un heterodúplex encebador-motlle, cal l'obertura de l'anell per part del complex multiproteic γ , el carregador de la pinça (Jeruzalmi *et al.* 2001a).

El complex carregador de pinça γ

El complex γ conté fins a 6 subunitats diferents necessàries tant pel carregament de la pinça β com per la coordinació de les diferents activitats enzimàtiques que tenen lloc a la forqueta de replicació.

Un complex γ mínim, que permet el carregament de la pinça β , conté 3 còpies de la proteïna γ i una còpia de cadascuna de les proteïnes δ i δ' (Jeruzalmi *et al.* 2001b). A la forqueta de replicació, dues subunitats γ del complex carregador de la pinça són

reemplaçades per dues subunitats τ per permetre l'associació del carregador de pinça amb les polimerases duals. γ i τ són producte del mateix gen, *dnaX*, i és diferencien en que γ és 34 kDa més petita a la regió C-terminal degut a un canvi de pauta de lectura traduccional que incorpora un sol aminoàcid abans de produir una truncació prematura de la proteïna. Precisament, τ empra aquesta regió C-terminal extra per interaccionar amb el *core* polimerasa i amb DnaB. Tant γ com τ hidrolitzen ATP i uneixen δ i δ' , modulant la unió i l'obertura de β_2 . La subunitat δ aïllada és capaç d'unir i obrir la pinça β . δ s'uneix a una regió hidrofòbica prop de d'interfície dels dímers de β , distorsionant aquesta superfície de contacte i provocant l'obertura de la pinça. Finalment, la subunitat δ' té un paper estructural, doncs funciona com a un cos rígid en el context del complex γ (Johnson *et al.* 2005; Pomerantz *et al.* 2007).

Anàlisis estructurals del complex carregador de pinça mínim ($\gamma_3\sigma\sigma'$) rebel·len que les subunitats que el conformen s'organitzen en espiral, deixant un forat entre les subunitats σ i σ' que proporciona accés al DNA al centre del complex γ , tal i com es mostra a la figura I.22. El dímer β és obert per la subunitat σ , i els residus de la pinça β que interaccionen amb σ es localitzen al domini N-terminal, indicant que l'anell β s'uneix per la part de sota del complex γ , on pot formar connexions amb totes les subunitats (Jeruzalmi *et al.* 2001a; Jeruzalmi *et al.* 2001b).

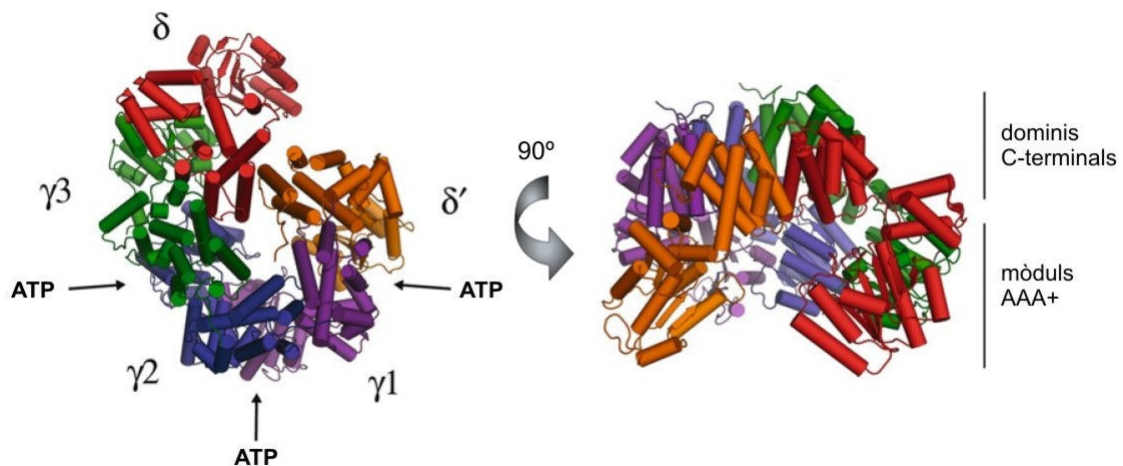


Figura I.22: Estructura del complex carregador de pinça mínim d' *E. coli*. Per bé que només les subunitats γ són capaces d'unir ATP, totes cinc pertanyen a la superfamília AAA+. Es mostren els tres llocs d'unió a ATP, a les interfícies entre subunitats. Codi PDB: 1JR3.

Adicionalment, a *E. coli* les proteïnes χ i ψ s'associen amb el complex γ per permetre el reconeixement de les proteïnes SSB i així facilitar el canvi de la primasa a la polimerasa a la cadena retardada. La subunitat χ s'uneix a SSB i es manté unida al carregador de pinça a través de la seva interacció amb ψ (Xiao *et al.* 1993; Glover *et al.* 1998; Yuzhakov *et al.* 1999).

2.1.2 L'helicasa replicativa: DnaB

L'helicasa DnaB és el primer component del replisoma en arribar a la futura forqueta de replicació i participa activament en el reclutament i assemblatge dels altres membres del replisoma.

L'activitat principal de DnaB dins del replisoma és convertir l'energia de l'ATP en treball mecànic de translocació i separació de cadena al front de la cadena retardada de la forqueta de replicació (LeBowitz *et al.* 1986). Ara bé, aquesta no és la única funció de DnaB sinó que també està implicada en coordinar la progressió de la forqueta amb l'activitat de l'holoenzim Pol III, així com en regular i coordinar la síntesi d'encebadors a la cadena retardada. Exerceix aquestes funcions regulatòries gràcies a interaccions directes amb la subunitat τ de Pol III i la primasa DnaG, respectivament (Johnson *et al.* 2005). L'activitat helicasa de DnaB, intrínscament baixa, augmenta en un factor de 10 en el context del replisoma gràcies a la seva interacció amb τ (Kim *et al.* 1996).

2.1.3 La primasa: DnaG

DnaG és una polimerasa de RNA especialitzada que sintetitza encebadors de RNA d'uns 10-12 pb que proporcionen un extrem 3' lliure a partir del qual les polimerases de DNA poden continuar la síntesi de cadena. És reclutada al replisoma mitjançant una interacció transitòria amb DnaB, que es repeteix cíclicament durant l'avanç del replisoma, doncs és clau en la coordinació del cicle de la cadena retardada (Tougu *et al.* 1996b).

La primasa s'organitza en tres dominis: un domini N-terminal d'unió a zinc implicat en l'estabilització de la unió de la primasa al seu DNA motlle (Corn *et al.* 2005); un domini polimerasa de RNA central, i un domini C-terminal d'interacció amb l'helicasa DnaB i les proteïnes SSB, el qual localitza la primasa a la forqueta de replicació.

DnaG és monomèrica en solució, però en aquesta forma no és activa, doncs cal que dues molècules de primasa interaccionin en *trans* per a produir un encebador. El domini polimerasa de DnaG aïllat sintetitza encebadors extremadament llargs a una taxa molt lenta. Necessita interaccionar amb el domini d'unió a zinc, el qual estabilitza la seva unió amb el ssDNA motlle i, conseqüentment, permet produir encebadors de la mida adequada i correctament espaiats. Com la zona d'enllaç entre ambdós dominis és massa curta per a que aquesta interacció es pugui donar intramolecularment, cal que dues molècules de primasa s'associïn i els seus dominis col·laborin. Aquesta cooperació entre subunitats augmenta l'activitat catalítica de la primasa, reforça la seva especificitat d'iniciació i permet que el domini d'unió a zinc actuï com fre molecular que restringeix tant la processivitat com la llargària dels encebadors (Corn *et al.* 2005; Corn *et al.* 2006).

Al replisoma, tres molècules de DnaG es troben unides a un hexàmer de DnaB i això possibilita que el domini d'unió a zinc d'una primasa veïna reguli la síntesi d'encebador de la primasa unida al DNA motlle.

2.2 Arquitectura del replisoma

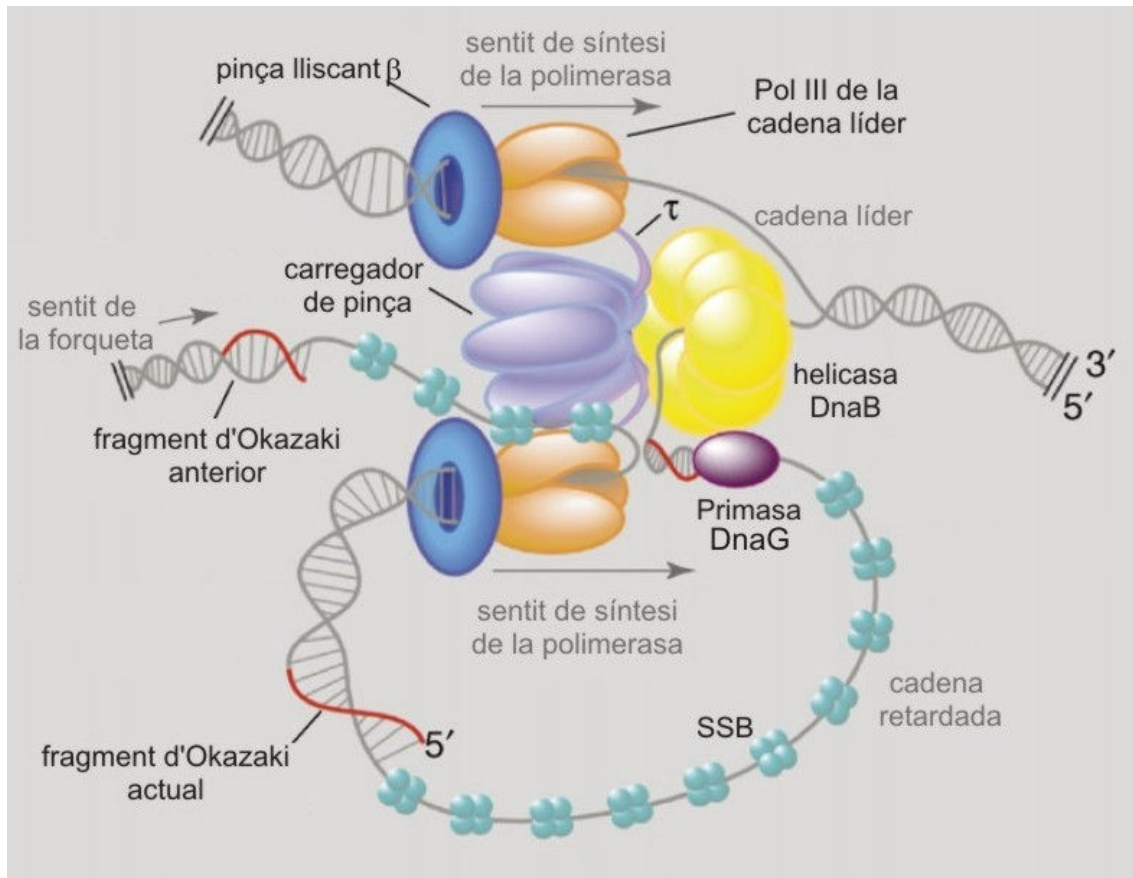


Figura I.23: Esquema de l'organització del replisoma durant la replicació del DNA. Per claredat només es representa una molècula de primasa, quan en realitat tres molècules de primasa interaccionen amb l'hexàmer d'helicasa i dues d'elles coordinen la seva activitat per a la síntesi eficient de l'encebador. Figura adaptada de Pomerantz i O'Donnell, 2007.

El replisoma emprava una polimerasa dual, amb dos *cores* $\alpha\epsilon\theta$ enllaçats mitjançant el complex γ , per dur a terme la síntesi simultània de les dues cadenes del DNA parental. Ara bé, com les polimerases de DNA només poden sintetitzar cadena en el sentit 5' a 3' i el DNA dúplex és antiparal·lel, la replicació del genoma s'ha de realitzar de manera asimètrica. La cadena que presenta polaritat 5'-3' es copiada de manera contínua, en la mateixa direcció en que es dona el desenrotllament de la forqueta de replicació, i s'anomena cadena líder. En canvi, l'altra cadena és copiada de manera discontinua, en fragments de DNA curts de 1-3 kb anomenats fragment d'Okazaki. Aquesta cadena s'anomena retardada i ha d'organitzar-se formant un llaç de cadena senzilla per tal de posicionar-se en el sentit adequat per poder ser copiada per la seva

Pol III corresponent a mesura que aquesta es mou amb el replisoma. Les proteïnes SSB embolcallen el llaç de ssDNA de la cadena retardada per tal d'estabilitzar-lo i protegir-lo de l'atac de les nucleases cel·lulars. Les SSB també tenen papers addicionals tan dins del replisoma com en el metabolisme general del DNA gràcies a les nombroses proteïnes amb les quals poden interaccionar (Shereda *et al.* 2008).

2.3 Interacció entre l'helicasa i la primasa

Tres molècules de DnaG s'uneixen a través del seu domini C-terminal (Tougu *et al.* 1994; Tougu *et al.* 1996a) a la cara N-terminal d'un hexàmer de DnaB (Chang *et al.* 2000; Bailey *et al.* 2007b; Wang *et al.* 2008), formant una associació directa que regula les funcions d'ambdues proteïnes. La primasa augmenta tant l'activitat ATPasa com helicasa de DnaB (LeBowitz *et al.* 1986; Bird *et al.* 2000). Al seu torn, DnaB modula l'activitat global de DnaG, així com la llargària dels encebadors sintetitzats per la primasa i la seva especificitat d'iniciació (Rowen *et al.* 1978; Lu *et al.* 1996; Bhattacharyya *et al.* 2000; Johnson *et al.* 2000).

Donat que DnaG sola és feblement activa *in vivo*, amb una taxa màxima d'uns tres encebadors per hora, l'activació de la primasa gràcies a la seva associació amb DnaB és essencial per a que la replicació pugui progressar a les ràpides taxes observades *in vivo*.

2.3.1 Model estructural

Les recents estructures de DnaB unida al domini d'interacció de DnaG (Bailey *et al.* 2007b) i del domini polimerasa de DnaG unit al DNA (Corn *et al.* 2008) possibiliten el modelatge del complex format per DnaB i DnaG a la forqueta de replicació, tal i com es mostra a la figura I.24.

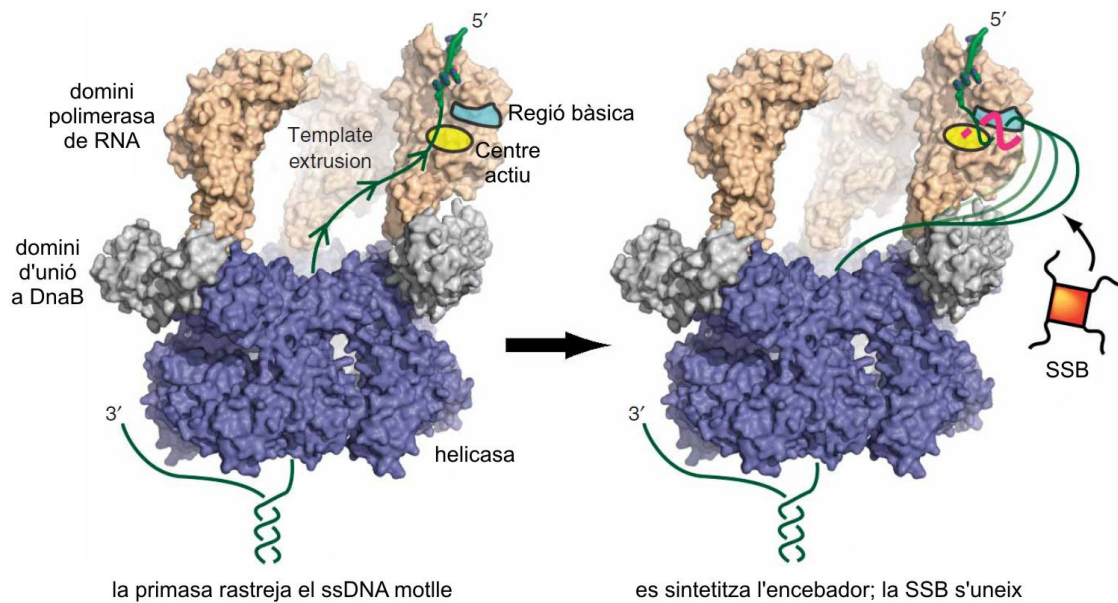


Figura I.24: Model estructural del complex helicasa-primasa a una forqueta. *Docking* del domini polimerasa de DnaG (rosa) unit a ssDNA (codi PDB: [3B39](#)) al complex entre un hexàmer de DnaB (blau) i tres dominis d'interacció amb l'helicasa de DnaG (gris) (codi PDB: 2REA). El solc de rastreig del DNA motlle queda emplaçat sobre l'helicasa, en una orientació que possibilita la seva associació al ssDNA que es alliberat de l'helicasa. Després de la captura del motlle i el seu rastreig a la recerca d'un lloc d'iniciació (esquerra), el ssDNA forma un llaç cap al centre actiu de la primasa, on es produeix la síntesi de l'encebador (fúcsia) que surt cap a la cara externa del complex helicasa-primasa (dreta). Al quedar a la cara externa, es facilita el posterior processat de l'encebador, per exemple, la seva transferència al complex carregador de pinça a través d'una interacció entre la primasa i la proteïna SSB. Per claretat, no s'ha representat el domini d'unió a zinc de la primasa.

DnaG uneix el ssDNA mitjançant una regió bàsica ortogonal al centre actiu del seu domini polimerasa, la qual forma un solc a través del qual llisca el DNA de cadena senzilla. La primasa usa aquest solc per rastrejar el ssDNA motlle a la recerca de les seqüències tribàsiques d'iniciació (Tougu *et al.* 1994; Corn *et al.* 2008). La interacció entre DnaB i DnaG orienta aquest solc de rastreig sobre la cavitat central de l'anell de l'helicasa, en una posició apropiada per unir el motlle lliure de la cadena retardada a mesura que aquest surt de la cavitat central de l'helicasa. A més, un cop sintetitzat l'encebador, l'heterodúplex RNA-DNA surt cap a l'exterior del complex helicasa-primasa, facilitant així la seva transferència final al *core* polimerasa (Corn *et al.* 2008).

El domini d'interacció amb l'helicasa de DnaG s'uneix a aquesta tapant la interfície entre dímers a la cara N-terminal de DnaB (Bailey *et al.* 2007b), fixant la simetria C3 d'aquesta capa. S'ha proposat que la primasa estimula l'activitat de

l'helicasa estabilitzant la conformació de l'hexàmer de DnaB sobre el DNA, acció que incrementaria la processivitat de desenrotllament. A més, la comunicació de la primasa cap al motor de l'helicasa és evident en una variant de G40P en la qual, quan s'interfereix amb zona d'interacció entre la capa N-terminal i la capa C-terminal de l'hexàmer de G40P, s'obté una proteïna que encara s'uneix a la primasa però que no és estimulada en les seves activitats ATPasa i helicasa en la mateixa proporció que la proteïna salvatge (Wang *et al.* 2008).

2.3.2 Control del cicle de la síntesi de fragments d'Okazaki

La interacció entre la primasa i l'helicasa determina la freqüència de la síntesi dels fragments d'Okazaki i, per tant, ajusta el rellotge de la forqueta de replicació (Tougu *et al.* 1996b). DnaG actua distributivament, dissociant-se de DnaB i associant-se de nou per sintetitzar cada fragment d'Okazaki (Wu *et al.* 1992; Johnson *et al.* 2000; Mitkova *et al.* 2003).

Al replisoma, l'hexàmer de DnaB actua com a estació port mòbil per incrementar la concentració local de motlle de DNA de cadena senzilla disponible per a la primasa, alhora que possibilita la colocalització de varies DnaG en una disposició adequada per a que els seus dominis interactuïn, modulant així la síntesi d'encebador.

Adicionalment, estudis d'una sola molècula (*single-molecule*) de la síntesi de la cadena líder catalitzada per DnaB i l'holoenzim Pol III mostren que la unió cooperativa de tres molècules de DnaG a un hexàmer de DnaB desestabilitza la forqueta de replicació, reduint la processivitat del replisoma per un factor de quatre (Tanner *et al.* 2008) Aquest efecte és similar a l'encallament del replisoma de T7 observada en sistemes similars de *single-molecule* quan la proteïna del gen 4 sintetitza un encebador, de totes maneres, en aquest cas la processivitat no es veia afectada (Johnson *et al.* 2007). Aquesta modulació de la replicació de la cadena líder a través de la interacció de la primasa amb DnaB podria ser una mecanisme per prevenir que la síntesi a la cadena líder adelanti la replicació de la cadena retardada durant una síntesi anormalment lenta d'encebadors a la cadena retardada (Marians 2008; Tanner *et al.* 2008).

2.3.3 Dissociació del complex i transferència de l'encebador al complex γ

La unió del domini d'interacció de DnaG a la proteïna SSB provoca que la primasa unida a un heterodúplex encebador-motlle es dissociï de l'helicasa. La primasa roman associada a l'encebador acabat de sintetitzar per estabilitzar el curt heterodúplex RNA-DNA i assegurar la transferència apropiada d'aquest a la polimerasa replicativa (Yuzhakov *et al.* 1999). Encara que la primasa és un monòmer en solució, l'estequiometria de la primasa unida a llocs encebats del DNA és de 2:1, cosa que suggereix un mecanisme d'alliberament concertat en que dues primases es dissocien conjuntament de l'hexàmer DnaB (Marians 2008).

Els encebadors sintetitzats de nou són emprats per la polimerasa de la cadena retardada amb una eficiència molt alta, doncs més del 90% dels encebadors sintetitzats per la primasa a la forqueta de replicació són usats per iniciar la síntesi de fragments d'Okazaki (Wang *et al.* 2008). Aquesta alta eficiència és producte d'una reacció concertada on la interacció entre la primasa i la proteïna SSB és reemplaçada competitivament per una entre SSB i la subunitat χ del carregador de pinça (Yuzhakov *et al.* 1999). Aquest intercanvi dissocia la primasa de l'encebador, possibilitant la càrrega de la pinça β de l'holoenzim Pol III a l'heterodúplex encebador-motlle i, per tant, la subsegüent unió de la subunitat catalítica (α) de l'holoenzim a l'extrem 3' de l'encebador.

2.4 Carregament de la pinça β i entrada de Pol III

En presència d'ATP, el carregador de pinça s'uneix a i obra la pinça β . La pinça s'obra formant una espiral de mà dreta, distribució que encaixa molt bé amb l'organització espiral de les subunitats del carregador de pinça. Dins del complex γ , és la subunitat σ la que obra la pinça, alineant el forat entre les subunitats σ i σ' amb el forat de la pinça oberta. Aquest complex obert de pinça β - carregador γ uneix el DNA encebat i el posiciona a través de la pinça oberta gràcies a les estretes connexions entre els dominis C-terminal del pentàmer del carregador de pinça, les quals impedeixen la

sortida del DNA per la part superior de l'estructura i l'obliguen a fer-ho pel forat entre σ i σ' (Bloom 2006).

Un cop el DNA es troba posicionat a l'interior de la pinça β , atreu els residus d'interacció amb el dsDNA de la pinça, induint el tancament d'aquesta al seu voltant. Al canviar d'una espiral oberta a una estructura tancada planar i inclinada, la pinça β ja no encaixa amb la superfície espiral del carregador de pinça, trencant així la seva connexió amb algunes de les subunitats del carregador. La subunitat del carregador de pinça que més fortament interactua amb β és la subunitat δ , la qual s'uneix a la butxaca d'unió de proteïna de β . Per tant, la unió del DNA motlle de cadena senzilla a la butxaca d'unió a proteïna de β pot ajudar a alliberar δ de β i així completar l'alliberament del carregador de pinça, un requisit necessari per a que la polimerasa uneixi β (Georgescu *et al.* 2008).

Quan el carregador de pinça es dissocia de β , la pinça roman al lloc encebada a través de la seva connexió amb el ssDNA motlle. En absència d'aquesta connexió, β podria lliscar fora de l'extrem 3'. Per tant, la connexió entre la pinça β i el ssDNA al lloc encebada serveix per mantenir β al lloc encebada, on se la necessita per interaccionar amb PolIII. Com l'holoenzim competeix amb el ssDNA per al lloc d'unió a proteïna de β , la interacció ssDNA- β es trenca quan la polimerasa s'associa amb la pinça, fet que facilita la difusió de β al llarg del DNA. Finalment, un cop associada a la seva pinça de processivitat β , l'holoenzim Pol III inicia la síntesi del nou DNA a l'extrem 3' de l'encebador.

2.5 Mecanismes per a la síntesi de la cadena líder i la retardada

L'extensió de la cadena líder només necessita d'un sol pas, o pocs, de formació d'encebadors i per tant, es dona de manera continua. Aquest punt de vista, àmpliament acceptat, es basa en assajos de replicació realitzats *in vitro* amb proteïnes altament purificades. Ara bé, un estudi *in vivo* recent dels intermediaris de replicació, suggereix que el procés de síntesi de la cadena líder s'interrompia freqüentment. A més, inactivacions ocasionals de la forqueta de replicació degut a danys al DNA o altres

impediments a llarg de la cadena líder desemboquen en noves etapes de formació d'encebadors, que són necessaris per la reiniciació de la replicació. Actualment s'està investigant la freqüència amb que s'interromp la síntesi de la cadena líder en condicions de creixement normals.

En canvi, la síntesi de la cadena retardada necessita múltiples encebadors perquè aquesta cadena s'ha de sintetitzar com a una sèrie de fragment discontinus. La síntesi de cadascun dels fragments de la cadena retardada a partir d'un encebador de RNA i continua fins que el fragment es troba l'extrem 5' del pròxim encebador riu avall, fet que provoca la dissociació de la polimerasa de DNA pol III de la cadena retardada. La porció 5' de RNA de cada fragment d'Okazaki és reemplaçada posteriorment per desoxiribonucleòtids per la polimerasa de DNA I, i els forats entre fragments de DNA adjacents són units entre ells per la lligasa.

A mesura que el replisoma progressa, el complex carregador γ assembla repetidament pinces β als llocs on la primasa ha sintetitzat un encebador. Al seu torn, la polimerasa de la cadena retardada es dissocia de la pinça β quan completa un fragment d'Okazaki, per tal de poder estendre un encebador de RNA riu amunt per al següent fragment d' Okazaki. La poca disponibilitat cel·lular de Pol III (10-20 molècules per cèl·lula) i els centenars de fragments d'Okazaki que cal generar, implica la necessitat d'un reciclatge ràpid i eficient de Pol III.

A la figura I.25 es mostra esquemàticament la progressió de l'etapa d'elongació i els processos que tenen lloc tant durant la síntesi de la cadena líder com durant la síntesi de la cadena retardada.

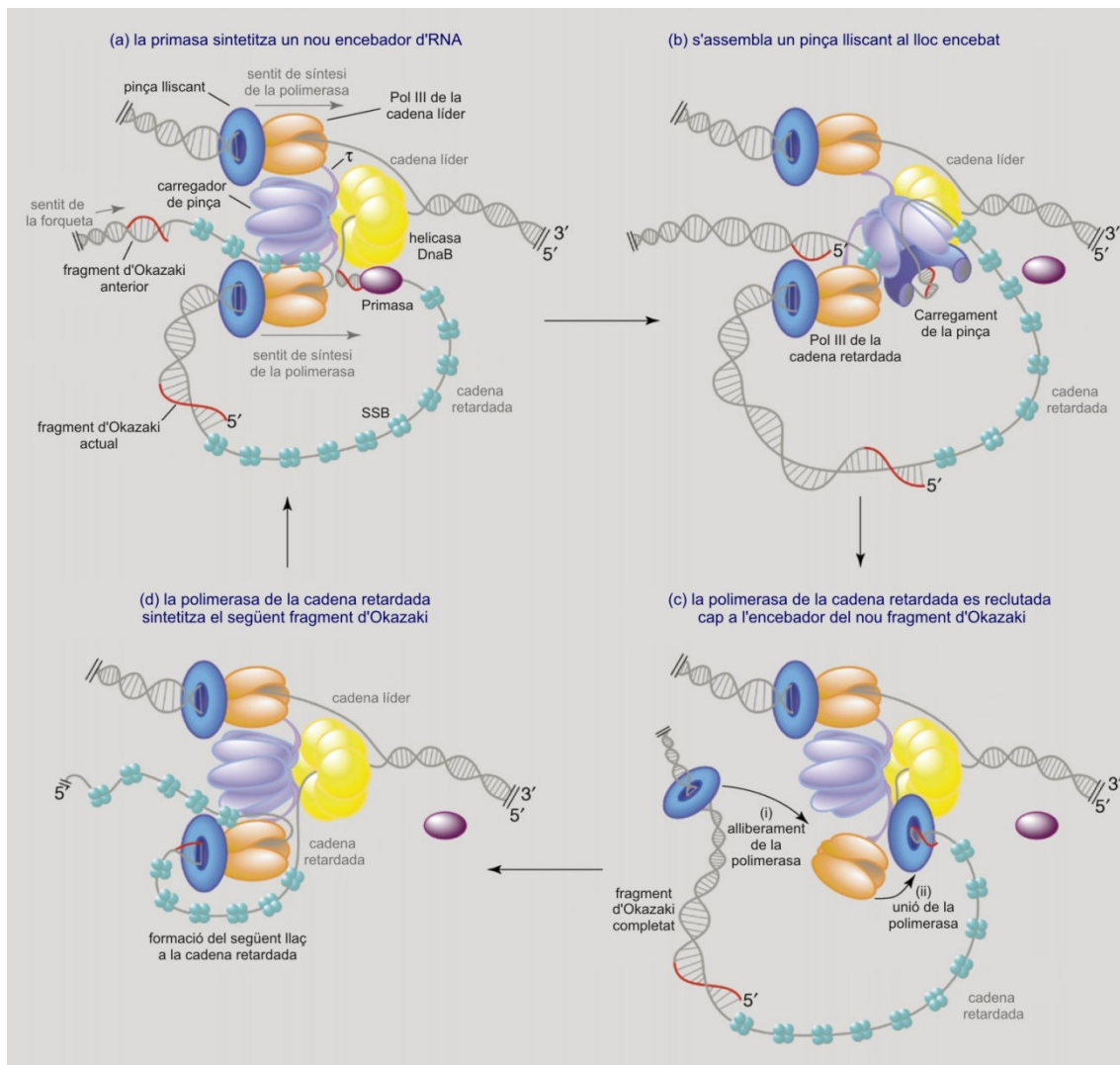


Figura I.25: Esquema de la progressió d l'etapa d'elongació de la replicació a *E. coli*.

La síntesi a la cadena líder és un procés contínu que només necessita que la primasa sintetitzi un o pocs encebadors. La cadena retardada es sintetitza de manera discontinua, en múltiples fragments d'Okazaki. **(a)** La síntesi d'un encebador de RNA per part de DnaG inicia el cicle d'un nou fragment d'Okazaki a la cadena retardada. Un cop sintetitzat l'encebador, la primasa es dissocia del DNA motlle. **(b)** El complex γ carregador de pinça empra l'energia de l'ATP per obrir i assemblar la pinça β al lloc que acaba de ser encebador per la primasa. El carregador de pinça es dissocia del DNA després de l'assemblatge de la pinça, deixant β a l'extrem 3' de l'encebador. **(c)** La polimerasa de la cadena retardada es dissocia de la pinça β després d'haver acabat un fragment d'Okazaki (i) i es recicla cap a una nova pinça β que es troba al lloc encebador (ii). La primera pinça β roman unida al fragment d'Okazaki acabat de completar. **(d)** La polimerasa de la cadena retardada inicia la síntesi de DNA del fragment d'Okazaki. Adaptat de Pomerantz *et al.*, 2007.

OBJECTIUS

L'objectiu principal ha estat la determinació de l'estructura tridimensional de l'helicasa DnaB bacteriana mitjançant les tècniques pròpies de la cristal·lografia de proteïnes. Es pretenia aprofundir en el funcionament d'aquesta proteïna en concret i de les helicases hexamèriques en general.

Per aconseguir aquest objectiu principal, es van marcar diferents objectius parcials, els quals es descriuen a continuació:

- Aïllar el gen codificant per a la proteïna DnaB de l'organisme font. En aquest treball es varen emprar dos organismes fonts diferents: *Aquifex aeolicus* i *Thermotoga maritima*. Ambdós són bacteries termòfiles i tenien el seu genoma completament seqüenciat quan es va iniciar aquest projecte.
- Produir quantitats elevades de proteïna DnaB gràcies a la seva sobreexpressió en cèl·lules hoste d'*Escherichia coli*.
- Purificar la proteïna recombinant fins obtenir-la en un grau de puresa suficient per a poder cristal·litzar-la.
- Obtenir cristalls de proteïna de qualitat suficient per generar dades de difracció de raigs X de resolució i qualitat adequada.
- Processar les dades de difracció i utilitzar-les per resoldre l'estructura tridimensional de la proteïna.
- Analitzar i validar l'estructura obtinguda, per intentar extreure'n un model de funcionament per a l'helicasa.

MATERIALS I MÈTODES

Per a una major claredat, la secció de materials i mètodes està organitzada en dos grans apartats, corresponents als dos conjunts diferenciats de metodologies que s'han tractant durant la realització d'aquest treball.

El primer apartat (A) comprèn les tècniques pròpies d'un laboratori de biologia molecular i està subdividit en tres blocs, els quals descriuen, respectivament, el clonatge, l'expressió i la purificació dels constructes proteics estudiats.

El segon apartat (B) descriu les tècniques específiques de cristal·lografia de proteïnes i està subdividit en dos blocs. El primer bloc descriu les tècniques de cristal·lització de proteïnes. El segon bloc comprèn la recollida de dades de difracció de raigs X i la resolució de l'estructura tridimensional de la proteïna d'interès.

A. TÈCNIQUES DE BIOLOGIA MOLECULAR

Aquesta secció comprèn les tècniques de biologia molecular emprades per produir les proteïnes objecte d'estudi en quantitats suficients i de forma soluble i homogènia, requisit previ a les tècniques de cristal·lografia de proteïnes.

Un cop escollida la proteïna que es vol caracteritzar estructuralment, primerament cal aïllar el seu gen del genoma de l'organisme font, introduir-lo en un sistema que permeti la sobreexpressió de la proteïna en un organisme hoste, realitzar l'expressió en les condicions òptimes per a la correcta transcripció i traducció de la proteïna d'interès i, finalment, separar-la de la resta de proteïnes produïdes per l'hoste. Així doncs, es poden distingir tres grans etapes en aquest procés: el clonatge, l'expressió i la purificació de la proteïna objecte d'estudi. Conseqüentment, aquesta secció està subdividida en tres grans blocs, corresponents a cadascuna d'aquestes etapes. Hi ha una petita introducció a cada bloc, on s'explica breument l'estratègia global seguida a cada pas.

I. CLONATGE DE LES PROTEÏNES D'INTERÈS

En aquesta secció es descriuen les tècniques de biologia molecular emprades per aïllar els gens codificants per a la proteïna DnaB dels dos termòfils estudiats (*Aquifex aeolicus* i *Thermotoga maritima*) i introduir-los als vectors d'expressió escollits, generant així els constructes que es van utilitzar en la sobreexpressió d'ambdues proteïnes. Igualment, es descriu l'obtenció dels constructes de la forma curta i del domini C-terminal de la DnaB d'*Aquifex aeolicus*.

Es van utilitzar tècniques de clonatge clàssiques basades en reaccions de digestió per endonucleases i posterior lligació dels productes. L'amplificació dels fragments de DNA d'interès es va fer mitjançant la tècnica de la reacció en cadena de la polimerasa (PCR).

MATERIALS

DNA motlle

DNA genòmic d'*Aquifex aeolicus VF5* i de *Thermotoga maritima MSB8*. S'obtingué del laboratori del Prof. Dr. Michael Thomm, de la Universitat de Regensburg.

Vectors

❖ pGFPCR.

Vector de clonatge que incorpora com a marcador el gen codificant per la proteïna fluorescent verda. El mecanisme de clonatge es descriu a mètodes.

❖ Vectors pET28a, pET29a.

Vectors comercials de sobreexpressió de proteïnes recombinants, de la casa Novagen. Les característiques d'aquests vectors es descriuen en detall a la secció de materials de la part d'expressió.

Enzims

❖ Polimerasa Vent_R[®].

De la casa comercial New England Biolabs. És una polimerasa de DNA d'alta fidelitat de còpia gràcies a la seva activitat correctora en sentit de 3' a 5'. Genera productes d'extrem roms. Aïllada del termòfil *Thermococcus litoralis*.

❖ Lligasa de T4.

De la casa Roche. És una lligasa de DNA, per tant catalitza la formació d'enllaços fosfodiester entre el grup hidroxil 3' d'un nucleòtid i el grup fosfat 5'

d'un altre. En les reaccions de lligació dins d'un protocol de clonatge, els nucleòtids lligats pertanyen a fragments de DNA lineals diferents, de manera que aquesta reacció serveix per unir-los en un sol fragment i generar nous constructes circulars (i, per tant, estables a l'interior de les cèl·lules bacterianes).

❖ **Enzims de restricció.**

Endonucleases NcoI, NdeI, SmaI, XbaI, XmaIII i XhoI. Totes de la casa comercial Fermentas.

Les dianes de tall, així com les característiques específiques de cada enzim, es resumeixen a la taula M.1.

| Enzim | Diana tall | T ^a òptima | Inactivació | Tampó (alternatius) |
|---------------------|--------------------------------|--------------------------|-------------|------------------------|
| NcoI | 5'-C [^] C A T G G-3' | 37° C | 65°C 20' | Tango 1X (Tango 2X, R) |
| NdeI | 5'-C A [^] T A T G-3' | 37° C | 65°C 20' | O (R, Tango 2X) |
| SmaI | 5'-C C C [^] G G G-3' | 30° C | 65°C 20' | Tango 1X |
| XbaI | 5'-T [^] C T A G A-3' | 37°C | 65°C 20' | Tango 1X (Tango 2X, B) |
| XmaIII [*] | 5'-C [^] G G C C G-3' | 37° C | 65°C 20' | Tampó específic |
| XhoI | 5'-C [^] T C G A G-3' | 37° C | 80°C 20' | R (Tango 2X) |

Taula M.1: Característiques dels enzims de restricció emprats durant el clonatge. Es mostra la seva seqüència de reconeixement, marcant el punt de tall amb el símbol “^”; la temperatura a la qual presenten una activitat òptima; el sistema d'inactivació i el tampó de la casa comercial on la seva activitat és màxima (100% per la primera opció, de 50 a 100% als alternatius). (*) Xma també s'anomena Eco521.

Encebadors

Els oligonucleòtids que s'empraren com a encebadors foren sintetitzats i purificats per la casa comercial Roche, que els serveix liofilitzats. Per a la seva utilització foren resuspesos en aigua MiliQ autoclavada, a una concentració final de 20mM.

La seqüència exacta i característiques de cada encebador es descriuen a la secció de mètodes, on també s'explica el seu disseny.

Kits comercials per a la purificació de DNA

❖ QIAprep Spin Miniprep kit.

De la casa Quiagen. Kit per a la realització de preparacions plasmídiques a petita escala a partir de cultius de nit bacterians.

❖ Ilustra GFX™

De la casa GE Healthcare. Kit per a la purificació i/o concentració de fragments de DNA de llargàries compreses entre 50 pb i 10 kb procedents de reaccions de PCR, digestions amb enzims de restricció o bandes retallades de gels d'agarosa.

Soca de clonatge

S'utilitzà la soca *E. coli* DH5 α de la casa comercial Invitrogen. El seu genotip és el següent: F⁻ ϕ 80 Δ lacZ Δ M15 Δ (lacZYA-argF)U169 *deoR recA1 endA1 hsdR17*(r_k⁻ m_k⁺) *phoA supE44 λ thi-1 gyrA96 relA1*. I presenta com a característiques rellevants per al clonatge:

- una major estabilitat dels inserts gràcies a la mutació *recA1*;
- un augment del rendiment i la qualitat de les minipreparacions de DNA plasmídic aïllat d'aquesta soca gràcies a la mutació *endA1*;
- la presència del marcador ϕ 80 Δ lacZ Δ M15, el qual permet la complementació α del gen de la β -galactosidasa i, per tant, el cribatge dels clons recombinants per diferenciació entre fenotips blaus i blancs.

MÈTODES

1. Determinació de les seqüències dels diferents constructes

La seqüència completa del gen codificant per a la proteïna DnaB d'*Aquifex aeolicus* VF5 (*aq_dnab* en endavant) es va obtenir de la base de dades per a genomes microbians complerts *Genome information broker* (GIB), accessible a l'adreça [web http://genome.nig.ac.jp](http://genome.nig.ac.jp). Aquesta seqüència es va analitzar amb el programa Webcutter2 (<http://rna.lundberg.gu.se/cutter2>) per determinar les dianes de restricció presents i així poder-les descartar per al clonatge.

Similarment, la seqüència completa del gen codificant per a la proteïna DnaB de *Thermotoga maritima* MSB8 (*th_dnab* en endavant) també es va obtenir de la base de dades GIB i es va analitzar amb el programa Webcutter2 per descartar l'ús per al clonatge de les dianes de restricció presents a la seqüència.

La seqüència codificant per al domini C-terminal d'Aq_DnaB es va inferir dels anàlisis d'alineament de seqüència de les diferents DnaB bacterianes descrites, emprant com a referència les dades bioquímiques i estructurals sobre els dominis de la proteïna DnaB descrites a la bibliografia, així com les prediccions d'estructura secundària i desordre per a Aq_DnaB. Per als alineaments de seqüència amb totes les proteïnes homòlogues a DnaB descrites s'emprà el programa PSIBlast (disponible *online* a <http://www.ebi.ac.uk/Tools/psiblast>) contra la base de seqüències proteïques UniProt. Per als alineaments entre les DnaB d'espècies bacterianes concretes, s'emprà el programa ClustalW (disponible *online* a <http://www.ebi.ac.uk/Tools/clustalw2>). Per a les prediccions d'estructura secundària s'empraren els programes: *Secondary Structure Consensus Prediction* de la plataforma bioinformàtica NPS@ (*Network Protein Sequence @nalysis* a <http://npsa-pbil.ibcp.fr>), PSIPRED (*Protein Structure Prediction Server* a <http://bioinf.cs.ucl.ac.uk/psipred>), JPred (<http://bioinf.cs.ucl.ac.uk/psipred>) i PredicProtein (<http://www.predictprotein.org>). Per últim, la predicció de desordre es realitzà amb els programes: FoldIndex (<http://bioportal.weizmann.ac.il/fldbin/findex>), DisEMBL (<http://dis.embl.de>), DISOPRED (<http://bioinf.cs.ucl.ac.uk/disopred>), Predict

Protein (<http://www.predictprotein.org>), PrDOS (<http://prdos.hgc.jp/cgi-bin/top.cgi>) i RONN (<http://www.strubi.ox.ac.uk/RONN>).

El constructe de la forma curta de la DnaB d'*A. aeolicus* també es va dissenyar en funció dels alineaments de seqüència i prediccions d'estructura secundària i desordre, emprant els mateixos programes que en cas anterior.

2. Disseny dels encebadors (*primers*)

Els encebadors pel constructe d'Aq_DnaB foren dissenyats amb dos objectius, que permetessin l'obtenció de la seqüència genòmica completa codificant per aquesta proteïna, alhora que incorporessin les dianes dels enzims de restricció que permetrien la seva posterior inserció al vector d'expressió. Aquest segon requisit és comú per a tots els constructes dissenyats.

Per a la forma curta d'Aq_DnaB, es dissenyà un encebador *reverse* addicional que permetia incorporar un codó de parada (*stop*) just darrera de la E439, de manera que els 29 últims aminoàcids de la proteïna no fossin transcrits. En endavant, aquest constructe s'anomenà Aq_DnaB [1-439], fent referència als aminoàcids que comprèn.

Per al domini C-terminal d'Aq_DnaB, es va emprar una parella d'encebadors que incorporessin un codó d'iniciació al residu T153 de la proteïna original (encebador *forward*), i un codó de parada a la mateixa posició (E439) que a la forma curta (encebador *reverse*). Aquest constructe s'anomenà Aq_DnaB-CT, on CT fa referència a que es tracta del domini C-terminal de la proteïna.

Finalment, per aïllar *th_dnab* es va dissenyar una parella d'encebadors que permetessin amplificar la seqüència codificant completa.

Un cop dissenyats, les seves propietats de cadascuna de les parelles d'encebadors van ser estudiades amb el programa NetPrimer (disponible *online* a www.premierbiosoft.com/netprimer), per tal de descartar dissenys amb alta propensió a

formar estructures secundàries tant intramolecularment com amb l'altre membre de la parella. Es prioritzaren aquelles parelles d'encebadors que presentaven temperatures de fusió similars, tenien com a mínim 10 bases correctament aparellades a banda i banda de la zona desaparellada (que incorpora la diana de restricció i/o els codons de parada generats de nou) i que acabaven preferentment en una base de G o C (per augmentar l'especificitat de la reacció d'amplificació).

A la taula M.2 es mostren les seqüències dels encebadors emprats per a cadascun dels quatre constructes.

| Nom | Seqüència nucleotídica | Diana |
|--------|---|--------|
| Fw-Aq | 5'-CGACTGTAAATTAATTCCATGGCATTGTGGATAAACT-3' | NcoI |
| Rv-Aq | 5'-AGCTACCCCGGATACTCGAGATAAAATTATAAAACC-3' | XhoI |
| Fw-Th | 5'-CGTTTTTTGAGAGGTGAGCATATGCGTGTCCCCC-3' | NdeI |
| Rv-Th | 5'-GTAGGCCTTTATGTCCCGGCCGAACACCTCCCTGG-3' | XmaIII |
| Fw-CT | 5'-GCGGAAAGTGCTACACATATGCAGTTTTACCATGTG-3' | NdeI |
| Rv-CT | 5'-GTGGTGGTGGTGTCTCGAGGTCATTCTAGGTTTGC-3' | XhoI |
| Fw-439 | 5'-CTGTAAATTAATTTACATATGCAATTTGTGGATAAAC-3' | NdeI |
| Rv-439 | 5'-TGGAGGTTGTTCTCGAGGTCATTCTAGGTTTGC-3' | XhoI |

Taula M.2: Encebadors emprats pel clonatge. Fw: encebador a l'extrem 5' de la seqüència a amplificar; Rv: encebador a l'extrem 3' del fragment a amplificar; Aq: constructe Aq_DnaB; Th: constructe Th_DnaB; CT: Aq_DnaB-CT; 439: Aq_DnaB[1-439]. La diana de tall que incorpora cada encebador està subratllada i l'enzim a que correspon indicat. En verd es mostra, el codó d'iniciació i en vermell el de parada. Els encebadors Rv_Aq i Rv-Th no contenen codó de parada perquè, per no incorporar mutacions a la seqüència aminoacídica, són complementaris a una zona riu avall del codó de parada del gen original.

3. Obtenció del gen: amplificació mitjançant PCR

S'utilitzà la tècnica de la PCR (Saiki *et al.* 1988) per aïllar i amplificar les seqüències d'interès. Per a les proteïnes senceres tant d'*Aquifex aeolicus* com de *Thermotoga maritima*, es va partir de DNA genòmic de la soca corresponent. Per a Aq_DnaB [1-439] i per a Aq_DnaB-CT, es va partir del gen sencer *aq_dnaB* clonat al vector d'expressió pET29[Aq_DnaB].

A totes les reaccions de PCR s'emprà la polimerasa de DNA amb activitat correctora Vent_R, utilitzant en cada cas les parelles d'encebadors dissenyades específicament per a cadascun dels constructes. Les parelles d'encebadors han estat descrites a l'apartat anterior (veure taula M.2).

Per tal d'obtenir una banda amplificada única per a cada constructe, es va optimitzar experimentalment tant la composició de la barreja de reacció com els paràmetres d'aquesta, és a dir, la temperatura i durada de la separació de cadenes inicial, la temperatura d'anellament (*annealing*) dels encebadors, la temperatura i durada de la fase d'extensió i el nombre de cicles. A les taules M.3 i M.4 es mostren, respectivament, la composició de la barreja de PCR i els paràmetres de reacció finals per a cadascun dels quatre constructes.

Els productes de PCR es van analitzar mitjançant una electroforesi en gels d'agarosa al 1.2%, que permeten una correcta separació de bandes de longitud compreses entre 0.4 i 6 kb.

Els productes de les reaccions de PCR positives es van purificar mitjançant el kit comercial Illustra GFX, seguint les instruccions del fabricant.

| Components de la barreja de reacció per a la PCR | | | | |
|--|--------------------|--------------------|------------|--------------------|
| | Aq_DnaB sencera | Aq_DnaB [1-439] | Aq_DnaB-CT | Th_DnaB sencera |
| DNA motlle | 0.70 ng/μl | 0,23 ng/μl | 0,23 ng/μl | 0.75 ng/μl |
| Vent _R [®] | 0.001 u/μl | 0.001 u/μl | 0.001 u/μl | 0.001 u/μl |
| <i>Fw primer</i> | 0.75 μM | 0.7 μM | 0.6 μM | 0.625 μM |
| <i>Rv primer</i> | 0.75 μM | 0.7 μM | 0.6 μM | 0.625 μM |
| Barreja de dNTP | 0.2 mM | 0.2 mM | 0.2 mM | 0.2 mM |
| Mg SO ₄ | 2 mM | 4 mM | 2 mM | 3 mM |
| Volum reacció | 100 μl | 100 μl | 50 μl | 100 μl |

Taula M.3: Composició de la barreja de reacció de PCR per als quatre constructes.

| Paràmetres de reacció de la PCR | | | | |
|---------------------------------|--------------------|--------------------|---------------|--------------------|
| | Aq_DnaB sencera | Aq_DnaB [1-439] | Aq_DnaB CT | Th_DnaB sencera |
| Desnaturalització inicial | 4' a 94°C | 2' a 94°C | 2' a 94°C | 4' a 94°C |
| Nombre de cicles: | 30 | 29 | 25 | 30 |
| - Desnaturalització | 1' a 94°C | 1' a 94°C | 1' a 94°C | 1' a 95°C |
| - <i>Annealing</i> | 1' a 58°C | 1' a 60°C | 1' a 58°C | 1' a 55°C |
| - Extensió | 2' 30'' a 72°C | 2' 30'' a 72°C | 2' a 72°C | 2' a 72°C |
| Extensió final | 10' a 72°C | 10' a 72°C | 10' a 72°C | 10' a 72°C |

Taula M.4: Paràmetres de reacció de PCR per als quatre constructes.

4. Lligació dels productes gènics amplificats als vectors d'expressió

Es varen seguir dues estratègies per generar els constructes d'expressió finals. Per a Aq_DnaB es va realitzar un lligació en dues etapes, a través d'un pas intermedi de lligació a el vector de clonatge pGFPCR. En canvi, per a Th_DnaB, Aq_DnaB[1-439] i Aq_DnaB-CT es va lligar directament el producte gènec amplificat al vector d'expressió heteròloga final.

4.1 Lligació a través d'un vector de clonatge

Per a la proteïna Aq_DnaB sencera, degut al baix rendiment de la reacció d'amplificació del gen i a la disminució de l'eficiència de tall de l'endonucleasa NcoI quan es troba a menys de 20 bases de l'extrem del fragment de DNA, es va decidir clonar primer el gen aïllat per PCR al vector de clonatge pGFPCR i, posteriorment, al vector d'expressió definitiu. La lligació al vector pGFPCR presentava l'avantatge de tractar-se d'una lligació entre fragments d'extremes roms i, per tant, no es necessitava la digestió del producte de PCR.

En el primer pas, el producte de PCR purificat es lligava directament al vector pGFPCR en una reacció de digestió i lligació simultània. El vector pGFPCR contenia una diana de tall per a l'endonucleasa SmaI, la qual deixava extrems roms. El producte de PCR també tenia extrems roms, doncs aquesta és una característica de la polimerasa Vent_R. La lligasa present a la mescla catalitzava la reacció de lligació entre els fragments de DNA amb extrems roms. Quan el producte de PCR s'inseria al vector de clonatge, es destruïa la diana per SmaI i aquest constructe ja no podia ser digerit de nou. En canvi, els vectors relligats reformaven la diana i eren tallats de nou, de manera que tornaven a quedar disponibles a la inserció del gen d'interès.

Adicionalment, el lloc d'inserció del fragment de PCR dins el vector pGFPCR trencava el gen que codifica per la proteïna de fluorescència verda (GFP: *green fluorescence protein*), de manera que les colònies transformades amb un vector positiu pel clonatge no eren capaces de sintetitzar la proteïna GFP i romanien blanques al ser exposades a llum ultraviolada. Conseqüentment, al pas final d'aquest protocol de lligació es seleccionaren les colònies blanques i es comprovà per digestió amb l'endonucleasa XbaI, les dianes del qual flanquegen el lloc d'inserció del gen d'interès al vector, que havien incorporat un insert de la mida correcta.

Subsegüentment, les colònies positives foren digerides amb les parelles d'enzims de restricció incorporades durant l'etapa de PCR al fragment gènic amplificat i el producte fou purificat a partir d'un gel d'agarosa. D'aquesta manera estava a punt per una segona etapa de lligació, aquesta vegada al vector d'expressió final, tal i com es descriu al pròxim apartat.

4.2 Lligació al vector d'expressió

Per als constructes Aq_DnaB [1-439], Aq_DnaB-CT i Th_DnaB, es va seguir una estratègia de lligació directe al vector d'expressió, partint directament dels productes de PCR purificats. Aquests foren incubats amb els enzims de restricció que permetrien la seva incorporació en el sentit correcte al vector d'expressió, que també

estava prèviament digerit amb els mateixos enzims. A la taula M.5 es mostra quin vector d'expressió s'utilitzà per cada proteïna.

| Constructe | Vector d'expressió | Dianes d'inserció |
|----------------|--------------------|-------------------|
| Aq_DnaB | pET28a | NcoI / XhoI |
| Th_DnaB | pET27b i pET29a | NdeI / XmaIII |
| Aq_DnaB-CT | pET29a | NdeI / XhoI |
| Aq_DnaB[1-439] | pET29a | NdeI / XhoI |

Taula M.5: Vectors d'expressió i dianes de tall emprades per a cada constructe proteic. La primera diana d'inserció és la de l'extrem 5' del gen, la segona la de l'extrem 3'.

Cal destacar que les parelles d'enzims de restricció escollides tenien tampons de reacció compatibles per als constructes Aq_DnaB, Aq_DnaB[1-439] i Aq_DnaB-CT, cosa que permetia la digestió doble de manera simultània. En canvi, per a Th_DnaB, els enzims no eren compatibles i calia realitzar la digestió de manera seqüencial, purificant el DNA entre ambdues reaccions per tal d'eliminar el tampó del primer enzim, que inactivaria el segon.

Les reaccions de digestió tant dels inserts com dels vectors es varen carregar íntegrament a un gel d'agarosa a 1.2% i la banda corresponent als fragments de DNA digerits va ser retallada del gel i purificada amb el kit Illustra GFX. Les mostres purificades es varen quantificar per densitometria en gel, per a poder calcular la relació insert:vector (3:1 o 5:1) que s'empraria a la reacció de lligació.

S'emprà la lligasa de T4 i un volum de reacció no superior als 20 µl per a tots els casos. La barreja de lligació s'incubà a temperatura ambient durant 2-3h i/o a 4°C durant la nit (*o/n*). Un cop transcorregut el temps d'incubació, es transformà el volum total de reacció a una soca d'*E. coli* de clonatge, en el nostre cas *E. coli* DH5α competents. Aquesta soca no té cap resistència pròpia, de manera que només aquelles cèl·lules que havien incorporat el vector d'expressió podien créixer a una placa d'agar LB suplementada amb l'antibiòtic al qual conferia resistència el plasmidi.

5. Cribatge dels clons positius

Es varen realitzar minipreparacions plasmídiques de les colònies obtingudes de la lligació i es comprovà, mitjançant tècniques de PCR o de digestió amb enzims de restricció, que el vector havia incorporat el gen d'interès. Si es feia una PCR amb els *primers* emprats al clonatge, els clons positius generaven una banda de la mida del gen d'interès. Igualment, si es digerida la preparació plasmídica amb els enzims de restricció emprats al clonatge, també es generava una banda de la mida de l'insert.

Es descartaren aquells clons corresponents a vectors relligats o que contenien un insert de mida diferent a l'esperada. S'escolliren 2 clons positius per a cada constructe i es van seqüenciar per tal de descartar la presència de mutacions a la cadena nucleotídica. La seqüenciació es va dur a terme al Servei de Seqüenciació del CSIC (per a Aq_DnaB i Th_DnaB) i a MacroGen Inc. (per a Aq_DnaB [1-439] i Aq_DnaB-CT).

II. EXPRESSIÓ HETERÒLOGA DELS DIFERENTS CONSTRUCTES DE DNAB A *ESCHERICHIA COLI*

Un dels requisits imprescindibles per l'estudi estructural de proteïnes mitjançant les tècniques de cristal·lografia de raigs X és poder disposar de grans quantitats de proteïna (en l'ordre de mil·ligrams) altament purificada. Aquest requisit fa gairebé indispensable la utilització de sistemes de sobreexpressió proteics, doncs són molt rars el casos en que l'aïllament de la font natural dóna rendiments prou elevats per satisfer les necessitats cristal·logràfiques.

El sistema més comunament emprat per a l'expressió de la proteïna d'interès utilitza com a hoste productora la bactèria *Escherichia coli*. Aquest mètode està especialment indicat en el cas de proteïnes bacterianes, que no requereixen subsegüents modificacions posttraduccionals, doncs és el sistema més fàcil, ràpid i econòmic. Conseqüentment, donat que l'helicasa DnaB és d'origen bacterià, s'utilitzà per a totes les proteïnes sobreexpressades en aquest treball.

A la figura M.1 és descriu esquemàticament el funcionament d'un sistema de sobreexpressió a *E. coli* a través de vectors d'expressió tipus pET amb promotors *T7lac* (Dubendorff *et al.* 1991; Studier 1991), que ha estat el sistema escollit.

Existeixen múltiples soques d'*E. coli* (DE3) amb característiques especials per tal d'afavorir l'expressió de proteïnes altament tòxiques per a *E. coli*, proteïnes procedents d'organismes amb un ús de codó diferent al d'*E. coli*, proteïnes que presenten ponts disulfur i/o proteïnes poc solubles que necessiten l'ajuda de xaperonines per plegar-se correctament. La taula M.7 de l'apartat de materials sumaritza les característiques diferencials de les soques d'*E. coli* utilitzades per expressar les proteïnes objecte d'interès d'aquest treball.

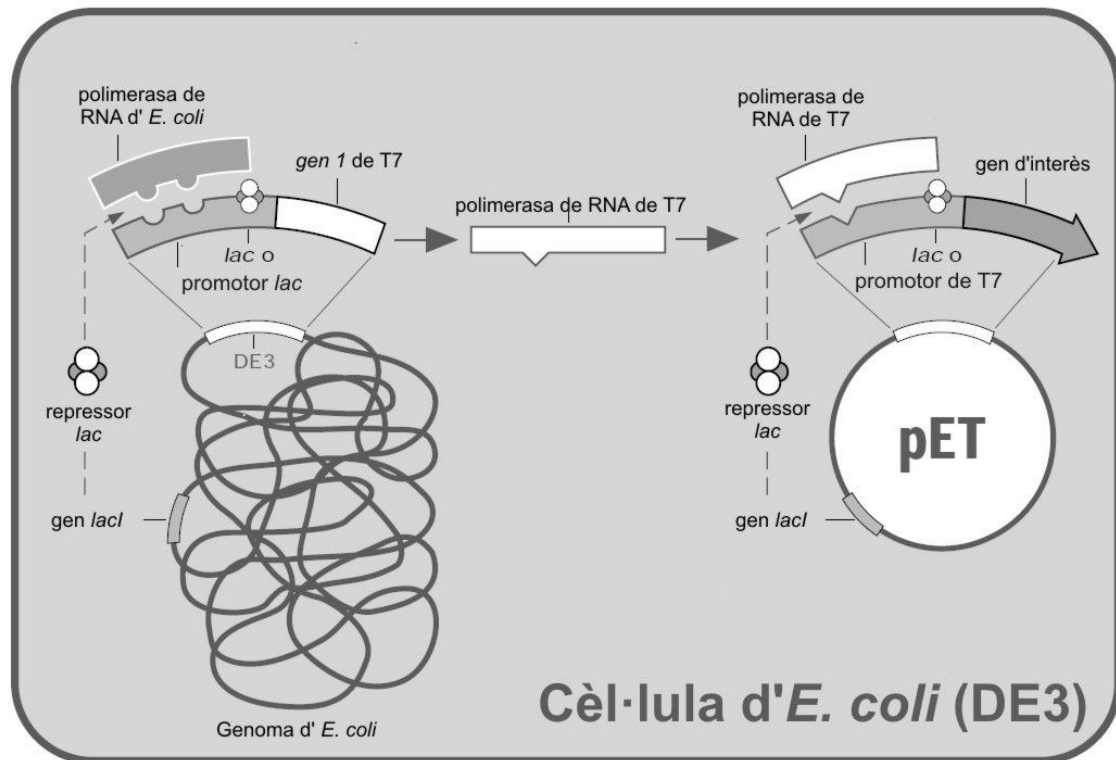


Figura M.1: Funcionament de la sobreexpressió proteica mitjançant vectors pET T7/lac a soques d' *E. coli* (DE3).

El sistema pET posa el gens d'interès sota control d'un promotor de T7, que no és reconegut per la polimerasa de RNA d' *E. coli*. Per tant, per a la transcripció del gen d'interès cal proporcionar a la cèl·lula hoste la polimerasa de RNA de T7. Per a aconseguir-ho, la sobreexpressió es realitza en soques d' *E. coli* lisògenes per al bacteriòfag DE3, les quals porten incorporades al seu genoma una còpia del gen de la polimerasa de T7 sota control del promotor *lacUV5*, induïble per IPTG (un anàleg no hidrolitzable de la lactosa, inductor natural dels promotors tipus *lac*). En una situació no induïda, el repressor *lac*, produït a partir dels gens *lacI* presents tant al genoma bacterià com al propi plasmidi, actua al promotor *lacUV5* del cromosoma de l'hoste per reprimir la transcripció de la polimerasa de RNA de T7 per part de la polimerasa de l'hoste i al promotor T7/lac del vector per bloquejar la transcripció del gen diana per part de qualsevol polimerasa de RNA de T7 que s'hagi sintetitzat basalment. Quan s'afegeix IPTG al medi, aquest s'uneix al repressor *lac*, provocant-li un canvi conformacional que impossibilita la seva unió als promotors *lac*. Conseqüentment, es produeix la síntesi de polimerasa de RNA de T7, que al seu torn transcriurà, a partir del promotor T7/lac, el gen d'interès del plasmidi d'expressió.

Finalment, s'inclou en aquesta secció un apartat addicional que comprèn l'expressió de proteïna recombinant amb els residus de metionina substituïts per selenometionina. És tracta d'una tècnica específica per cristal·lografia de proteïnes que permet obtenir cristalls derivatitzats, que incorporen àtoms de seleni a la seva xarxa cristal·lina, útils per a la posterior determinació de l'estructura tridimensional de la proteïna.

MATERIALS

A. GENERALS

Vectors d'expressió

Tots els vectors d'expressió emprats en aquest treball són comercials, de la casa Novagen, i utilitzen un promotor tipus *T7lac* per controlar l'expressió de la proteïna d'interès. A la taula M.6 es mostren els tres vectors usats i les característiques específiques de cadascun d'ells.

| Vector | Mida (pb) | Resistència | Promotor | Cua a C_term | Diana tall | Cua a N_term |
|---------------|-----------|-------------|----------|--------------|------------|--------------|
| pET27b | 5,414 | Kan | T7lac | pelB | PE | HSV+His(x6) |
| pET28a | 5,369 | Kan | T7lac | His(x6) | Trombina | His(x6) |
| pET29a | 5,371 | Kan | T7lac | S | Trombina | His(x6) |

Taula M.6: Característiques dels vectors d'expressió emprats. Kan: kanamicina; *T7lac*: promotor dual per a la polimerasa de T7 controlat pel sistema *lac*; pelB: pèptid senyal per a l'expressió periplasmàtica de la proteïna d'interès; PE: peptidasa del pèptid senyal; S: proteïna de fusió S per augmentar la solubilitat de la proteïna recombinant; HSV: proteïna de fusió per augmentar solubilitat; His(x6): oligopèptid de 6 histidines per ajudar en la purificació mitjançant cromatografies d'afinitat.

Medis de cultiu bacterians:

❖ Luria Bernani (LB)

Medi ric amb la següent formulació per a un litre: 10 g de triptona, 5 g d'extracte de llevat, 10 g de NaCl. El valor de pH s'ajusta a 7.0 i, finalment, el medi s'esterilitza autoclavant-lo.

❖ **Terrific Broth (TB)**

Medi ric amb la següent formulació per a un litre: 12 g de triptona, 24 g d'extracte de llevat, 9.4 g de KH_2PO_4 i 2.2 g K_2HPO_4 . Valor de pH ajustat a 7.2 i autoclavat.

És una formulació més rica en nutrients que la del medi LB, cosa que permet assolir majors densitats cel·lulars i, consegüentment, això pot afavorir la producció d'una major quantitat de proteïna recombinant.

❖ **Superior BrothTM**

Medi nutritiu de formulació comercial, distribuït per la casa Athena Enzyme Systems. Està dissenyat per augmentar l'expressió de proteïnes recombinants a *E. coli*. La seva formulació complexa proporciona una font de nitrogen, vitamines i nivells de glucosa moderats a les cèl·lules en creixement. Està tamponat a $\text{pH } 7.2 \pm 0.2$ per prevenir l'acidificació del medi deguda als metabòlits produïts pel metabolisme de la glucosa de les bactèries.

Com aquest medi produeix densitats de creixement molt més altes que els medis habitualment emprats per a cultius bacterians, de tres a quatre vegades superiors, cal induir l'expressió de la proteïna d'interès a valors de densitat òptica d'aproximadament el doble dels que s'utilitzarien en cultius crescuts en medi LB.

Solucions d'antibiòtics

❖ **Solució mare de kanamicina a 15 mg/ml.**

Dissolt en aigua destil·lada i esterilitzat per filtració. Emmagatzematge a -20°C . La solució de treball és a 15 $\mu\text{g/ml}$.

❖ **Solució mare de cloramfenicol a 34 mg/ml.**

Dissolt en etanol i esterilitzat per filtració. Emmagatzematge a -20°C . La solució de treball és a 34 $\mu\text{g/ml}$.

❖ **Solució de mare tetraciclina a 5 mg/ml.**

Dissolt en etanol i esterilitzat per filtració. Emmagatzematge a -20°C. La solució de treball és a 12.5 µg/ml.

Solució d'IPTG

Anàleg no hidrolitzable de la glucosa que s'utilitza com a inductor de l'expressió de la proteïna d'interès en sistemes de sobreexpressió regulats per un promotor tipus *T7lac*. Per a més detalls sobre el seu funcionament, consultar la introducció d'aquest apartat.

Tampó PBS (phosphate buffered saline)

Tampó de resuspensió cel·lular i extracció de proteïnes estàndard, àmpliament utilitzat en biologia molecular.

La seva formulació per a un volum final d'un litre és: 8 g de NaCl, 0.2 g de Na₂HPO₄, 1.44 g de KH₂PO₄. El valor de pH s'ajusta a 7.4 i la solució s'esterilitza a l'autoclau.

Soques bacterianes d'expressió

Les múltiples soques d'*Escherichia coli* emprades en aquest treball, així com el seu genotip i les seves característiques fenotípiques més rellevants, es descriuen a la taula M.7.

Totes són soques comercials de les cases Invitrogen, Novagen i Stratagene.

| Soca Bacteriana Genotip | Característiques fenotípiques principals |
|---|--|
| BL21 B F <i>ompT hsdS_B(r_B⁻ m_B⁻) gal dcm</i> | Aquesta soca és deficient en les proteases <i>lon</i> i <i>omp T</i> , cosa que la fa apropiada per l'expressió heteròloga de proteïnes no tòxiques. |
| BL21(DE3) B F <i>ompT hsdS_B(r_B⁻ m_B⁻) gal dcm λ(DE3)</i> | Soca derivada de BL21. Conté el lisògen DE3, el qual porta el gen de la RNA polimerasa de T7 (sota el control del promotor <i>lacUV5</i>) i el <i>lacI^q</i> . Adequada per a l'expressió de proteïnes a partir de vectors tipus pET. |
| BL21(DE3)pLysS B F <i>ompT hsdS_B(r_B⁻ m_B⁻) gal dcm λ(DE3) [pLysS Cam^r]</i> | Soca derivada de BL21(DE3). El plasmidi pLysS codifica pel lisozim de T7, un inhibidor de la polimerasa de T7 que redueix i gairebé elimina l'expressió a nivell basal del gen d'interès abans de l'inducció. Per tant, aquesta soca és apropiada per a l'expressió heteròloga de proteïnes tòxiques. |
| BL21 StarTM B F <i>dcm ompT hsdS_B(r_B⁻ m_B⁻) gal λ(DE3) rne131</i> | Soca derivada de BL21(DE3). La RNA polimerasa de T7 sintetitza mRNA més ràpidament que les polimerases de RNA d' <i>E. coli</i> , de manera la transcripció procedent del promotor de T7 no està acoblada a la traducció a <i>E. coli</i> . Això produeix transcrits de mRNA que no es troben protegits per ribosomes i, conseqüentment, estan subjectes a la degradació enzimàtica per part de RNases endògenes. Aquesta soca està mutada al gen <i>rne</i> (que codifica per la RNasa E), una de les principals fonts de la degradació de mRNA, de manera que es millora l'estabilitat dels transcrits de mRNA i, conseqüentment, l'expressió de proteïna de vectors amb promotors de T7 pot incrementar-se. |
| TunerTM (DE3) F <i>ompT hsdS_B (r_B⁻ m_B⁻) gal dcm lacY1 λ(DE3)</i> | Derivada de la soca BL21(DE3). Mutada al gen de la permeasa <i>lac</i> (<i>lacZY</i>). Aquesta mutació permet una entrada uniforme de l'IPTG a totes les cèl·lules del cultiu, cosa que produeix un nivell d'inducció uniforme depenent de la concentració d'IPTG. Ajustant la concentració d'IPTG es pot regular l'expressió des de nivells molt baixos fins als nivells robustos totalment induïts. Nivells d'expressió baixos poden millorar la solubilitat i l'activitat de proteïnes difícils. |
| OrigamiTM B F <i>ompT hsdS_B(r_B⁻ m_B⁻) gal dcm lacY1 ahpC λ(DE3) gor522::Tn10 trxB (Kan^r, Tet^r)</i> | Soca derivada de Tuner TM (DE3). Té mutacions als gens de la reductasa de tioredoxina (<i>trxB</i>) i de la reductasa de glutatí (<i>gor</i>), les quals faciliten enormement la formació de ponts disulfur al citoplasma. Per reduir la formació de ponts disulfur intermoleculars inespecífics, només es recomana aquesta soca per proteïnes que necessitin la formació de ponts disulfur per a plegar-se correctament. Combina les característiques favorables dels hostes BL21(DE3), Tuner TM i Origami en un sola soca. |
| OrigamiTM B (DE3)pLysS F <i>ompT hsdS_B(r_B⁻ m_B⁻) gal dcm lacY1 ahpC λ(DE3) gor522::Tn10 trxB [pLysS] (Cam^R, Kan^R, Tet^R)</i> | Soca derivada de Origami TM B, que incorpora addicionalment el plasmidi pLysS. |
| RosettaTM (DE3) B F <i>ompT hsdS_B(r_B⁻ m_B⁻) gal dcm λ(DE3) [pRARE Cam^r]</i> | Derivada de la soca BL21(DE3). Ha estat dissenyada per millorar l'expressió de proteïnes eucariotes que contenen codons rars per a <i>E. coli</i> . Aquesta soca proporciona els tRNAs pels codons AGG, AGA, AUA, CUA, CCC i GGA en un plasmidi compatible resistent a cloramfenicol, proporcionant així una traducció "universal" que d'altra manera estaria limitada per l'ús de codó <i>E. coli</i> . Els gens dels tRNAs estan sota el control dels seus promotors natius. |
| Rosetta(DE3) [pLysS] F <i>ompT hsdS_B(r_B⁻ m_B⁻) gal dcm λ(DE3) [pLysSRARE Cam^r]</i> | Soca derivada de Rosetta TM (DE3), que incorpora addicionalment el plasmidi pLysS. |
| Rosetta-gami B F <i>dcm ompT hsdS_B (r_B⁻ m_B⁻) gal lacY1 ahpC λ (DE3) gor522::Tn10 trxB [pRARE (Cam^r, Kan^r, Tet^r)]</i> | Derivada de la soca Origami TM B. Combina una formació facilitada de ponts disulfur al citoplasma gràcies a les mutacions a <i>trxB/gor</i> amb una expressió millorada de proteïnes eucariòtiques que contenen codons rars per <i>E. coli</i> , doncs proporcionen els tRNAs per AGG, AGA, AUA, CUA, CCC, GGA. |
| BL21-CodonPlus[®](DE3)-RIL B F <i>dcm+ ompT hsdS_B(r_B⁻ m_B⁻) Tet^r gal λ(DE3) endA Hte [argU ileY leuW Cam^r]</i> | Soca derivada de BL21 (DE3). Conté còpies extra dels gens dels tRNAs d' <i>E. coli</i> <i>argU</i> , <i>ileY</i> , i <i>leuW</i> . Els tRNAs codificats per aquests gens reconeixen, respectivament, els codons AGA/AGG (arginina), AUA (isoleucina) i CUA (leucina). |
| BL21(DE3)[pT-groE] B F <i>dcm ompT hsdS_B(r_B⁻ m_B⁻) gal λ(DE3)[pT-groE Cam^r]</i> | Soca derivada de BL21(DE3). La presència del plasmidi pT-groE confereix majors possibilitats d'un plegament proteic correcte gràcies a la co-expressió de les xaperonines GroESL. |

Taula M.7: Característiques genotípiques i fenotípiques de les soques d'expressió d'*E. coli* emprades en aquest treball.

B. ESPECÍFICS PER PRODUIR PROTEÏNA AMB SELENOMETIONINA

Solucions:

❖ **Solució d'aminoàcids a 250 mg/ml (5X).**

Es prepara dissolent en 80 ml d'aigua doblement destil·lada 100 mg de cadascun dels següents 18 aminoàcids: Ala, Arg, Asn, Asp, Cys, Gln, Glu, Gly, His, Ile, Leu, Lys, Phe, Pro, Ser, Thr i Val.

A part, es dissolen 100 mg de Tyr i Trp en 400 ml d'aigua tamponada amb fosfat. Cal afegir lentament HCL 1M mentre es manté la solució en agitació fins que aquests dos aminoàcids es dissolguin completament.

Un cop preparades ambdues solucions, es barregen i s'enrasen a 400 ml.

La solució s'esterilitza per filtració a través d'un diàmetre de porus de 0.22 µm.

❖ **Solució d'aminoàcids repressors a 250 mg/ml (5X).**

Esterilitzada per filtració. Per a 6 ml de solució, afegir 100 mg de cadascun dels aminoàcids Thr, Lys i Phe, i 50 mg de cadascun dels aminoàcids Ile, Leu i Val.

❖ **Solució 20% (p/v) de glucosa.**

Esterilitzada preferiblement per filtració.

❖ **Solució 1M de MgSO₄.**

Autoclavada.

❖ **Solució 1M de CaCl₂.**

Autoclavada.

❖ **Solució de tiamina a 10 mg/ml.**

Esterilitzada per filtració.

❖ **Solució de biotina a 5 mg/ml.**

Esterilitzada per filtració. Per tal de que es dissolgui, cal dissoldre primer els mg de biotina necessaris en la meitat de volum, afegir 1M NaOH mentre es manté la solució en agitació fins que aquesta es torni transparent, i llavors ajustar al volum final.

❖ **Solució de SeMet a 60 mg/ml.**

Esterilitzada per filtració. Normalment cal afegir alguna gota de 10M NaOH per ajudar a la solubilització.

Medis de cultiu bacterians:

❖ **Medi M9**

Medi de cultiu mínim, esterilitzat amb l'autoclau. Els ingredients per a un litre de medi 10X són: 80 g de Na_2HPO_4 , 40 g de KH_2PO_4 , 5 g de NH_4Cl i 5 g de NaCl .

❖ **Medi TES**

És una solució d'elements traça, que s'esterilitza per filtració. La formulació d'un litre de medi 100X és de:

- 5 g de EDTA (dissoldre en 800 ml d'aigua destil·lada i ajustar el pH a 7.5, després afegir la resta de components del medi. És important fer-ho així per garantir la solubilització de tots els components.)
- 0.83 g de $\text{FeCl}_3 \times 6 \text{H}_2\text{O}$
- 84 mg de ZnCl_2
- 13 mg de $\text{CuCl}_2 \times 2 \text{H}_2\text{O}$
- 10 mg de $\text{CoCl}_2 \times 6 \text{H}_2\text{O}$
- 10 mg de H_3BO_3
- 1.6 mg de $\text{MnCl}_2 \times 6 \text{H}_2\text{O}$

MÈTODES

1. Proves d'expressió a petita escala.

L'objectiu d'aquestes proves era comprovar l'expressió total de la proteïna d'interès per part dels diferents clons positius obtinguts a l'etapa de clonatge, per així descartar aquells que no produïen la proteïna desitjada i/o seleccionar aquells que tenien uns nivells més alts de sobreexpressió.

El vector d'expressió que contenia el gen d'interès era transformat per xoc tèrmic a la soca bacteriana escollida. S'escollien colònies aïllades per inocular tubs d'assaig amb 4 ml de medi de cultiu LB, suplementat amb l'antibiòtic adequat per a cada constructe. Els cultius eren incubats a 37° C en agitació a 250 rpm fins que assolien una densitat òptica a 600 nm de 0.5, moment en que s'induïa l'expressió de la proteïna recombinant addicionant IPTG a una concentració final de 2 mM.

Per tal de monitoritzar la sobreexpressió proteica al llarg del temps, es recollien alíquotes de 200 µl del cultiu bacterià a diferents intervals, començant pel moment en que s'induïa el cultiu. Les mostres recollides es centrifugaven a 4000 g per separar les cèl·lules del medi de cultiu. Posteriorment, es resuspensia el pellet cel·lular en tampó PBS i es sotmetia la solució resultant a cicles de congelació-descongelació per trencar les cèl·lules i alliberar les proteïnes. Es tornava a centrifugar la mostra i el sobrenedant es carregava a un gel SDS-PAGE per tal d'analitzar les proteïnes presents.

Les condicions d'expressió d'aquestes proves a petita escala eren astringents pel que fa a la temperatura i la concentració d'agent inductor, ja que només interessava determinar si la proteïna recombinant era o no expressada, no optimitzar-ne la seva recuperació en forma soluble. Al emprar condicions astringents, aquestes proves es podien fer en temps molt curts (normalment poques hores de sobreexpressió) i així les primeres valoracions sobre l'expressió d'un determinat constructe s'obtenien molt ràpidament.

2. Proves d'expressió pilot a mitjana escala

Per determinar les condicions d'expressió en que s'obtenia una major quantitat de proteïna soluble, es duren a terme proves d'expressió pilot emprant combinacions de diferents soques bacterianes, medis de cultiu, concentracions d'inductor, temperatura d'inducció i durada d'aquesta. Addicionalment, es varen testar diferents moments d'inducció, a valors de DO_{600} entre 0.4 i 0.6 unitats. Per a una major claredat, així com per evitar repeticions, les combinacions específiques provades per a cadascuna de les quatre proteïnes que s'han expressat en aquest treball es descriuen a l'apartat de resultats.

Aquestes proves d'expressió a escala pilot tenien en comú un volum de cultiu de 200 ml i el protocol de determinació de la fracció proteica recuperada en la forma soluble vers la recuperada de manera insoluble. Es va descartar fer proves emprant volums de cultiu més petits, doncs els resultats obtinguts eren rarament escalables a les expressions a gran escala.

El protocol s'iniciava transformant per xoc tèrmic el vector que contenia el constructe a testar a la soca bacteriana escollida. Tots els medis de cultiu es suplementaven amb l'antibiòtic per al qual conferia resistència el constructe utilitzat. Es picava una colònia aïllada amb la nansa de Kolle per utilitzar-la d'inòcul per a un cultiu de 4 ml en medi LB, que servia de precultiu. Aquest precultiu es creixia a 37° C i en agitació a 250 pm *o/n*. L'endemà, s'inoculava el precultiu de nit a un cultiu de 600 ml de medi LB, el qual es creixia a 37°C i agitació a 250 rpm fins que aquest assolía una $DO_{600} \approx 0.5$. Es refredava el cultiu fins a la temperatura d'expressió (si s'esqueia) i s'induïa amb la concentració d'IPTG escollida. Es mantenia la sobreexpressió a la temperatura desitjada i en agitació durant el temps escollit, agafant alíquotes a diferents intervals. Finalment, els cultius es treien dels incubadors i es recuperaven les cèl·lules per centrifugació.

La taula M.8 descriu esquemàticament el protocol emprat per determinar la fracció de la proteïna total expressada de manera soluble sota les condicions d'expressió testades.

| Protocol per a la determinació de la fracció soluble |
|--|
| ✓ Agafar, en condicions d'esterilitat, una mostra de 25 ml de cultiu |
| ✓ Centrifugar 20' a 4000g per pelletejar les cèl·lules |
| ✓ Resuspendre el sobrenedant en 1ml de tampó de resuspensió |
| ✓ Afegir-hi DNAsa i inhibidors de proteases |
| ✓ Trencar les cèl·lules per sonicació |
| ✓ Centrifugar 30' a 40.000g |
| ✓ El sobrenedant representa la fracció soluble i el pellet les restes cel·lulars juntament amb les proteïnes expressades de forma insoluble. |
| ✓ Carregar una mostra de sobrenedant i pellet en un gel SDS-PAGE per comparar |

Taula M.8: Protocol per a la determinació de la fracció expressada de manera soluble.

3. Sobreexpressió a gran escala

Un cop determinades les condicions òptimes per a l'expressió de cadascun dels constructes, calia realitzar l'expressió a major escala per tal d'obtenir una quantitat de proteïna suficient per a la subsegüent etapa de cristal·lització.

L'escalat de l'expressió es feia en flascons Erlenmeyer de 2 l amb solcs a la seva base i omplerts només amb 600 ml del medi d'expressió escollit, per així assegurar una bona oxigenació. Els cultius es creixien a 37° C i 250 rpm fins assolir la densitat òptica desitjada, moment en el qual s'atempaven, s'induïen amb la concentració d'IPTG prèviament determinada i s'incubaven a la temperatura d'expressió durant el temps adequat.

Un cop finalitzada l'expressió, les cèl·lules es separaven del medi de cultiu per centrifugació a 4500 g durant 30 minuts. El pellet cel·lular resultant bé es congelava ràpidament per submersió en nitrogen líquid i s'emmagatzemava a -80° C fins a la seva utilització, o es resuspensia en tampó de resuspensió cel·lular per al seu processament immediat.

4. Expressió de proteïna amb selenometionina

A continuació es descriu en detall com es va preparar proteïna derivatitzada amb selenometionina emprant una soca d'*E. coli* no auxotròfica, és a dir, la utilitzada normalment per l'expressió de la proteïna nativa. Els primers protocols desenvolupats per marcar les proteïnes amb selenometionina aconsellaven la utilització d'una soca auxotròfica per metionina (Doublet 1997), però aquest requisit s'obvià, doncs amb el protocol presentat a continuació també s'assolien alts índexs d'incorporació de SeMet i la utilització d'una soca normal permetia que el creixement bacterià fos més ràpid. Aquest protocol estava optimitzat -en quan a concentració d'inductor, temperatura i durada de la sobreexpressió- per a l'expressió de les proteïnes Aq_DnaB sencera i Aq_DnaB [1-439].

4.1 Consideracions prèvies

El primer dia calia preparar minicultius bacterians que servien d'inòcul, tants com flascons de cultiu es volguessin emprar en la sobreexpressió a gran escala. Per cada minicultiu, calia picar una colònia aïllada a un tub amb 4ml de medi LB suplementat amb l'antibiòtic adequat. Aquests minicultius es creixien a 37° C *o/n*.

Igualment, era important tenir preparades i autoclavades totes les solucions específiques per a aquest protocol, descrites a materials, així com autoclavar els flascons Erlenmeyer que es volien utilitzar per l'expressió amb 400 ml d'aigua MiliQ a cadascun.

4.2 Protocol

- Per assolir un volum de medi de cultiu final de 600 ml, calia afegir a cada flascó:
- 120 ml de solució d'aminoàcids (5X) filtrada. Si s'observava precipitat a aquesta solució, s'escalfava a 60-70° C durant menys de 5 minuts i s'agitava fins que es redissollessin els precipitats.

- 60 ml de medi M9 (10X) autoclavat
- 6 ml de medi TES (100X) filtrat
- 12 ml de solució de glucosa 20% (p/v) autoclavada
- 0.6 ml MgSO₄ 1M autoclavat
- 0.18 ml CaCl₂ 1M autoclavat
- 1.2 ml biotina a 5mg/ml filtrada
- 0.6 ml tiamina a 10 mg/ml filtrada
- 0.12 ml SeMet a 50 mg/ml
- els ml adequats de l'antibiòtic corresponent a la resistència incorporada al plasmidi d'expressió

Un cop preparats, s'inoculava cada Erlemmeyer amb un precultiu de 4 ml crescut *o/n*. Era convenient eliminar el medi LB en que s'havien crescut aquestes cèl·lules per centrifugació, resuspendre el pellet cel·lular en medi M9 i fer servir aquesta solució com a inòcul.

Els cultius es creixien a 37° C i agitació fins que assolien una DO₆₀₀ = 0.5-0.6, moment en que es refredaven a temperatura ambient abans d'afegir, preservant les condicions d'esterilitat, a cada flascó de cultiu:

- 3.6 ml de la solució d'aminoàcids repressors de la síntesi de metionina.
- 0.6 ml de la solució de SeMet a 50 mg/ml
- 6 ml de glucosa 20% (p/v)
- 60 ml de solució d'aminoàcids (opcional)
- 0.6 ml de biotina a 5mg/ml
- 0.3 ml de tiamina a 10 mg/ml

S'incubaven ls cultius durant 30-40 minuts a 17° C i 220 rpm per esgotar la metionina intracel·lular i després s'induïa l'expressió de la proteïna d'interès afegint 100 µl d' IPTG 1M.

Els cultius induïts s'incubaven a 17° C i 220 rpm durant 5 dies tenint cura de suplementar els cultius d'acord amb la següent pauta:

- al cap d'un dia i mig: amb 6 ml de glucosa a cada Erlemmeyer

- al cap de tres dies: amb 6 ml de glucosa, 0.6 ml de biotina, 0.3 ml de tiamina i 0.5 ml de Se-Met

Finalment, s'efectuava la recol·lecció de les cèl·lules per centrifugació i es purificava la proteïna derivatitzada amb SeMet de la mateixa manera que la proteïna nativa.

III. PURIFICACIÓ DE LES PROTEÏNES D'INTERÈS

Per a realitzar les proves de cristal·lització, es convenient obtenir una mostra de proteïna amb un grau de puresa mínim d'un 95 % (determinat a partir de gels SDS-PAGE tenyits amb el colorant blau de Coomassie) i homogènia (sense presència d'agregats o diferents estats oligomèrics). Per a complir aquests requisits és necessari dissenyar i optimitzar protocols de purificació específics per a cadascuna de les proteïnes estudiades.

El procés de purificació es pot dividir en tres etapes bàsiques: la captura de la proteïna d'interès, la seva purificació preliminar i l'etapa final de poliment (*polishing*). Normalment, també s'inclou dins del protocol de purificació el procés inicial d'extracció i clarificació, durant el qual s'obté l'extracte cel·lular cru (propi del sistema d'expressió emprat) i s'eliminen de la mostra la matèria particulada i altres contaminants que no són compatibles amb les tècniques cromatogràfiques.

L'objectiu bàsic de l'etapa de captura és separar ràpidament la proteïna d'interès de les proteases cel·lulars presents a la mostra inicial, per tal d'evitar al màxim la degradació de la proteïna d'interès, alhora que aquesta és concentrada i estabilitzada. Durant l'etapa de purificació preliminar l'objectiu és eliminar la major part de les impureses presents a l'extracte inicial, ja es tracti d'altres proteïnes, àcids nucleics o endotoxines cel·lulars. Finalment, a l'etapa de poliment l'objectiu és assolir la màxima puresa eliminant qualsevol traça d'impureses que encara es trobin presents a la mostra, així com substàncies relacionades amb la proteïna que es vol purificar.

La purificació es realitza en varies etapes cromatogràfiques successives, durant les quals es combinen les diferents tècniques de separació de proteïnes existents, doncs cadascuna d'elles separa en funció d'una característica concreta de la proteïna d'interès. Per tal de maximitzar el rendiment final de la purificació, cal minimitzar el nombre de

passos de purificació emprats, doncs a cadascun és perd inevitablement una part de la proteïna diana.

En aquest treball, per a la purificació intermèdia i el poliment final de les diferents formes de l'helicasa DnaB estudiades, s'emprà la cromatografia de bescanvi iònic, que permet separar les diferents proteïnes presents a una mostra en funció de la seva càrrega superficial neta al pH de treball, i la cromatografia de filtració en gel, que permet separar en funció del volum aparent de la proteïna. A les figures M.2 i M.3 s'esquematitza, respectivament, el procés de separació en una cromatografia de bescanvi iònic i de gel filtració. Per a l'etapa de captura no s'utilitzà cap tècnica cromatogràfica, sinó un tractament de xoc tèrmic on s'aprofitava que les proteïnes d'interès són termòfiles i han estat sobreexpressades en un organisme mesòfil i, per tant, la resta de proteïnes presents a l'extracte inicial no són resistents a les temperatures altes.

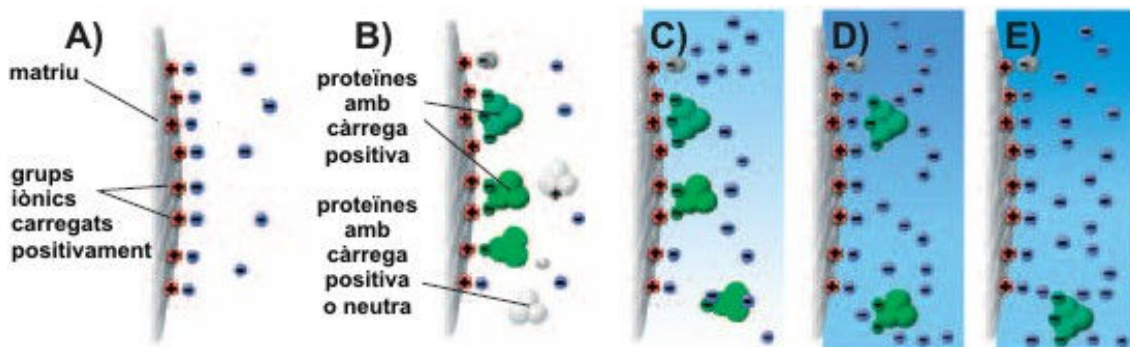


Figura M.2: Principis del bescanvi aniónic. A) S'empra un intercanviador aniónic, és a dir, una matriu amb grups carregats positivament. Inicialment, s'equilibra la columna amb un tampó de baixa força iònica, el pH del qual s'escull en funció de la proteïna d'interès. B) S'aplica la barreja proteica a la columna. Les proteïnes que presenten una càrrega neta superficial negativa ($\text{pH treball} > \text{pI}$), estableixen interaccions reversibles amb els grups carregats de la matriu i queden retingudes a la columna. Les proteïnes amb càrrega neutra o positiva ($\text{pH treball} \leq \text{pI}$), s'elueixen. C) S'aplica un gradient de força iònica creixent per a provocar l'elució de les proteïnes unides a la matriu. A mesura que la força iònica augmenta, els contraions del tampó competeixen cada vegada més fortament per les càrregues de la matriu amb les proteïnes unides. D) A mesura que la força iònica augmenta, les proteïnes amb menor càrrega neta al pH de treball s'elueixen en primer lloc. E) Les proteïnes amb major càrrega neta necessiten forces iòniques més elevades per eluir-se, doncs es troben més fortament retingudes per la columna. Com la relació entre la càrrega superficial neta i el pH de treball és única per a cada proteïna, aquesta tècnica permet una elució diferencial de les diferents proteïnes unides. Es sol escollir un pH de treball al qual la proteïna d'interès quedi retinguda a la reïna i s'elueixi aproximadament a la meitat del gradient salí. Els mateixos principis s'apliquen pel bescanvi catiónic, però amb matrius carregades negativament, que uneixen proteïnes amb càrrega neta positiva.

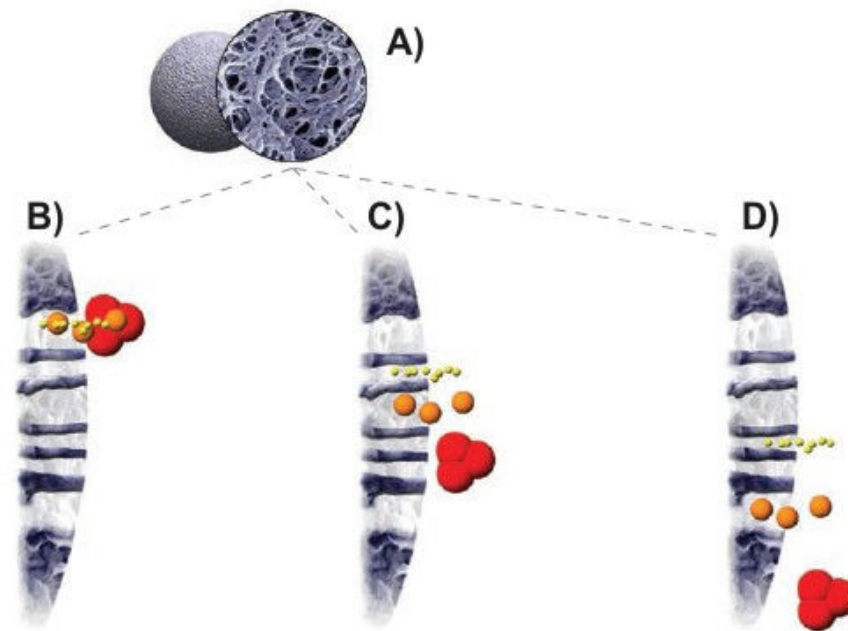


Figura M.3: Procés de gel filtració (exclusió molecular). **A)** Una matriu per a cromatografia d'exclusió molecular està formada per partícules poroses esfèriques inerts, que s'empaqueten dins la columna i són posteriorment equilibrades amb tampó. El tampó omple tant els porus de la matriu com l'espai entre les partícules que la formen. **B)** La mostra s'aplica a la columna. **C)** El tampó i la mostra es mouen a través de la columna. Les molècules presents a la mostra es difonen a través de la matriu en funció de la seva mida. Les molècules més grosses que els porus de la matriu no poden difondre's cap a l'interior d'aquests, passen a través dels espais entre partícules i són eluïdes ràpidament. Les molècules petites es difonen completament cap a l'interior dels porus i són retardades. Les molècules de mida intermèdia poden difondre o no a l'interior dels porus i són retardades en menor mesura. **D)** Com existeixen múltiples capes de partícules poroses que poden retardar o no les molècules en funció de la seva mida, a mesura que el tampó passa a través de la columna es produeix un retard diferencial dels components de la barreja proteica inicial. Els components de major pes molecular són eluïts primer.

MATERIALS

Inhibidors de proteases.

S'emprà la barreja comercial d'inhibidors de proteasa Complete[®] de la casa Roche. Es distribueix en forma de pastilles solubles en medi aquós que contenen un còctel d'inhibidors de proteases efectius contra proteases de serina, proteases de cisteïna, metal·loproteases i calpaïnes. Està especialment indicat per extractes cel·lulars. La concentració final dels inhibidors dissolts en 50ml de tampó és la següent: quimiotripsina 1.5 µg/ml, termolisina 0.8 µg/ml, papaïna 1 mg/ml, pronasa 1.5 µg/ml, extracte pancreàtic 1.5 µg/ml i tripsina 0.002 µg/ml.

Solució de DNasa.

Solució mare a 2 mg/ml en aigua destil·lada emmagatzemada a -20° C.

Membrana de diàlisi:

S'utilitzaren les membranes de cel·lulosa regenerada Spectra/Por[®], de la casa SpectrumLabs, d'un MWCO entre 12.000 i 14.000 Daltons.

Concentradors

Els sistemes emprats en aquest treball concentren per ultrafiltració de la mostra a través d'una membrana de cel·lulosa regenerada quan el sistema es sotmet a centrifugació. La força centrífuga condueix el solvent i els soluts de baix pes molecular a través de la membrana, mentre que els macrosoluts (mida superior al diàmetre de porus de la membrana de cel·lulosa) són retinguts i van incrementant progressivament la seva concentració. Un paràmetre important d'aquest mètode de concentració és el NMWL (*nominal molecular weight limit*), doncs els soluts més grans d'aquest valor queden retinguts pel filtre.

S'empraren 3 tipus diferents de concentradors, depenent dels volums inicial i finals amb els quals es treballava:

- ❖ **Centriprep YM-30 i YM-50.** (NMWL de 30.000 i 50.000, respectivament). De la casa Millipore. Especial per mostres biològiques amb molts soluts, és a dir, quan ens trobem en les primeres etapes de purificació. Per a volums inicials d'entre 5 i 15ml, amb un volum final de la mostra concentrada de 2 ml.
- ❖ **Centricon Ultracel[®] YM-10, YM-30, YM-50** (NMWL de 10.000, 30.000 i 50.000, respectivament). De la casa Millipore. Volum inicial màxim: 2 ml. Volum final: 0.5ml.
- ❖ **Microcon Ultracel[®] YM-30** (NMWL de 30.000). De la casa Millipore. Volum inicial màxim: 500 µl. Volum final: 5-15 µl

Columnes de cromatografia líquida (FPLC)

Totes les columnes emprades són comercials i s'utilitzaren acoblades a sistemes de purificació automàtics ÄKTA (models Basic, Explorer i Purifier) controlats mitjançant el *software* UNICORN. Tant les columnes com els sistemes són de la casa GE Healthcare. A la taula M.9 es resumeixen les característiques funcionals de les columnes.

| Columna | Tècnica | Etapa | Rang de separació |
|------------------------|----------------|--------------------|--------------------------------------|
| Hitrap Q XL 5ml | Bescanvi | Captura/Intermèdia | Proteïnes $pI > pH_{\text{treball}}$ |
| Hitrap S XL 5ml | Bescanvi | Captura/Intermèdia | Proteïnes $pI < pH_{\text{treball}}$ |
| Superose 6 10/300 GL | Gel filtració | Intermèdia/Final | 50 – 600 KDa |
| Superdex 200 10/300 GL | Gel filtració | Intermèdia/Final | 10 – 600 KDa |

Taula M.9: Característiques funcionals de les columnes cromatogràfiques emprades. Es detalla la tècnica cromatogràfica que utilitzen i l'etapa de purificació per la qual es recomana el seu ús, així com el seu rang de separació característic, segons el punt isoelèctric de la proteïna per les columnes de bescanvi iònic i segons el pes molecular en les de gel filtració.

MÈTODES

1. Determinació del tampó de resuspensió cel·lular òptim

El primer tampó de resuspensió cel·lular (I1) utilitzat es va formular d'acord amb allò descrit a la bibliografia prèvia d'expressió i purificació de la proteïna DnaB tant d'*E. coli* (Arai *et al.* 1981a; Fass *et al.* 1999) com del termòfil *Thermus aquaticus* (Kaplan 2000). La composició d'aquest tampó inicial era la següent: 50 mM Tris pH 7.6 10 % (p/v) sacarosa 150 mM NaCl i 5 mM EDTA.

Posteriorment, per tal d'intentar millorar la solubilitat proteica de la proteïna recombinant, es van dissenyar un conjunt de tampons amb diferents valors de pH, força iònica i presència d'agents desnaturalitzants (urea a diferents concentracions) o solubilitzadors (diferents tipus de detergents), amb l'objectiu de determinar experimentalment la composició òptima del tampó d'extracció i decidir si calia fer canvis al tampó inicial. La taula M.10 mostra els diferents tampons testats. Aquestes proves d'optimització es van dur a terme pels constructes només per tres dels constructes amb que s'ha treballat: Aq_DnaB, Aq_DnaB CT i Th_DnaB.

Per a les proves de solubilitat s'utilitzaren mostres de 25 ml de cultiu bacterià procedents de cultius crescuts i induïts sota les mateixes condicions. Aquestes mostres eren processades tal i com es descriurà en el protocol de purificació per al constructe Aq_DnaB (apartat 3.1), però només fins arribar a l'etapa de tractament tèrmic. En aquest punt, es parava la purificació i es carregaven mostres equivalents de la fracció soluble i de la insoluble en gels SDS-PAGE, els quals eren posteriorment tenyits amb el colorant blau de Coomassie per tal de valorar a quina fracció es trobava majoritàriament la proteïna d'interès i en quina proporció. Per tal d'obtenir uns gels amb les bandes proteiques més ben definides, era convenient netejar les mostres de pellet de possibles restes lipídiques amb acetona segons els següent protocol: afegir 2 ml d'acetona, vortexar, congelar o deixar 15' en gel, centrifugar 5' a 30.000 g i eliminar l'acetona pipetejant.

| Tampons d'optimització de l'etapa d'extracció |
|--|
| PH1: 50 mM Na Acetate pH 5, 50 mM NaCl, 5 mM EDTA |
| PH2: 50 mM MES pH 6, 50 mM NaCl, 5 mM EDTA |
| PH3: 50 mM Tris pH 7, 50 mM NaCl, 5 mM EDTA, |
| PH4: 50 mM Tris pH 8, 50 mM NaCl, 5 mM EDTA, |
| PH5: 50 mM Tris pH 9, 50 mM NaCl, 5 mM EDTA |
| S1: 50 mM Tris pH 7.6, 0.1 M NaCl, 5 mM EDTA |
| S2: 50 mM Tris pH 7.6, 0.5 M NaCl, 5 mM EDTA |
| S3: 50 mM Tris pH 7.6, 1 M NaCl, 5 mM EDTA |
| S4: 50 mM Tris pH 7.6, 0.1 M KCl, 5 mM EDTA |
| S5: 50 mM Tris pH 7.6, 1 M KCl, 5 mM EDTA |
| U1: 50 mM Tris pH 7.6, 0.5 M Urea, 50 mM NaCl, 5 mM EDTA |
| U2: 50 mM Tris pH 7.6, 1 M Urea, 50 mM NaCl, 5 mM EDTA |
| U3: 50 mM Tris pH 7.6, 2 M Urea, 50 mM NaCl, 5 mM EDTA |
| U4: 50 mM Tris pH 7.6, 3 M Urea, 50 mM NaCl, 5 mM EDTA |
| U5: 50 mM Tris pH 7.6, 4 M Urea, 50 mM NaCl, 5 mM EDTA |
| U6: 50 mM Tris pH 7.6, 5 M Urea, 50 mM NaCl, 5 mM EDTA |
| D1: 20mM Tris pH 7.6, 50 mM NaCl, 0.2% (p/v) NP 40 |
| D2: 20mM Tris pH 7.6, 50 mM NaCl, 0.2% (p/v) triton X-100 |
| D3: 20mM Tris PH 7.6, 50 mM NaCl, 0.2% (p/v) Tween-20 |
| D4: 20mM Tris pH 7.6, 50 mM NaCl, 0.2% (p/v) dodecil maltòsid |

Taula M.10: Composició dels tampons d'optimització de la solubilitat de la proteïna d'interès durant l'extracció. Es designen amb el prefix PH aquells que testen diferents valors de pH; amb S els de diferent força iònica; amb U els de diferent concentració d'urea i amb D els que contenen detergents.

2. Determinació del protocol de xoc tèrmic òptim

Per realitzar un tractament de xoc tèrmic adequadament és convenient treballar amb tampons de fosfat, doncs aquests presenten un canvi molt petit en el seu pH com a conseqüència de canvis en la temperatura de treball, a diferència dels tampons basats en Tris que poden arribar a presentar variacions de varis punts. Conseqüentment, en aquest treball s'empraren tampons de fosfat de potassi.

Partint d'un cultiu de 200 ml, es van realitzar varies alíquotes de 2 ml, cadascuna de les quals es va sotmetre a un tractament de xoc tèrmic amb una combinació de temperatura i temps diferent (consultar taula M.11). Es va fer un duplicat de cada condició.

| | 50° C | 55° C | 60° C | 65° C | 70° C | 75° C | 80° C | 85° C | 90° C |
|-----|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 5' | | | | X | X | X | X | X | X |
| 10' | | X | X | X | X | X | X | X | X |
| 15' | X | X | X | X | X | X | X | X | X |
| 20' | X | X | X | X | X | X | X | X | X |
| 25' | X | X | X | X | X | X | X | X | |
| 30' | X | X | X | X | X | X | X | | |

Taula M.11: Combinacions de temperatura i durada de les proves d'optimització de l'etapa de xoc tèrmic. En horitzontal es mostren les 9 temperatures diferents provades i en vertical els 6 temps. Es marquen amb una creu les combinacions que es varen testar. Les no marcades no es varen provar perquè eren massa suaus o massa extremes.

Per a la proteïna Aq_DnaB sencera, que va ser la primera amb que es va treballar, es van testar dos tampons diferents per al protocol de xoc tèrmic, amb presència o absència de clorur de magnesi. Per a tots els altres constructes es van emprar sempre tampons amb clorur de magnesi.

Per a la forma Aq_DnaB[1-439], addicionalment es va testar un protocol de xoc tèrmic doble, durant el qual primer es sotmetia la mostra a un tractament tèrmic suau (menys de 60°C), es separaven les proteïnes que es mantenien solubles de les precipitades per centrifugació a 30.000 g durant 30 minuts, i seguidament es realitzava un segon xoc tèrmic més astringent.

3. Protocol de purificació

A continuació es descriuen els protocols de purificació aplicats finalment per a cadascuna de les proteïnes sobreexpressades en aquest treball: Aq_DnaB sencera, Aq_DnaB [1-439], Aq_DnaB-CT i Th_DnaB sencera.

Per a la proteïna DnaB d'*A. aeolicus* sencera i la forma curta d'aquesta, el protocol era molt similar, doncs el segon es va basar en el primer, fent les modificacions pertinents en quant a temperatura i durada del xoc tèrmic. Conseqüentment, la purificació d'ambdues proteïnes es descriurà en un mateix apartat.

Per contra, els protocols de purificació per al domini C-terminal d'Aq_DnaB i la DnaB sencera de *T. maritima* presentaven diferències importants i es descriuran en apartats propis.

3.1 Purificació d'Aq_DnaB i d'Aq_DnaB[1-439]:

Aquest protocol constava d'una etapa de captura inicial mitjançant tractament tèrmic i dues etapes cromatogràfiques successives, una cromatografia de bescanvi iònic i una d'exclusió molecular. Estava dissenyat per a ser realitzat en tres dies consecutius de treball.

3.1.2 Primer dia.

3.1.2.1 Extracció i clarificació

Es partia d'un pellet cel·lular. El pellet procedent d'un litre de cultiu original es resuspensia en 20 ml de tampó de resuspensió R1. Durant aquest procés s'afegia una pastilla d'inhibidors de proteases per cada 50 ml de solució final. Igualment, per facilitar el trencament del DNA genòmic, s'afegia DNasa a raó de 50 µl per litre de cultiu original d'una mare a 2 mg/ml.

Un cop eliminats tots els grumolls cel·lulars, es sotmetia la suspensió resultant a 4-6 cicles de 10'' de sonicació a 0.4 de freqüència i 0.9 de continuïtat i 10'' de descans. La solució es mantenia en gel durant tot el procés de trencament cel·lular per evitar el sobreescalfament de la mostra, que portaria a desnaturalització proteica. Seguidament, la mostra es centrifugava a 40.000 g durant 1 h i es descartava el pellet, que contenia les restes cel·lulars i les proteïnes expressades de manera insoluble en cossos d'inclusió.

3.1.2.2 Etapa de captura mitjançant un tractament de xoc tèrmic

Es sotmetia el sobrenedant procedent de la centrifugació anterior a un tractament tèrmic per precipitar les proteïnes d'*E. coli* (mesòfiles) però no la sobreexpressada, doncs era de termòfil. El tractament de xoc tèrmic es realitzava submergint la mostra completament en un bany calent a 80° C durant 30'. La temperatura i la durada d'aquest tractament es determinaren experimentalment per cada proteïna, tal i com s'ha descrit a l'apartat 2.

Després del xoc tèrmic, es centrifugava la mostra a 45.000 g durant 1 h i s'eliminava el pellet, que contenia les proteïnes precipitades.

3.1.2.3 Preparació de la mostra per a la cromatografia de bescanvi aniònic

El sobrenedant procedent de l'etapa de xoc tèrmic es sotmetia a precipitació del DNA per tractament amb el polímer polietilenimina (PEI) a una concentració final de 0.01 % (v/v). La solució mare de PEI al 10 % (v/v) s'afegia gota a gota i es barrejava immediatament per inversió. S'incubava en agitació a la s'nia a 4° C durant 30 minuts. La solució s'havia de tornar opaca.

Es centrifugava la mostra a 15.000 g durant 1 h. Es descartava el pellet i el sobrenedant es sotmetia a diàlisi per tal de canviar a un tampó basat en Tris-HCl, doncs els fosfat no era adequat per al següent pas de purificació. S'introduïa la mostra dins d'un sac de diàlisi de mida de porus de 12.000-14.000 Da i es dialitzava enfront de 2 litres de tampó A1, a 4° C i en agitació. Al cap d'un mínim de 3 h, es canviava el sac que contenia la mostra a un nou recipient amb 2 l més de tampó A1 i es dialitzava *o/n*, en fred i agitació.

3.1.3 Segon dia: bescanvi aniònic

Acabada la diàlisi, s'extreïa la mostra del sac i es filtrava mitjançant un filtre de xeringa de 0.22 µm de diàmetre per eliminar les restes de PEI abans de carregar-la en

una columna de FLPC, doncs aquest polímer podria provocar problemes de pressió a la columna.

El primer pas de purificació cromatogràfic es realitzava mitjançant una columna de bescanvi aniónic Hitrap Q XL de 5ml equilibrada amb el tampó A1. Un cop carregada la mostra, es feia un rentat amb el tampó A1 de 3 volums de columna o fins que es tornava a l'absorbància base. Llavors s'el·luïa la mostra amb un gradient lineal amb el tampó B1 de 22 volums de columna. Es recolliren fraccions de 3 ml.

Els pics d'absorbància a 280nm s'analitzaren per SDS-PAGE i es van reunir les fraccions més riques en la proteïna d'interès.

3.1.4 Tercer dia: gel filtració

Com el volum resultant d'agrupar les fraccions procedents de la columna de bescanvi era massa gran per carregar-lo a la següent etapa cromatogràfica, es concentrà la mostra per ultracentrifugació fins a un volum final de 1 ml. Un cop concentrada, la mostra s'aplicà en dues punxades, per no disminuir la resolució de la tècnica, a una columna de gel filtració Superose 6 equilibrada amb el tampó G1.

S'analitzaren els pics d'absorbància a 280nm per SDS-PAGE, es concentraren les fraccions on hi havia la proteïna d'interès i es mesurà la concentració final de la mostra pel mètode de Bradford (Bradford 1976).

La identitat i l'homogeneïtat de la mostra purificada d'Aq_DnaB es va comprovar per espectrometria de masses.

Finalment, abans de realitzar les proves de cristal·lització, es canvià el tampó de la proteïna a un tampó mínim (tampó M) per tal que les condicions de cristal·lització tinguessin el màxim efecte sobre la proteïna i en fos més fàcil la seva cristal·lització. També es va filtrar la mostra a través d'un porus de 0.22µm de diàmetre. Aquesta etapa era comuna per a les quatre proteïnes purificades en aquest treball.

A la taula M.12 es mostra la formulació dels tampons emprats en la purificació d'ambdues proteïnes.

| Tampons de purificació per a Aq_DnaB i Aq-DnaB [1-439] |
|---|
| R1: 50 mM fosfat potàssic pH 7.6, 10% (p/v) sacarosa, 200 mM NaCl, 5mM DTT, 1 mM EDTA |
| A1: 50 mM Tris pH 7.6, 10% (p/v) sacarosa, 150 mM NaCl, 5 mM MgCl ₂ , 2mM DTT |
| B1: 50 mM Tris pH 7.6, 10% (p/v) sacarosa, 1M NaCl, 5 mM MgCl ₂ , 2mM DTT |
| G1: 50 mM Tris pH 7.6, 200 mM NaCl, 5 mM MgCl ₂ |
| M: 20 mM Tris pH 7.6, 50 mM NaCl |

Taula M.12: Tampons emprats en la purificació d'Aq_DnaB i Aq-DnaB [1-439]. Tots han estat filtrats a través d'un filtre de xeringa de 0.22µm de diàmetre i desgasats per bomba de buit.

3.1.5 Característiques diferencials de la purificació d'Aq_DnaB[1-439]

Per a la purificació d'Aq-DnaB [1-439], el protocol presentava només un lleuger canvi, doncs la temperatura i durada òptimes per a xoc tèrmic eren inferiors als de la proteïna sencera. En aquest cas, la mostra s'escalfava a 75° C durant 15'.

3.2 Purificació de Th_DnaB

El pellet cel·lular es resuspensia en el tampó R2, afegint inhibidors de proteases i DNasa abans de sotmetre la mostra a sonicació. Els paràmetres de sonicació eren els mateixos que per a les proteïnes descrites anteriorment (apartat 3.1).

A continuació, es separava la fracció soluble de la insoluble mitjançant centrifugació a 40.000 g durant una hora. El sobrenedant es sotmetia a un tractament de precipitació del DNA amb PEI al 0.015 % (v/v), es filtrava a través d'un filtre de xeringa de 0.22 µm i es carregava a una columna HitrapQ XL de 5ml prèviament equilibrada amb el tampó A2. L'elució de la proteïna es feia mitjançant un gradient lineal de 20 volums de columna fins arribar a la concentració salina del tampó B2. Els

pics d'absorbància a 280 nm s'analitzaren per SDS-PAGE i es reuniren les fraccions riques en Th_DnaB.

Aquestes fraccions, un cop concentrades, es carregaven a una columna Superose 6 GL, equilibrada enfront el tampó G2. De nou, s'analitzaren les fraccions mitjançant gels d'acrilamida i es van reunir les fraccions corresponents a la proteïna d'interès.

Finalment, es diluïa la mostra en tampó G2 fins assolir un volum final de 6 ml i es sotmetia a un tractament de xoc tèrmic suau (10' a 65° C). Es centrifugava a 45.000 g durant 1 h, i la fracció soluble es diluïa en 8 ml de tampó M2, es concentrava i es filtrava. La identitat de la mostra es va comprovar per *molecular fingerprinting* al Servei de Proteòmica del Parc Científic de Barcelona.

A la taula M.13 es mostra la formulació dels tampons emprats en aquesta purificació.

| Tampons de purificació per a Th_DnaB |
|---|
| R2: 50 mM Tris pH 7.6, 10% (p/v) sacarosa, 300 mM NaCl, 5mM DTT, 1 mM EDTA |
| A2: 50 mM Tris pH 7.6, 10% (p/v) sacarosa, 200 mM NaCl, 5 mM MgCl ₂ , 2mM DTT |
| B2: 50 mM Tris pH 7.6, 10% (p/v) sacarosa, 1M NaCl, 5 mM MgCl ₂ , 2mM DTT |
| G2: 50 mM Tris pH 7.6, 250 mM NaCl, 5 mM MgCl ₂ |
| M2: 20 mM Tris pH 7.6, 100 mM NaCl |

Taula M.13: Tampons emprats en la purificació de Th_DnaB. Tots han estat filtrats a través d'un filtre de xeringa de 0.22 µm de diàmetre i desgasats per bomba de buit.

3.3 Purificació d'Aq_dnaB CT

El pellet cel·lular es processà igual que en el cas de la proteïna Aq_DnaB (apartat 3.1), fins arribar a l'etapa de diàlisi posterior al tractament per xoc tèrmic.

Per aquesta proteïna es van realitzar diferents proves per determinar el tampó més adequat per a l'etapa de cromatografia de bescanvi iònic, doncs el pH de treball emprat per a la proteïna sencera era massa proper al pI teòric d'Aq_DnaB CT. Es varen testar valors de pH entre 5.5 i 9 i, consegüentment, l'ús tant de bescanvi aniònic com catiònic. La taula M.14 mostra els diferents tampons emprats.

| Tampons de bescanvi iònic per a Aq_DnaB-CT |
|---|
| S1: 50 mM Bis-Tris pH 5.5, 10% (p/v) sacarosa, 100 mM NaCl, 5 mM MgCl ₂ , 2mM DTT |
| S2: 50 mM Bis-Tris pH 6, 10% (p/v) sacarosa, 100 mM NaCl, 5 mM MgCl ₂ , 2mM DTT |
| S3: 50 mM Bis-Tris pH 6.5, 10% (p/v) sacarosa, 100 mM NaCl, 5 mM MgCl ₂ , 2mM DTT |
| S4: 50 mM Bis-Tris pH 7, 10% (p/v) sacarosa, 100 mM NaCl, 5 mM MgCl ₂ , 2mM DTT |
| Q1: 50 mM Tris pH 7.6, 10% (p/v) sacarosa, 100 mM NaCl, 5 mM MgCl ₂ , 2 mM DTT |
| Q2: 50 mM Tris pH 8, 10% (p/v) sacarosa, 100 mM NaCl, 5 mM MgCl ₂ , 2 mM DTT |
| Q3: 50 mM Tris pH 8.5, 10% (p/v) sacarosa, 100 mM NaCl, 5 mM MgCl ₂ , 2 mM DTT |
| Q4: 50 mM Tris pH 9, 10% (p/v) sacarosa, 100 mM NaCl, 5 mM MgCl ₂ , 2 mM DTT |

Taula M.14: Tampons emprats en l'optimització de l'etapa de bescanvi iònic per a Aq_DnaB-CT. Es designen amb la lletra S els tampons amb pHs vàlids per a una columna Hitrap S i amb la lletra Q els tampons amb pHs vàlids per una Hitrap Q. Cadascun dels tampons té la seva parella amb la mateixa formulació però 1M de NaCl, per poder fer el gradient de força iònica.

La mostra procedent de les fraccions més riques en Aq_DnaB CT eluïda de la columna de bescanvi iònic es concentrava i era carregada en una columna Superose 6 equilibrada amb el tampó G3 (50 mM tampó 250 mM NaCl 5 mM MgCl₂), el qual usa el mateix agent tamponador i valor de pH utilitzats durant l'etapa de bescanvi.

La identitat de la mostra es va comprovar per *molecular fingerprinting* al servei de proteòmica del Parc Científic de Barcelona.

B. MÈTODES DE CRISTAL·LOGRAFIA DE RAIGS X

En aquesta secció es descriuen els mètodes propis de les tècniques de cristal·lografia de proteïnes, posant especial èmfasi en aquells aspectes que s'han explorat més durant la realització d'aquest treball.

Un experiment de cristal·lografia de raigs X comença amb l'obtenció de cristalls únics de proteïna, els quals són exposats a un feix de raigs X per tal de recollir-ne el patró de difracció. Aquestes dades de difracció són posteriorment tractades i les fases determinades per tal de reconstruir l'estructura molecular de la proteïna. El model resultant és afinat i validat, abans de procedir a la seva descripció i a la discussió de les seves implicacions biològiques.

Durant aquest procés, hi ha varis punts que poden esdevenir un coll d'ampolla. En aquest treball, el pas limitant ha estat aconseguir cristalls que fossin capaços de generar patrons de difracció de qualitat, a partir dels quals poder procedir a la determinació de l'estructura. Per aquest motiu, s'ha invertit un gran esforç en la millora dels primers cristalls obtinguts i, conseqüentment, s'han explorat sistemàticament un nombre considerable de tècniques de millora de la qualitat dels cristalls descrites a la bibliografia actual. Per això la secció que recull aquestes tècniques és molt extensa.

Igualment, un cop aconseguides dades de difracció i donat la baixa qualitat d'aquestes, s'han explorat diversos mètodes per a la resolució del problema de les fases.

Per contra, les etapes finals de refinament i validació de l'estructura no han estat abordades en aquest treball. Per això no es descriuen en aquesta secció, tot i formar part de la metodologia pròpia de la cristal·lografia de macromolècules.

I. CRISTAL·LITZACIÓ

En aquesta secció es descriuen les tècniques emprades per obtenir cristalls de proteïna, així com els diversos protocols de millora de la qualitat de difracció dels raigs X d'aquests cristalls. Es diferencien els mètodes de millora precristal·lització i els postcristal·lització. Donat que ja ha estat explicat a la secció de clonatge, aquí no s'inclouen com a protocols de millora previs a la cristal·lització la utilització de proteïnes homòlogues d'altres organismes ni el redisseny del constructe proteic per generar un fragment més compacte, tot i que en realitat es tracta d'una aproximació molt eficaç.

També es descriu com obtenir cristalls de proteïna que incorporin àtoms pesats, els quals són imprescindibles per a resoldre l'estructura tridimensional de la proteïna mitjançant mètodes experimentals.

Finalment, es detallen els protocols utilitzats en aquest treball per estabilitzar els cristalls, crioprotegir-los adequadament i col·locar-los en un suport adequat per a la seva congelació, emmagatzematge, transport i, més importantment, per ser sotmesos a un feix de raigs X.

MATERIALS

Plaques de cristal·lització:

❖ **Plaques VDX™.**

De la casa Hampton Research. Plaques de plàstic òpticament transparent, adequades per experiments de difusió de vapor per gota penjant. Formades per 24 pous, al voltant dels quals s'ha aplicat greix per a facilitar el segellat amb el cobreobjectes.

❖ **Plaques Cryschem™.**

De Hampton Research. Plaques de, de 24 pous, al centre de cadascun del quals hi ha un reservori còncav per a una gota assegurada.

❖ **Plaques Greiner-CrystalQuick™.**

De Jena Biosciences. Plaques de poliestirè transparent de 96 reservoris distribuïts en 8 files i 12 columnes. Adequades per a experiments *high-throughput* en nanogota assegurada. Cada reservori conté, a la part superior, tres pouets quadrats de fons pla per a col·locar-hi les gotes assegurades, les quals s'equilibraran contra un mateix reservori.

Solucions de cribatge de les condicions de cristal·lització (*screens*)

Els *screens* emprats en aquest treball segueixen les formulacions de les cases comercials, tot i que es van fer a la Plataforma Automatitzada de Cristal·lografia (PAC) del Parc Científic de Barcelona a partir de solucions mares i emprant el robot de dispensació de volums Tecan. A la taula M.18 mostren totes les solucions de cribatge

que s'han provat durant l'etapa de cristal·lització de les quatre proteïnes estudiades en aquest treball. Les solucions mares es prepararen manualment i totes elles van ser filtrades i suplementades amb àzida sòdica 0.05%.

| Codi PAC | Casa comercial | Formulació |
|----------|--|--|
| LMB1 | Hampton Research | Crystal screen I – II |
| LMB2 | Jena Biosciences | Wizard I –II |
| LMB3 | Hampton Research | Sulfat d'amoni, PEG 6K/LiCl, Quick phosphate, NaCl |
| LMB4 | Hampton Research | PEG 6K, MPD, Membfac |
| LMB5 | Hampton Research | Natrix, Complex screen |
| LMB6 | Hampton Research | Crystal Screen Lite, PEG-Ion |
| LMB7 | Jena Biosciences | Wizard Cryo I – II |
| LMB8 | Jena Biosciences | Jena Biosciences Screen (JBS) 1 – 2 – 3 – 4 |
| LMB9 | Jena Biosciences | JBS 5 – 6 – 7 – 8 |
| LMB10 | Jena Biosciences / Molecular Dimensions | JBS 9 – 10; Clear Strategy I pH 4.5 i 5.5 |
| LMB11 | Molecular Dimensions | Clear Strategy I pH 6.5, 7.5, 8.5; CSII pH 4.5 |
| LMB12 | Molecular Dimensions | Clear Strategy II pH 5.5, 6.5, 7.5, 8.5 |
| LMB13 | Hampton Research | Index |
| LMB14 | Hampton Research | Salt RX |
| LMB15 | Molecular Dimensions | MemStart, MemSys |
| LMB16 | Hampton Research | Screens A/S Ion, Citrat i Mc Pherson |
| LMB17 | Hampton Research | Protein – DNA screen |
| LMB18 | Hampton Research | PEG 400, PEG 4K/LiCl, Àcid Malònic |

Taula M.15: Solucions de cribatge de les condicions de cristal·lització. S'indica el codi adjudicat a cada *screen* per la PAC, la casa comercial que el formulà originàriament i el nom que li donà.

Cobreobjectes OptiClear™

De Hampton Research. Cobreobjectes quadrats (22 mm) de plàstic, lliures d'activitat RNasa i de superfície hidrofòbica. Emprats per dipositar un gota penjant, alhora que per segellar el reservori, durant un experiment de cristal·lització per difusió de vapor.

Vials de congelació

❖ CrystalCap Magnetic™

Com a vials de congelació s'han emprat les bases i tapes de Hampton Research. La base és magnètica, així com el sistema d'acoblament entre la base i la tapa, cosa que permet la seva utilització en els sistemes de muntatge a la base del goniòmetre automàtics de les línees de sincrotró del ESRF.

❖ Pins metàl·lics

Als vials de congelació s'han afegit dos tipus de varetes metàl·liques que difereixen en els sistemes que empren per atrapar els cristalls al seu extrem:

- **CryoLoops™** premuntats de la casa Hampton Research. Empra llaços de niló de diferents diàmetres, que es poden escollir entre valors de 0.025 a 1.0 mm.
- **LithoLoops™** premuntats de la casa comercial Molecular Dimensions. Empra una reixeta d'un material que no interfereix amb la difracció dels raigs X. Els cristalls queden retinguts a aquesta reixeta. És un sistema particularment dissenyat per a poder pescar cristalls fràgils i de mides molt petites i poder centrar-los correctament a les línees de microfocus del sincrotró.

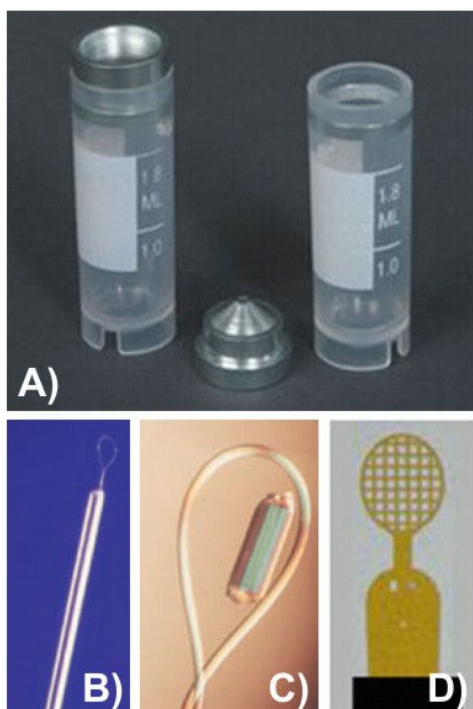


Figura M.4: Vials de congelació.

A) Bases i tapes magnètiques de Hampton Research.

B) Vareta d'acer inoxidable amb el *loop* a l'extrem.

C) *Loops* de niló de Hampton Research.

D) *LithoLoops* en reixeta de Molecular Dimensions

MÈTODES

1. Cristal·logènesi

Les proteïnes poden ser induïdes a formar cristalls si es sotmeten a les condicions apropiades, sota les quals la proteïna purificada pateix una precipitació lenta i ordenada a partir d'una solució aquosa. Durant la cristal·logènesi, les molècules de proteïna individuals s'alineen entre elles en sèries repetitives de cel·les unitats mitjançant l'adopció d'una orientació consistent. La xarxa cristal·lina que es forma és manté unida per interaccions no covalents. Els cristalls de proteïna són inherentment fràgils ja que, com les proteïnes tenen superfícies de formes irregulars, es formen grans canals de solvent a l'interior dels cristalls.

La cristal·lització es dona en dues fases, la nucleació i el creixement del cristall. La nucleació és el pas durant el qual les molècules de proteïna disperses al solvent comencen a reunir-se en grups (*clusters*), els quals esdevenen estables si aconsegueixen assolir una mida crítica. Si no, es redissolen. Aquests grups estables formen els nuclis de cristal·lització. És a la fase de nucleació on les molècules de proteïna es disposen d'una manera definida i periòdica, la qual determina l'estructura de la xarxa cristal·lina. L'aparició dels cristalls, així com el seu augment de mida, es produeix gràcies al subsegüent creixement dels nuclis que han aconseguit assolir la mida crítica. Una de les paradoxes de la cristal·logènesi és que les condicions òptimes per a la nucleació dels cristalls no són les ideals per permetre el seu posterior creixement. Això és degut a que la nucleació espontània és més probable quan els nivells de sobresaturació són elevats, mentre que el creixement lent i ordenat de cristalls grans està afavorit per nivells de saturació més baixos. Per tant, l'experiment de cristal·lització ideal hauria de desacoblar la nucleació del creixement per tal de satisfer els requisits diferencials d'ambdós successos.

1.1 El diagrama de fases de la cristal·logènesi

El procés de cristal·lització es pot il·lustrar gràficament mitjançant un diagrama de fases que mostra quin estat de la proteïna (líquid, cristal·lí o sòlid amorf/precipitat) és estable sota un determinat conjunt de paràmetres de cristal·lització. Aquest diagrama permet quantificar la influència de paràmetres com ara la concentració de la proteïna, precipitant(s), additiu(s), valor de pH i temperatura sobre la producció de cristalls. Per tant, els diagrames de fases constitueixen la base per al disseny de les possibles condicions de creixement de cristalls d'una determinada proteïna. A la figura M.5 es representa esquemàticament un diagrama de fases típic.

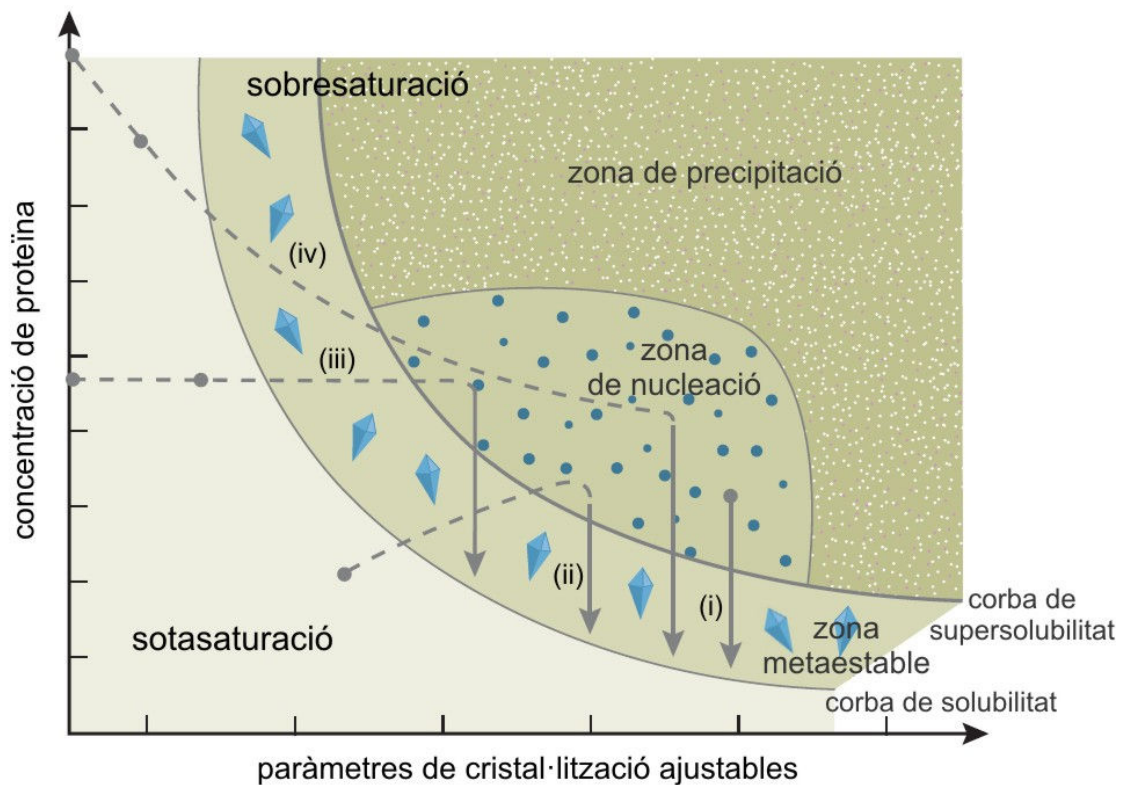


Figura M.5: Diagrama de fases de cristal·lització típic. Dins del gràfic es diferencien quatre àrees: una àrea de molt alta sobresaturació on la proteïna precipitarà; una àrea de sobresaturació moderada on es dona nucleació espontània; una àrea de més baixa saturació, just per sota de la zona de nucleació, on els cristalls són estables i poden créixer però on no es produiran més nucleacions (aquesta àrea es coneix com a zona metaestable i es on es donen les millors condicions per al creixement de cristalls grans ben ordenats); i una àrea de sotasaturació on la proteïna es troba totalment dissolta i, per tant, mai hi cristal·litzarà. Les fletxes representen l'evolució seguida pels diferents mètodes de cristal·lització: (i) *microbatch*, (ii) difusió de vapor, (iii) diàlisis i (iv) difusió d'interfície lliure. Figura adaptada de (Chayen *et al.* 2008)

1.2 Determinació de les condicions de cristal·lització

L'obtenció de cristalls útils per difracció de raigs X depèn d'un gran nombre de factors (puresa de la proteïna, pH, temperatura, concentració de proteïna, presència d'additiu/s, concentració i tipus de precipitant/s). A més, cada proteïna requereix unes condicions particulars per a la seva cristal·lització, les quals són difícils de predir i, a la pràctica, es determinen gràcies a un procés de prova i error durant el qual s'exploren nombroses combinacions d'aquests factors.

Si es coneixen les condicions de cristal·lització d'una proteïna semblant (homòleg d'un organisme diferent o un constructe diferent), es pot començar per variacions entorn a aquesta condició, doncs la nova proteïna pot cristal·litzar en unes condicions semblants. Aquesta aproximació també sol ser vàlida per a proteïna derivatitzada amb selenometionina o complexes amb diferents substrats, en aquest cas prenent com a referència les condicions de la proteïna nativa.

Normalment, però, el procés de cerca s'inicia de zero i llavors s'utilitzen formulacions estàndard (*screens*) que combinen de forma racional els diferents precipitants, sals i tampons usats en cristal·lografia de proteïnes. Aquest procés de cribatge permet determinar unes primeres condicions de cristal·lització (on solen obtenir-se microcristalls o cristalls no aïllats) per a una determinada proteïna, que posteriorment seran optimitzades, ja de forma específica, per assolir cristalls únics i grans. En aquest treball, els *screens* inicials solien posar-se a dues temperatures diferents (20° i 4°C) i, sovint, també a diferents concentracions de proteïna.

1.3 Mètodes de cristal·lització de proteïnes.

Hi ha quatre mètodes estàndards per produir cristalls de proteïna (*microbatch*, difusió de vapor, diàlisi i difusió d'interfície lliure), cadascun dels quals aborda el procés de transició entre la fase de nucleació i la de creixement de manera diferent (veure figura M.5). En aquest treball s'utilitzà exclusivament per a totes les proteïnes la difusió de vapor, que és el mètode més habitual i reproduïble.

1.3.1 Cristal·lització per difusió de vapor

El principi de la cristal·lització per difusió de vapor és molt senzill. Una gota formada per una mescla de solució de proteïna i solució de cristal·lització es col·loca en equilibri de vapor amb un reservori líquid de solució de cristal·lització. Entre la gota i el reservori s'estableix un equilibri per difusió de les espècies volàtils (aigua i dissolvents orgànics) fins que la pressió de vapor entre els dos punts s'iguala. La concentració de l'agent precipitant és major al reservori, doncs a la gota la solució precipitant ha estat mesclada amb la solució de proteïna. Així, l'intercanvi d'aigua es produeix des de la gota cap a la solució del pou, fet que provoca una reducció en el volum d'aquesta, de manera que augmenta progressivament la concentració dels seus constituents. Si s'han aconseguit trobar els percentatges de precipitant i proteïna adequats, la proteïna arriba a la sobresaturació i comença a cristal·litzar quan la gota i el reservori han assolit o estan a punt d'assolir l'equilibri.

La cristal·lització per difusió de vapor es pot realitzar en gota penjant o en gota assegurada, la figura M.6 mostra ambdós muntatges.

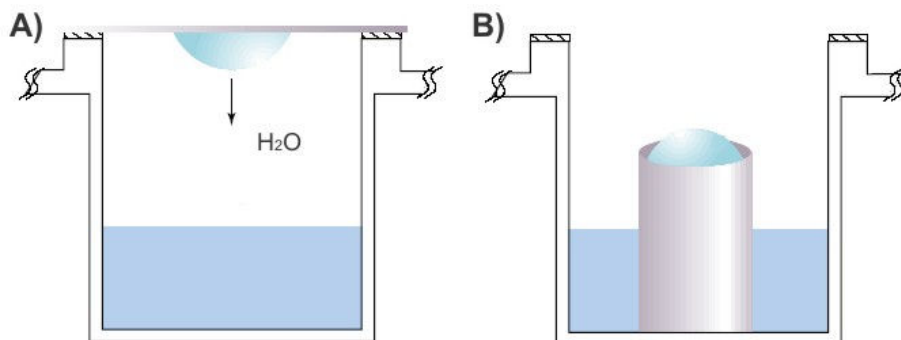


Figura M.6: Cristal·lització per difusió de vapor en gota penjant (A) i gota assegurada (B).

A) La gota amb la mescla de solució de proteïna i de precipitant es diposita sobre un cobreobjectes, que posteriorment es segella damunt del reservori que conté la solució de cristal·lització. **B)** La gota es col·loca en un petit reservori còncav situat dalt d'un pedestal al centre del reservori gran que conté la solució de precipitant. El sistema es tanca segellant la cambra amb cinta adhesiva òpticament transparent. Tant la superfície del cobreobjectes en el cas (A) com del dipòsit còncav del cas (B) és hidrofòbica, per així optimitzar la relació S/V de la gota i afavorir l'intercanvi de vapor. Els volums de gota poden variar dels 2 als 6 μl , amb diferents relacions entre el volum de solució de proteïna i de precipitant. El volum de solució de cristal·lització dipositada al reservori sol ser de 1 ml en el cas de gota penjant i de 0.5 ml en gota assegurada.

En aquest treball, s'utilitzaren els dos muntatges, normalment gota penjant durant el cribatge de les condicions de cristal·lització i gota assegurada durant l'optimització d'aquestes.

1.4 Utilització de sistemes robòtics

La cristal·lització emprant sistemes robòtics utilitza els mateixos components que els descrits a l'apartat anterior, però presenta l'avantatge de que són molt més ràpids i reproduïbles. Es treballà dins l'escala nanomètrica, habitualment amb gotes de 200 nl (100 nl solució proteïna + 100 nl de solució precipitant), equilibrades sobre un reservori que contenia 200 µl de solució de precipitant. Això permetia utilitzar menys quantitat de proteïna purificada durant el procés de cribatge inicial, doncs amb menys de 10 µl de proteïna es podien testar 96 condicions diferents, quan amb la metodologia clàssica això només donava per 10 condicions. A més, al tractar-se de gotes més petites, amb una relació superfície/volum més favorable, el procés de cristal·lització esdevenia molt més ràpid.

El desavantatge principal d'aquesta tècnica és que a vegades les condicions determinades en nanogota no es pogueren escalar fàcilment a microgota i, per tant, no es pogueren generar cristalls de mida suficient.

1.5 Detecció de les condicions de cristal·lització positives

Un cop s'hagueren posat les gotes de cristal·lització, aquestes es deixaren incubar en condicions controlades de temperatura i en absència de vibracions. Per a descobrir la formació de cristalls, es monitoritzaren les gotes sota una lupa a diferents intervals de temps, típicament: cada dia durant la primera setmana, cada setmana durant el primer mes, cada 15 dies durant els tres primers mesos i després un cop cada mes.

En el cas de les plaques de 96 pouets amb nanogotes, aquestes foren introduïdes a la Crystal Farm, on quedaren identificades per un codi de barres únic, s'incubaren a la temperatura escollida i es sotmeteren a anàlisi fotogràfic automàtic als intervals de temps que havia determinat l'usuari. Aquestes fotos eren guardades ordenades

cronològicament pel sistema i es podien consultar *online* des de qualsevol PC, emprant un *software* que també permetia classificar-les manualment en funció de si presentaven o no cristalls, esferolits o precipitat.

2. Tècniques de millora de la qualitat dels cristalls

2.1 Mètodes previs a la cristal·lització.

En un experiment de cristal·lització ideal, un cop s'han format els nuclis, la concentració de proteïna a la solució baixa, portant així naturalment al sistema cap a la zona metastable on es produeix el creixement del cristall, sense que es formin nous nuclis. Ara bé, l'experiment ideal no sempre es dóna i sovint es produeix un excés de nucleació, fenomen que desemboca en la formació de nombrosos cristalls de baixa qualitat. Igualment, a vegades es formen cristalls visualment macos que no produeixen patrons de difracció satisfactoris. Normalment, aquests dos tipus de cristalls no poden ser millorats canviant només els paràmetres de cristal·lització. Per tant, l'objectiu de les tècniques de millora de la qualitat dels cristalls prèvies a la seva formació és dissenyar mètodes que permetin portar l'experiment de cristal·lització de la fase de nucleació a la de creixement, per així garantir la formació de cristalls grans de qualitat.

Hi ha tres estratègies principals de millora precristal·lització: el *seeding*, la dilució i la utilització d'olis. En aquest treball s'han aplicat totes tres i, per tant, a continuació es descriuen en detall.

2.1.2 Separació física de les fases de nucleació i creixement: *seeding*.

Una manera d'evitar els problemes de la nucleació és saltar-se aquesta fase totalment. Això es pot aconseguir inserint nuclis ja formats directament a una nova gota de cristal·lització que es trobi en condicions metastables. D'aquesta manera, s'assegura la presència d'un sol o un nombre limitat de nuclis, que així poden créixer per convertir-se en cristalls únics grans. La millor forma per aconseguir-ho és per "sembrar" (*seeding* en anglès) de la zona metastable amb cristalls petits ja crescuts (de la mateixa proteïna o

a vegades d'una proteïna relacionada), microcristalls, cristalls esmicolats, un precipitat ordenat o fins i tot substrats porosos (Chayen *et al.* 2001). Aquests actuen com a “llavors” per al creixement dels nous cristalls (Bergfors 2003).

Les tècniques de *seeding* es poden classificar en dos grups basant-nos en la mida de les llavors utilitzades:

- *Macroseeding*: transferència d'un sol cristall, normalment de 5-50 µm.
- *Microseeding*: transferència de llavors submicroscòpiques, massa petites per ser distingides individualment.

La selecció del cristall que farà de llavor és molt important tant en el *macroseeding* com en el *microseeding*. Els defectes s'acumulen en els cristalls a mesura que aquests creixen, per tant cal evitar els cristalls més vells i grans com a fonts de llavors. Si no es disposa de cristalls, també es pot intentar fer *seeding* amb qualsevol fase sòlida generada durant el cribatge (olis, esferolits o precipitats cristal·lins), doncs poden presentar algun grau d'ordenació i per tant, actuar com a nuclis per al posterior creixement de cristalls. En aquest casos, s'utilitza la tècnica del *seeding* en estria (veure més endavant).

2.1.2.1 Protocol *macroseeding*

Es van triar cristalls recents no gaire grans. Per tal de crear una superfície de creixement adequada i alhora eliminar les restes de les condicions inicials (que podrien contenir nuclis addicionals), es van sotmetre els cristalls que s'utilitzaren com a llavors a successius rentats amb una solució de menor concentració de precipitant. Amb aquests rentats, els cristalls es fongueren lleugerament i generaren una superfície de creixement fresca. Normalment, es submergia el cristall parental en solució de rentat durant 1 minut i es repetia 2-3 vegades.

Paral·lelament, es varen preparar diverses gotes equilibrades a la zona metastable de la sobresaturació, l'àrea del diagrama de fases on el creixement del cristall està afavorit però no es dona nova nucleació. Un cop rentats, els cristalls eren transferits a aquestes gotes. Com no es solia determinar el diagrama de fases de la

proteïna (consumeix molta proteïna i temps), normalment les condicions d'aquestes gotes es determinaven empíricament disminuint la concentració de la proteïna i/o del precipitant(s) que havien generat els cristalls parentals. Així doncs, es feia un cribatge de condicions òptimes de *seeding* fins trobar aquelles en que els cristalls no es redissolguessin, sinó que creixessin per generar nous cristalls de més qualitat.

Una opció completament diferent de la de baixar les concentracions de proteïna i/o precipitant era sembrar les llavors a unes condicions no relacionades en absolut amb aquelles que permeteren obtenir les llavors. Aquesta tècnica s'anomena *cross-seeding*. La transferència de cristalls a solvents dràsticament diferents de les solucions originals es fa en moltes ocasions per facilitar la incorporació d'àtoms pesats o substrats enzimàtics.

2.1.2.2 Protocol de microseeding

El més crític d'aquesta tècnica era controlar el nombre de llavors que es transferien per tal de que no es produís de nou una nucleació excessiva. Les condicions de les gotes on es sembraren les microllavors es determinaren d'igual manera que en el *macroseeding*. Existeixen diverses tècniques tant per generar les microllavors com per introduir-les a la nova gota, a continuació es descriuen en detall les tècniques emprades en aquest treball:

- *Seeding* en estria

És la manera més fàcil i més ràpida de transferir les llavors, per bé que també la menys reproduïble. Aquest mètode usa un pèl de bigoti animal (normalment de gat o conill) com a vareta de *seeding*, la qual es feia tocar fent pinzellades sobre la superfície del cristall parental per tal de despendre i atrapar els nuclis entre les cutícules solapades que formen el pèl. El pèl llavors es passava a través de la nova gota, dipositant les llavors al llarg de estria que marca el seu recorregut.

- *Seeding* a partir de cristalls polvoritzats

Els cristalls dels quals s'obtidran les microllavors es poden esmicolar emprant homogeneïtzadors de teixits, ionització, vortexant, amb "boletes" (*seed beads*) o varetes

de vidre. En aquest treball s'empraren els *seed beads* de Hampton Research per garantir una correcta polvorització del cristalls i una mostra de microllavors més homogènia.

Un cop s'havia generat un estoc de microllavors, se'n feien dilucions en un tampó estabilitzador per tal d'obtenir solucions amb diferents quantitats de nuclis. La dilució amb el nombre de nuclis òptim s'havia de determinar experimentalment en cada cas, però normalment es trobava entre 10^{-2} i 10^{-7} .

Els nuclis es poden pipetejar de les diferents dilucions de la solució estoc com a petites alíquotes o ser transferits amb una vareta de *seeding* que es submergeix a la barreja de microllavors per tal que n'atrapi algunes i després es fa tocar, es barreja o es mou fent estries per sobre la superfície de la nova gota. En aquest treball, s'emprà preferentment el primer sistema, doncs els resultats eren més reproduïbles.

2.1.2.3 Millorar la qualitat de les llavors: *seeding* serial.

La qualitat dels cristalls emprats com a llavor, tant per macro com per *microseeding*, es pot millorar mitjançant el *seeding* serial (o *seeding* repetitiu). En el *seeding* serial, els cristalls resultants de la primera ronda d'experiments són emprats com a llavors per posar noves gotes. Els cristalls que apareixen en aquestes noves gotes, seran, al seu torn, utilitzats de nou com a llavors per les següents gotes i així successivament. Les llavors poden ser introduïdes a les noves gotes com a macrollavors, per sembra en estria o bé polvoritzant-los per generar microllavors. Els dos últims mètodes foren els emprats més habitualment en aquest treball.

Poden caldre fins a 7-10 rondes de *seeding* per obtenir cristalls de bona qualitat. Normalment, es continuava els procés només si s'observaven millores a les dues primeres rondes.

2.1.3 Limitar el nombre de nuclis a la prova de cristal·lització: separació dinàmica de la nucleació i el creixement

Una alternativa al *seeding*, que presenta l'avantatge de no requerir la manipulació de llavors ni nucleants, és canviar les condicions de cristal·lització per portar l'experiment de la zona de nucleació a la zona metastable abans que es formi un nombre excessiu de nuclis. La nucleació necessita un grau de sobresaturació més elevat que el creixement. Així doncs, es pot aconseguir transferint, després d'un cert temps d'incubació, el cobreobjectes que conté la gota a sobre d'un nou reservori amb menys concentració d'agent precipitant. Aquesta tècnica s'anomena dilució i ha produït increments significatius en l'ordre del cristall en nombroses proteïnes (Saridakis *et al.* 2000; Saridakis *et al.* 2003). Alternativament, hi ha altres mètodes que usen canvis de temperatura entre la fase de nucleació i la de creixement.

2.1.3.1 Protocol de dilució per gota penjant

Es prepararen uns 6-10 reservoris que contenien solucions amb concentracions de precipitants que donarien gotes transparents si aquestes proves de cristal·lització es deixessin incubar sota aquestes condicions. Es cobriren aquests reservoris amb cobreobjectes i es deixaren equilibrar. Paral·lelament, es realitzaren uns 6-10 experiments de cristal·lització sota condicions que produirien molts cristalls de baixa qualitat.

Passat un determinat temps després d'haver posat les gotes de cristal·lització (basant-nos en quan van sortir els primers cristalls en els *screens* inicials), es transferia un dels cobreobjectes a un reservori que contenia una concentració de precipitant inferior. Aquesta transferència havia de durar uns 1-2 segons com a màxim. Després d'un interval de temps superior, es transferia un segon cobreobjectes i després un tercer, un quart i així successivament. Normalment es provaren varis intervals separats per 1-2 hores. Calia deixar una o dues gotes sense transferir i una o dues gotes sota les condicions de baixa concentració de precipitant perquè actuessin com a control.

Les gotes transferides s'incubaren a la temperatura escollida i es controlà l'aparició de cristalls al llarg del temps, tenint en compte que el temps de formació de cristalls seria més llarg a la gota transferida comparat amb l'experiment control que no havia estat transferit, és a dir, al temps que trigaven les condicions de cristallització originals a generar cristalls.

S'esperava que alguns cristalls, a alguns dels temps de transferència i sobre algunes de les condicions de dilució provades foren menys nombrosos i tinguessin un ordre intern major que els cristalls originals. En canvi, les gotes que havien estat transferides massa aviat produïrien gotes transparents, mentre que aquelles que havien estat transferides massa tard donaren cristalls de poca qualitat. Calia tenir en compte aquests resultats per posteriors rondes de millora.

2.1.3.2 Feeding i weeding

Una altra manera de manipular la solució en comptes de les llavors, són els mètodes de *feeding* (en anglès, alimentar) i de *weeding* (literalment, treure les males herbes).

Al *feeding*, s'afegeix més proteïna a la gota, cosa que té dues conseqüències: la gota es dilueix, disminuint la sobresaturació, alhora que hi ha més proteïna disponible per als nuclis en creixement. Aquest era un mètode molt efectiu si els cristalls que havien nucleat a la zona làbil havien exhaurit tota la proteïna dissolta.

Al *weeding*, els microcristalls en excés són tretts de la gota de manera que la proteïna que roman en solució està disponible només als pocs cristalls seleccionats que es conserven. Al realitzar aquest mètode, calia anar molt en compte per evitar l'evaporació a la gota de cristallització. Com això era difícil de controlar, normalment el *weeding* no tingué èxit.

2.1.4 Influenciar la cinètica de cristal·lització: utilització d'olis per alentir la difusió de vapor.

Un dels problemes més comuns de la cristal·lització mitjançant difusió de vapor és que sovint es generen nombrosos cristalls petits en comptes de cristalls grossos aïllats. Això normalment succeeix perquè el procés de cristal·lització es dona massa ràpidament. Per tant, si s'aconsegueix alentir la cinètica de cristal·lització, es podrien generar menys cristalls de major qualitat.

El mètode més senzill i el més àmpliament utilitzat d'influenciar la cinètica en un experiment de difusió de vapor és emprar una barrera d'oli sobre el reservori que conté la solució de cristal·lització per així alentir l'evaporació. No són necessaris canvis en les condicions de cristal·lització ni en els recipients de cristal·lització emprats. Varies proteïnes problemàtiques han estat cristal·litzades amb èxit mitjançant aquest mètode (Mayans *et al.* 1998; Isupov *et al.* 2004; Schubot *et al.* 2005). A la figura M.7 s'esquematitza el muntatge necessari.

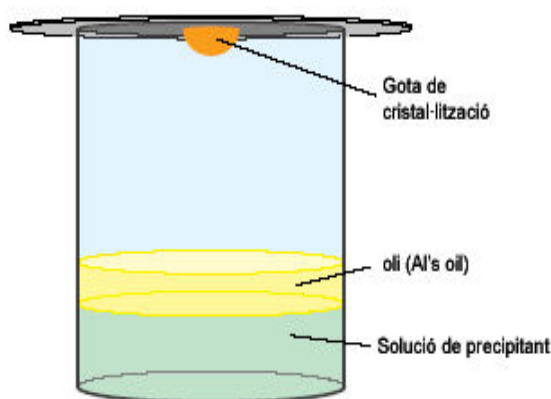


Figura M.7: Utilització d'olis com a barrera per alentir la difusió en un experiment de gota penjant. L'experiment de gota penjant es realitza en una placa normal. Un determinat volum d'oli (barreja d'olis de silicona i parafina) es situa sobre la solució de cristal·lització del reservori, just abans d'invertir el cobreobjectes on s'ha dipositat la gota i segellar així el pou.

El tipus d'oli utilitzat és crític per al resultat final de l'experiment. La parafina no permet una difusió suficient del vapor procedent del reservori, cosa que provoca l'assecament de la gota. Per contra, l'oli de silicona gairebé no suposa cap barrera a la difusió. Normalment, s'empren barreges d'ambdós olis i la taxa de difusió es pot

controlar canviant els ràtios de parafina i silicona a la mescla oleica final. A part del tipus d'oli, un altre factor important per controlar la velocitat de difusió del vapor és el gruix de la capa d'oli que es col·loca sobre la solució del reservori. Es sol variar entre gruixos d' 1-3 mm (corresponents a 0.1-1ml), amb valors entre 0.3 – 0.7 ml sent els més habituals. En aquest treball es va utilitzar l'oli comercial Al's Oil (Hampton Research), el qual es una barreja 50:50 (v/v) de parafina i silicona. El paràmetre que es va variar va ser el volum d'oli addicionat a sobre el reservori.

2.2 Tècniques de millora postcristal·lització

2.2.1 Annealing

Les tècniques de criocristal·lografia (descrites a la següent secció) poden provocar desordres en la xarxa cristal·lina i, consegüentment, un augment de la mosaicitat i una resolució de la difracció reduïda. Aquest dany, que és especialment important en els cristalls grans o amb alt contingut de solvent, es creu que és conseqüència d'un congelat desigual i a la expansió diferencial del solvent i la xarxa cristal·lina.

Si el cristall s'ha congelat malament, es pot sotmetre a *annealing* (ciclat entre temperatures baixes i altes) per provar de reduir la mosaicitat i incrementar la resolució. L'*annealing* es pot dur a terme de dues maneres, bé bloquejant la corrent de gas del criogen temporalment (1-2 segons) o posant el cristall de nou al crioprotector, permetent que s'equilibri i després tornant-lo a congelar ràpidament. És un mètode fàcil i ràpid, durant el qual la concentració de l'agent crioprotector s'ajusta ella mateixa, canviant així les seves propietats tèrmiques de manera que la contracció del solvent quan es refreda es troba més propera a la contracció de la xarxa cristal·lina del cristall. L'*annealing* té més possibilitats d'èxit si la concentració inicial d'agent crioprotector està per sobre, més que no per sota, de l'òptim. Això implica que és millor tenir massa agent crioprotector al tampó de crioprotecció que no pas massa poc, doncs sempre es pot sotmetre el cristall a *annealing*..

2.2.2 Deshidratació dels cristalls

L'aigua juga un paper crucial en el manteniment de l'estructura i l'activitat de les molècules de proteïna tant en solució com en forma cristal·lina. La reducció del contingut de solvent del cristall pot fer que aquest s'empaqueti millor i augmenti el seu ordre intern, estenent així la resolució dels patrons de difracció. Normalment, el contingut de solvent dels cristalls de proteïna ronda el 50%. La deshidratació sol funcionar millor en cristalls amb un contingut de solvent superior a aquest valor.

Els protocols de deshidratació de cristalls són molt diversos i inclouen des de l'assecat a l'aire de la gota que conté el cristall, fins a la transferència gradual d'aquest a solucions de deshidratació. Existeix també un aparell comercial que controla la humitat relativa del cristall (*Proteros Free Mounting System* de Rigaku), el qual permet realitzar protocols de deshidratació (i també de rehidratació) més reproduïbles i eficaços. A la figura M.8 es descriuen esquemàticament els 4 protocols de deshidratació que s'han emprat en aquest treball.

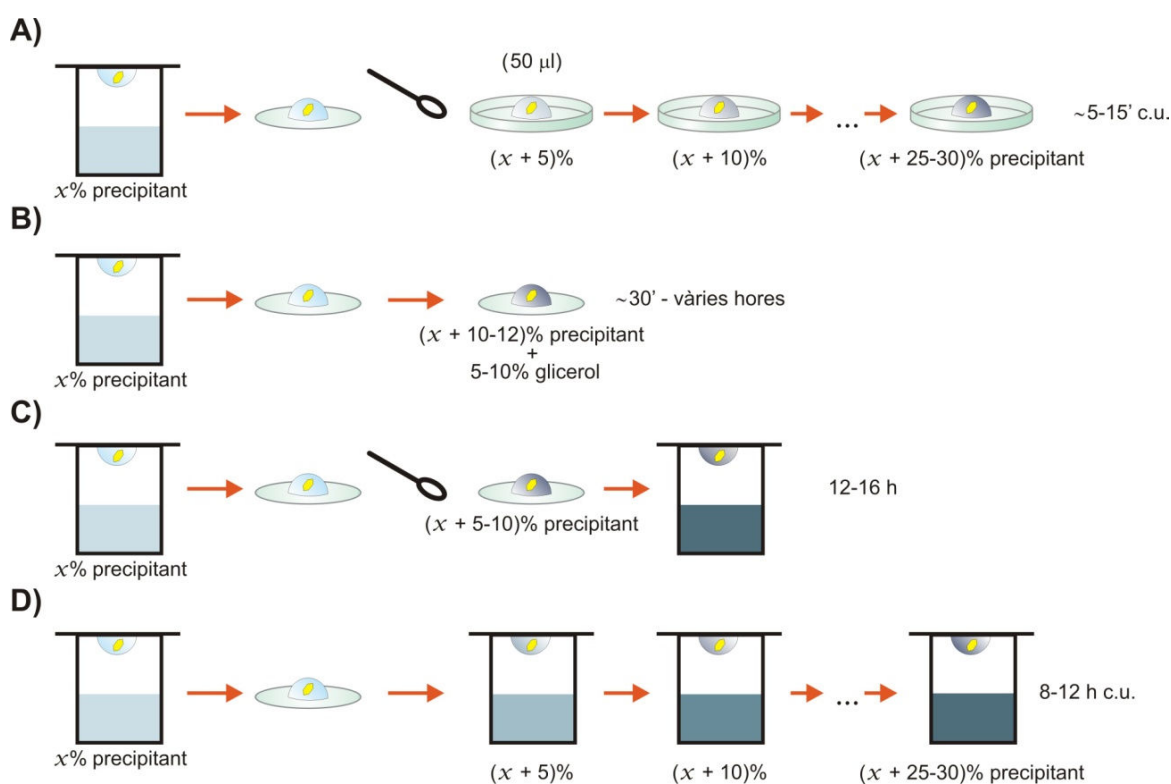


Figura M.8: Mètodes de deshidratació de cristalls.

Veure llegenda a la pàgina següent.

A) Transferir seqüencialment els cristalls de la gota de cristal·lització a gotes de 5 µl que continguin quantitats creixents de solució de deshidratació. Depenent de l'estabilitat del cristall, la concentració d'agent deshidratant es pot incrementar del 5 % fins al 30 % (p/v), i en passos de 0.5 % (p/v) fins a de 5 % (p/v). El temps de *soaking* també pot variar de 5 a 15 minuts (en aquest cas els cristalls són deshidratats a l'aire) a dies (en aquest cas les gotes que s'estan deshidratant s'equilibren contra un reservori que conté la solució deshidratant). **B)** Afegir lentament solució deshidratant a la gota que conté el cristall (aproximadament unes vuit vegades el volum inicial de la gota de cristal·lització) i deixar deshidratar a l'aire durant més de 30'. Com a solució deshidratant s'empra la condició de cristal·lització amb un increment del 10-12 % en la concentració de precipitant i una addició d'un 5-10 % (p/v) d'agent crioprotector, com ara glicerol **C)** Transferir el cristall de la gota de cristal·lització a un gota penjant de 5 µl de solució deshidratant i equilibrar contra un reservori amb la mateixa solució deshidratant. Solució deshidratant: condició de cristal·lització que conté un 5-10 % (p/v) més d'agent precipitant. Temps d'incubació: 12-16 h. **D)** Després del creixement del cristall, equilibrar la gota de cristal·lització contra reservoris que continguin concentracions creixents d'agents deshidratants. Temps d'incubació: 8-12 h a cada pas. Solució deshidratant: solució mare amb concentracions creixents bé de precipitant o de PEG de baix pes molecular, glicerol o MPD. La concentració s'incrementa en passos del 5 %.

La solució de deshidratació emprada era la solució mare amb una quantitat creixent de precipitant. A vegades també es s'emprava la solució mare suplementada amb concentracions creixents d'agents crioprotectors. En aquest cas, la crioprotecció i la deshidratació es fusionaven, doncs, de fet, una estratègia per crioprotegir els cristalls es reduir-ne la solvatació.

Finalment, es provà la combinació de la deshidratació amb altres tractament post-cristal·lització, com ara l'*annealing* o el *soaking*, doncs també pot produir millores en la qualitat de difracció i la resolució dels cristalls de proteïna (Chayen *et al.* 2008).

2.2.3 Submersió a major concentració de precipitant sense deshidratació

El *soaking* postcristal·lització és similar a la deshidratació dels cristalls en el fet que ambdós processos impliquen submergir els cristalls de proteïna en solucions que contenen concentracions creixents de precipitant o crioprotector. De totes maneres, la deshidratació implica la contracció de la xarxa cristal·lina i la disminució del contingut de solvent del cristall mentre que el *soaking* postcristal·lització sense deshidratació no implica un canvi en la cel·la unitat o el contingut de solvent, encara que també pot portar a un notable millora en la qualitat de difracció del cristall. De fet, a la literatura hi ha varis exemples on el *soaking* dels cristalls en solucions de més força iònica,

crioprotectants o solucions que contenen àtoms pesants, donen com a resultat millores en la qualitat del cristall (Heras *et al.* 2005).

2.2.4 Entrecreuament químic de cristalls per difusió de vapor.

S'empra per poder manipular millor els cristalls fràgils durant els canvis de solució i la seva congelació, doncs reforça la xarxa cristal·lina del cristall, fent-lo més resistent a l'estrès mecànic.

En aquest treball, com a agent d'entrecreuament s'emprà el glutaraldehyd, el qual funciona principalment sobre residus de lisina, per tant, el nombre de lisines i la temperatura eren les dues variables a tenir en compte a l'hora de definir el temps de reacció. Les amines interfereixen amb el glutaraldehyd, de manera que, abans de començar el *crosslinking*, s'havia de substituir el sulfat d'amoni i/o els tampons de Tris presents a la mostra de proteïna mitjançant diàlisi.

El cobreobjectes que contenia el cristall es transferia a un nou reservori amb solució precipitant i un micropont (Hampton Research) que contenia una gota assegurada (2-5 µl) de 25 % (v/v) glutaraldehyd a pH 3. S'equilibrava durant 30-60 minuts (en algun cas es va arribar a 6 hores), transcorreguts els quals s'aturava el procés col·locant el cobreobjectes sobre un nou reservori amb solució precipitant fresca (Lusty 1999).

2.2.5 Metilació reductiva de les lisines exposades

És un mètode de millora ràpid i fàcil d'utilitzar, amb el qual s'han obtingut resultats espectaculars en varies proteïnes (Walter *et al.* 2006). Consisteix en l'alquilació reductiva dels grups amino lliures, és a dir, dels residus de lisina i de l'extrem N-terminal de la proteïna, mitjançant una reacció senzilla en dues etapes en presència d'un agent alquilant, una solució de formaldehyd, i un agent reductor, el complex borà-dimetilamina (ABC).

El protocol que s'utilitzà en aquest treball és una modificació del protocol del kit de metilació comercial de Jena biosciences, basant-se en les recomanacions descrites a

l'article (Walter *et al.* 2006). Es detalla on s'han produït canvis per evitar la precipitació de la proteïna durant el procés i augmentar així el rendiment. Totes les manipulacions havien de realitzar-se a 4°C.

Abans de començar la reacció de metilació, s'intercanviava el tampó en que es trobava la proteïna a un tampó que no contingués grups amino lliures i/o alcohols, com ara tampons basats en HEPES o fosfat. Els tampons basats en Tris-HCl interferien en la reacció de metilació i, per tant, no eren adequats. Igualment, calia evitar la presència d'agents reductors com ara DTT, TCEP, β -mercaptoetanol, doncs podrien interferir en la reacció de metilació. Per als diferents constructes de DnaB s'emprà el següent tampó: 20 mM HEPES pH 7.6 100 mM NaCl 5 mM MgCl₂. Igualment, la proteïna era concentrada a ~ 1 mg/ml (per proteïna nativa) o ~ 0.5 mg/ml (per proteïna marcada amb selenometionina).

Seguidament, es realitzava la reacció de metilació tal i com es descriu a la taula M.16. Calia tenir en compte que la proteïna podia precipitar ja en el primer pas de metilació. Si ho feia, calia repetir el procés a una concentració de proteïna menor. La proteïna precipitada era irrecuperable.

| Reacció de metilació de les lisines exposades |
|--|
| ✓ Per a cada ml de proteïna, afegir a la solució de proteïna: |
| - 20 μ l de solució ABC a 1M |
| - 40 μ l de solució de formaldehid a 1M |
| ✓ Barrejar suaument pipetejant varies vegades |
| ✓ Incubar en gel durant dues hores |
| ✓ Afegir 20 μ l de solució ABC a 1M i 40 μ l de solució de formaldehid a 1M |
| ✓ Barrejar suaument pipetejant varies vegades i incubar en gel durant 2 hores |
| ✓ Afegir 10 μ l de solució ABC a 1M |
| ✓ Barrejar suaument i incubar a 4°C durant 20 hores |
| ✓ Aturar la reacció intercanviant el tampó de la solució a un basat en Tris per tal de neutralitzar qualsevol capacitat metiladora que quedés a la solució |

Taula M.16: Protocol de metilació de les lisines exposades.

Després de la reacció de metilació, es sotmetia la mostra a una cromatografia de gel filtració per acabar d'eliminar el tampó i els reactius utilitzats durant la reacció de metilació. Aquesta etapa cromatogràfica també era molt útil perquè donava indicis de si el procés havia funcionat: un sol pic estret al cromatograma indicava la presència d'una sola espècie. Calia tenir en compte que les proteïnes metilades tendeixen a córrer de manera diferent, normalment una mica més ràpidament, en les columnes de gel filtració.

Finalment, s'analitzava mitjançant espectrometria de masses el material metilat i no metilat, per tal de comprovar la eficiència del procés i ja es podien realitzar les proves de cristallització.

3. Derivatització dels cristalls amb àtoms pesats

L'obtenció de cristalls derivats amb àtoms pesats (de número atòmic alt) és imprescindible per a la utilització dels mètodes de fasejat experimental, tal i com es descriurà a la secció de resolució de l'estructura cristal·lina. A continuació es detallen les tècniques que permeten incorporar àtoms pesats als cristalls de les proteïnes estudiades en aquest treball.

3.1. Tractaments precristallització

Podem diferenciar els tractaments de cocrystallització, en que la proteïna és incubada amb l'àtom pesat prèviament i/o durant el procés de cristallització; i la producció de proteïna derivatitzada, en la qual s'han substituït els residus de metionina per selenometionina. Aquest últim mètode està explicat a l'apartat d'expressió i es va emprar per produir derivats de selenometionina per a Aq_DnaB sencera i Aq_DnaB [1-439].

Un cas especial de cocrystallització és quan s'empren nucleòtids halogenats, els quals incorporen un àtom de brom o iode a alguna base nitrogenada, aprofitant l'habilitat de les proteïnes estudiades per a unir una gran varietat de nucleòtids (Biswas

et al. 1986; Bujalowski *et al.* 1993). Aquesta tècnica es va utilitzar amb la proteïna Aq_DnaB[1-439].

3.2. Tractaments postcristal·lització: *soaking* dels cristalls en solucions d'àtoms pesats.

Un cop s'havien format els cristalls, aquests s'incubaren en una solució estabilitzadora durant uns 10 minuts, transcorreguts els quals es transferiren a una solució que contenia l'àtom pesat que es volia incorporar. Es pretenia aconseguir que l'àtom pesat penetrés a l'interior del cristall a través dels seus canals de solvent i que quedés integrat a la xarxa cristal·lina de forma ordenada. Tot i que es podia intentar inferir quin tipus d'àtom pesat s'incorporaria ordenadament al cristall basant-nos en les característiques fisicoquímiques de la proteïna (Garman *et al.* 2003), a la pràctica calia provar nombrosos compostos diferents fins trobar el més adequat en cada cas.

La durada i concentració del procés de *soaking* s'ha de determinar experimentalment per a cada compost i forma cristal·lina. Els protocols clàssics recomanen emprar temps de *soaking* llargs, que poden arribar a varis dies, i concentracions de solució d'àtom pesat baixes (0.1- 20 mM). A l'altre extrem, es trobaven els protocols de derivatització ràpids, els quals empen concentracions d'àtom pesat altes, properes a la saturació, i temps molt curts (pocs minuts). Un possible avantatge d'aquest darrer mètode és que es generen cristalls de derivat més isomòrfics amb els nadius (Sun *et al.* 2002). El desavantatge principal és que normalment els cristalls no sobreviuen a tractaments tant dràstics. En aquest treball es testaren ambdues estratègies.

Era important realitzar un *back-soaking* abans de congelar els cristalls. Aquest consistia en treure els cristalls de la solució de derivatització i sotmetre'ls a varies etapes curtes de rentat en una solució estabilitzadora que no contenia l'àtom pesat. D'aquesta manera s'aconseguia eliminar l'àtom pesat unit inespecíficament a la superfície del cristall o desordenat a l'interior dels canals de solvent.

4. Criocristal·lografia

4.1 El dany per radiació malmet els cristalls

Quan s'exposa un cristall a un feix de raigs X, aquests poden ser absorbits i difractats per la disposició periòdica i ordenada dels electrons dels àtoms presents al cristall. Aquest fenomen possibilita la determinació de l'estructura tridimensional de la proteïna, un cop s'han mesurat experimentalment les intensitats dels raigs difractats. De totes maneres, la interacció entre els raigs X i la matèria sòlida és un fenomen molt més complex, doncs, a part de la dispersió elàstica que causa la difracció, els fotons dels raigs X incidents poden perdre part de la seva energia al interaccionar amb la matèria i també poden transferir la seva energia i ionitzar àtoms (fotoionització).

Així doncs, els cristalls de proteïna exposats a un feix de raigs X pateixen danys per radiació, els quals provoquen que aquests perdin el seu ordre intrínsec i, conseqüentment, que les dades de difracció generades perdin progressivament tant resolució com qualitat. Aquest procés de degradació és molt ràpid si els cristalls es mantenen a temperatura ambient durant la recollida de dades, de manera que un cristall exposat a un feix intens de llum de sincrotró sovint només resisteix uns pocs segons. En canvi, quan els cristalls es mantenen a temperatures criogèniques durant la recollida, aquest procés s'alenteix molt, cosa que permet recollir més dades d'un sol cristall i que aquestes siguin de major qualitat.

Els danys per radiació que es produeixen en un cristall al qual es fa incidir un feix de raigs X són de dos tipus: primaris i secundaris. Els danys per radiació primaris es produeixen degut a la fotoionització dels àtoms del cristall. Els fotoelectrons resultants, que es solvaten, inicien processos fotoquímics que modifiquen químicament les molècules del cristall: trenquen enllaços, amb la conseqüent formació d'espècies altament reactives com els radicals lliures. Els danys per radiació secundaris són deguts a la difusió al llarg del cristall d'aquests radicals lliures. Els danys per radiació primaris depenen de la dosi de raig X absorbida pel cristall i són independents de la temperatura a que es troba la mostra, és a dir, són una part inevitable de la recollida de dades de difracció. Per contra, els danys per radiació secundaris es redueixen enormement a

temperatures criogèniques, doncs a aquestes temperatures els radicals lliures tenen taxes de difusió molt més baixes, de manera que no viatgen tant ràpid a través del cristall.

La congelació dels cristalls és molt recomanable, doncs permet reduir el dany per radiació, alhora que facilita l'emmagatzematge i el transport dels cristalls. De fet, la col·lecció de dades a temperatures criogèniques (al voltant dels 100K) s'ha convertit en un mètode vital per la cristal·lografia de proteïnes ja que redueix el dany per radiació i augmenta la vida del cristall, permeten així la recol·lecció de conjunts de dades complerts d'un sol cristall. Els desavantatges de la criocristal·lografia són que les condicions de crioprotecció han de ser establertes empíricament i que la mosaicitat del cristall sol augmentar quan aquest és congelat.

4.2 Objectius de la criocristal·lografia

El principi bàsic de la criocristal·lografia és que, durant la congelació, cal evitar la formació de gel cristal·lí a l'interior de la mostra, doncs aquest distorsiona l'ordre intern del cristall (degut a l'expansió produïda durant la seva formació) i també interfereix en el patró de difracció de la proteïna (aparició d'anells de gel). Així doncs, el cristall s'ha de congelar tant ràpid que, al finalitzar el procés, l'aigua de l'interior dels canals de solvent es trobi en l'estat vitri i no pas en el cristal·lí. Per l'aigua pura, la vitrificació de gotes micromètriques s'assoleix emprant temps de congelació de l'ordre del 10^{-5} segons. Clarament, aquesta és una velocitat de congelació inassolible pel solvent dels espais intersticials dels cristalls. Ara bé, aquest obstacle es pot superar reemplaçant part de l'aigua de la solució mare amb un agent crioprotector ("anticongelant") que augmenta a 1-2 s el temps disponible per a que el procés de congelació resulti en vitrificació i no en cristal·lització.

Hi ha un segon requisit, bàsic per a l'èxit de la congelació dels cristalls. La solució crioprotectora emprada no ha de degradar el cristall ni per atac a la seva superfície ni per xoc osmòtic. Per tant, a l'hora d'escollir quin agent crioprotector provar, primer cal considerar els components de la solució mare, doncs si un agent

crioprotector ja està present a la solució mare en concentracions baixes, normalment la concentració es pot augmentar sense efectes negatius.

4.3 Crioprotectors

Bàsicament, en cristal·lografia de proteïnes s'utilitzen dos tipus d'agents crioprotectors:

- aquells que penetren als canals de solvent del cristall. S'inclouen en aquest grup la majoria d'agents crioprotectors emprats en criocristal·lografia: glicerol, MPD, alcohols, polietilenglicols de baix pes molecular, sucres, criosalts, etc.

- aquells que embolcallen el cristall. En aquest grup hi trobem diversos olis, els quals no penetren al cristall, sinó que proporcionen una barrera entre la superfície del cristall i l'aire. Per tal d'obtenir uns resultats òptims amb aquest tipus de crioprotectors, cal eliminar la major part del solvent que envolta el cristall abans de submergir-lo en l'oli crioprotector. Exemples: Paratone N, parafina, oli mineral d'alta densitat.

Històricament el glicerol ha estat el crioprotector més utilitzat, però això no implica que sigui sempre el millor. Al tractar-se del primer crioprotector que es sol provar, normalment és només el més convenient, doncs no es duen a terme subsegüents optimitzacions si aquest ja dóna resultats adequats. En casos difícils, es poden provar combinacions de diferents crioprotectors i, fins i tot, emprar crioprotectors penetrants com a primer cas i oli com a pas final.

Les solucions crioprotectors es formulen substituint l'aigua de la solució en que ha crescut el cristall per l'agent crioprotector. Normalment, per evitar que el cristall es dissolgui durant el procés de crioprotecció, es parteix d'una solució estabilitzadora, no de la solució mare. La solució estabilitzadora emprada en aquest treball consisteix en la condició de cristal·lització multiplicant per 1.2 les concentracions dels precipitants presents.

Per tal de determinar la concentració de crioprotector que calia afegir a la solució estabilitzadora, la solució crioprotectora putativa era muntada primer sola (sense

cristall) en un *loop*, es congelava en criogen i es sotmetia a un test de difracció. S'escollia aquella solució crioprotectora que no generava anells de gel al patró de difracció. Com la concentració de crioprotector es dilueix lleugerament al afegir-hi el cristall i la solució mare que l'envolta, s'optà per una aproximació més segura, consistent en augmentar la concentració de crioprotector en un 2-5% respecte a la concentració determinada en les proves sense cristall.

4.4 Protocols de crioprotecció

El cas ideal és quan la solució mare del cristall ja contenia una concentració suficient d'agent crioprotector, de manera que no calia afegir-ne més. El cristall llavors es podia pescar directament de la solució on havia crescut i ser congelat, minimitzant així la seva manipulació. En alguns casos fou possible modificar les condicions de cristallització inicials per aconseguir que els nous cristalls creixessin ja en presència de crioprotector.

En la majoria dels casos, però, es va haver de transferir el cristall de la solució de creixement a la de crioprotecció abans de poder-lo congelar. Com a regla general, s'intentà manipular el cristalls el mínim possible, doncs aquesta manipulació podia causar danys a la superfície del cristall així com la seva deshidratació. Aquests traumes provocarien un increment de la mosaïcitat del cristall i, consegüentment, dades de pitjor qualitat i menor resolució.

Hi ha varies maneres de transferir el cristall de la gota en que ha crescut a la solució crioprotectora escollida. A continuació es descriuen les emprades en aquest treball.

4.4.1 Transferència directa a la solució crioprotectora final.

El cristall era transferit, amb un *loop* (o, per cristalls fràgils, amb un aparell de succió fet a partir d'una pipeta Pasteur i un tros de tub flexible), de la gota on havia crescut directament a una nova gota a la concentració final de crioprotector. Es deixava

allà perquè el crioprotector penetrés al cristall un temps que podia variar des d'un segon a minuts o dies, depenent de l'estabilitat del cristall dins la solució crioprotectora. El temps d'immersió del cristall en la solució crioprotectora (*soaking* en anglès) més habitual era d'entre 1 i 3 minuts.

4.4.2 Addició gradual de l'agent crioprotector.

Es tractava d'una estratègia més suau pels cristalls, consistent en canviar progressivament les condicions de la solució precipitant en que aquests es trobaven fins arribar a les de la solució crioprotectora final. Al protocol emprat en aquest treball, per tal de minimitzar les manipulacions directes a que es sotmet el cristall, es canviava la solució en que es trobava el cristall sense traslladar-lo (transferència *in situ*).

En primer lloc, el cristall s'incubava en una solució estabilitzant durant 5 o 15 minuts. A continuació, i mantenint-lo sempre en la solució estabilitzant, s'augmentava la substància que faria de crioprotector en petits percentatges fins arribar al valor desitjat. El percentatge d'augment a cada pas, així com la seva durada, es determinaven empíricament, doncs eren específics per a cada proteïna. Era recomanable vigilar sota la lupa binocular que els cristalls no es deformessin durant el procés (per això anava bé provar primer el protocol amb cristalls no útils pels experiments de difracció).

En cas de que els cristalls s'haguessin crescut mitjançant la tècnica de la gota assegurada, es treia la cinta adhesiva de manera que la gota que contenia el cristall quedés exposada. S'examinava la gota sota la lupa binocular i s'afegia, amb molta cura i al costat oposat a on hi havia el cristall, un volum igual al de la gota de la primera solució de crioprotecció. Es tornava a col·locar la cinta adhesiva sobre el pou original. Si els cristalls s'havien crescut per gota penjant, es treia el cobreobjectes que contenia la gota de sobre el pou i es col·locava en una placa de Petri petita. S'envoltava la gota de gotes més petites amb la solució crioprotectora final i s'afegia a la gota que contenia el cristall el mateix volum de la primera solució crioprotectora. En ambdós casos, s'incubava uns 5 minuts a la temperatura en que s'havien crescut els cristalls. Llavors s'obrien les gotes, es retirava un volum igual a l'afegit durant el pas anterior i s'addicionava el mateix volum de la solució 2. Es repetien els passos anteriors amb les següents

solucions de *soaking* fins arribar a la solució crioprotectora final. Després del pas final, en què s'incubava durant 5' la gota que contenia el cristall amb la solució crioprotectora, ja es podia procedir a la congelació dels cristalls.

4.4.3 Equilibrat per pressió de vapor enfront de la solució crioprotectora final.

Es tracta del mètode menys invasiu, doncs la penetració de l'agent protector al cristall és més lenta que en els mètodes d'immersió. El cobreobjectes amb la gota penjant que contenia el cristall era segellat sobre un reservori amb 1 ml de solució crioprotectora per permetre l'equilibrat dels components de vapor. El temps d'equilibrat variava d'una nit a varis dies, transcorregut el qual el cristall es sotmetia a un *soaking* ràpid en la solució crioprotectora final.

4.4.4 Incorporació de l'agent crioprotector mitjançant diàlisi.

En els casos més difícils, les solucions crioprotectors es dialitzaven lentament a l'interior del cristall. Metodològicament és el mètode més complicat, fet que comportà que només s'utilitzés com a recurs final.

4.5 Muntatge dels cristalls

Un cop crioprotegit, calia pescar el cristall del tampó crioprotector i congelar-lo, submergint-lo ràpidament en nitrogen líquid o en el flux de nitrogen gas. Tots els cristalls d'aquest treball foren congelats en nitrogen líquid, seguint el protocol que es detalla a continuació.

El cristall es pescava amb un llaç de niló (*loop* en anglès), a l'interior del qual hi quedava retingut gràcies a la tensió superficial del líquid que l'envoltava. Era convenient escollir un *loop* de mida similar a la del cristall per tal de minimitzar el volum de líquid (és a dir, reduir dispersió dels raigs X i així maximitzar la relació senyal/soroll). Per a alguns cristalls, es van emprar *loops* de reixeta, anomenats

litoloops. Aquests presentaven, com a característiques distintives, una menor retenció de líquid al voltant del cristall i major facilitat per pescar cristalls molt petits o en forma de placa fina (hi queden retinguts més fàcilment i és més fàcil centrar-los a les línies microfocalitzades de sincrotró). A la figura M.4 de materials es mostren els diferents tipus de *loops* emprats en aquest treball.

Per pescar el cristall, primer era convenient remoure lleugerament el líquid que conforma la gota per tal de provocar corrents que ajudessin a aixecar el cristall del fons de la gota. Si hi estava enganxat, es podia emprar una agulla especial (Crystal ProbeTM de Hampton Research) per alliberar-lo amb molt de compte. Un cop el cristall surava a la superfície de la gota, es portava el *loop* lateralment en direcció al cristall, seguint el pla horitzontal, de manera que el cristall quedés encerclat. Tot seguit, el *loop* es portava cap amunt, de manera que el niló que el forma trencava la tensió superficial de la gota. Al mantenir el pla del *loop* perpendicular a la superfície de la gota, es minimitzaven les forces que actuaven sobre el cristall i també es minimitzava el gruix de la capa de líquid que recobria el cristall un cop aquest estava al *loop*.

Quan el cristall es trobava dins del *loop*, aquest es submergia ràpidament en nitrogen líquid, col·locant la base que subjectava el *loop* dins d'un vial de crioprotecció, que ja es trobava submergit al criogen. Un cop congelats, els vials que contenen els cristalls es mantenen en nitrogen líquid dins de contenidors especials fins la recollida de dades de difracció, moment en el qual es col·locaren al difractor, on eren mantinguts a temperatures criogèniques durant la recollida gràcies a un flux de nitrogen gas que incidia sobre el cristall.

Els dos temps claus en aquest procés eren el temps que comportava la transferència del cristall del tampó crioprotector al criogen, i el temps que trigava el cristall a congelar-se per sota de la transició vítria al voltant de 155K. Per a uns resultats òptims, aquests dos temps havien de ser tan curts com fos possible. El primer per evitar la deshidratació de la superfície del cristall, cosa que provocaria un increment en la mosaicitat. El segon per minimitzar la mosaicitat i la probabilitat de formació de gel.

Adicionalment, calia vigilar tres factors més que afectaven la taxa de congelació i es trobaven sota control de l'experimentador: el ràtio superfície volum del cristall, la mida del *loop* i el volum de tampó crioprotector. Un cristall amb una gran taxa S/V, en un *loop* de mida equivalent a la seva i que tingués la capa de tampó crioprotector o oli de menor gruix al seu voltant, generaria un patró de difracció de raigs X d'una qualitat molt superior que un cristall amb les característiques oposades (Garman *et al.* 2006).

II. DIFRACCIÓ DE RAIGS X I RESOLUCIÓ DE L'ESTRUCTURA

1. Experiment de difracció de raigs X

L'ús d'una radiació electromagnètica per visualitzar objectes requereix que la radiació tingui una longitud d'ona comparable amb les distàncies més petites que es volen resoldre. Les longituds d'ona característiques dels raigs X comprenen valors molt similars a la distància entre àtoms de carboni units, per tant, és l'adequada per estudiar l'estructura molecular.

La difracció procedent d'una sola molècula seria massa petita per a poder ser mesurada. Per tant, s'empra un conjunt tridimensional ordenat de molècules, és a dir, un cristall, per amplificar el senyal. Un cristall ordena un nombre molt elevat de molècules de proteïna en la mateixa orientació, de manera que les ones difractades en fase poden sumar-se i elevar la senyal a un nivell mesurable. Si l'ordre intern del cristall és pobre, els raigs X no seran difractats cap a angles alts o alta resolució i, per tant, les dades no podran generar una estructura molecular detallada. Per contra, si el cristall està ben ordenat, llavors la difracció serà mesurable a angles alts o alta resolució i en resultarà una estructura detallada. Els raigs X són difractats pels electrons de l'estructura i, conseqüentment, el resultat d'un experiment de difracció de raigs X és un mapa tridimensional que mostra la distribució dels electrons a l'estructura proteica. Com els electrons estan majoritàriament localitzats al voltant dels nuclis, el mapa de densitat electrònica ens dóna una representació força bona de la molècula.

Un cristall es comporta com una reixeta de difracció, la qual genera efectes tant constructius com destructius en el patró de difracció, de manera que apareix al detector com una sèrie de punts discrets que s'anomenen reflexions. Cada reflexió conté informació de tots els àtoms de l'estructura i, per tant, cada àtom contribueix a la intensitat de cada reflexió.

Tal i com passa amb totes les radiacions electromagnètiques, els raigs X estan caracteritzats per una amplitud d'ona i una fase. Per tal de recombinar un patró de difracció, es necessiten aquests dos paràmetres per a cada reflexió. Desafortunadament, només les amplituds poden ser mesurades experimentalment, la informació de fase no. Això s'anomena el problema de la fase.

1.2 Recollida de dades de difracció

Els cristalls, muntats en un *loop*, es col·loquen sobre un capçal de goniòmetre, el qual permet que la mostra sigui posicionada acuradament en el feix de raigs X. Per a la recollida de dades a temperatures criogèniques, un feix de nitrogen gas manté el cristall a 100 K durant tot l'experiment de difracció. Els raigs X focalitzats emergeixen d'un tub estret anomenat col·limador i glopegen el cristall per produir un patró de difracció, el qual es grava al detector de raigs X.

Per a una recollida de raigs X correcta, primer cal determinar la simetria del cristall, el paràmetres de la cel·la unitat, l'orientació del cristall i el límit de resolució. Aquesta informació ens permet establir una estratègia de recollida que maximitzi tant la resolució com la completesa del conjunt de dades.

El mètode de recollida consisteix en rotar el cristall a un augment d'angle petit i fix, típicament 1° , i gravar el patró de difracció de raigs X que es genera a cada pas. Si el patró de difracció està molt ple, amb les diferents taques de difracció molt juntes, cal reduir l'angle de rotació de manera que cada punt pugui ser resolt a la imatge. Com més baixa sigui la simetria del cristall, més dades s'han de recollir. S'usen programes d'indexat, MOSFLM o XDS, per predir el grau de rotació, temps d'exposició i imatges que cal recollir per obtenir un conjunt de dades complet.

1.3 Processament de les dades de difracció

El conjunt de imatges obtingudes durant la recollida de dades difracció de raigs X s'ha de processar per tal d'obtenir un fitxer amb totes les reflexions mesurades. Cada reflexió prové d'un pla de difracció i té una intensitat pròpia que depèn de la quantitat de densitat electrònica que conté el pla. Aquest fitxer de reflexions conté les intensitats corresponents a cada pla i és necessari per a qualsevol càlcul cristal·logràfic.

1.3.1 Indexat, integració i escalat

El processament de les dades el podem dividir en tres etapes: indexat, integració i escalat. El procés d'indexat consisteix en determinar el grup espacial, l'orientació del cristall i la seva cel·la unitat a partir de les imatges de difracció. Això ens permetrà assignar a cada imatge una orientació, i a cada reflexió el seu índex de Miller hkl .

El següent pas és la integració, que consisteix en mesurar la intensitat de cada reflexió particular, tot ajustant la distribució d'intensitat a una gaussiana. Cada reflexió s'envolta d'una caixa que s'utilitzarà per calcular el soroll de fons associat a cada reflexió i mesura d'intensitat. Això es fa degut a que el soroll de fons no és homogeni en tota la superfície de la imatge i, a més, a determinades zones d'aquesta s'observen ombres o difracció difusa produïdes pel solvent desordenat del cristall. Tot el conjunt de mesures es recull en un fitxer que conté, entre d'altres, els índexs de cada reflexió, amb la seva intensitat integrada i la seva desviació estàndard. Tots aquests paràmetres s'utilitzaran en el següent pas, que és l'escalat de les dades.

L'escalat tracta de promitjar totes les reflexions repetides recollides i reconstruir les que estan disperses en varies imatges a causa de la mosaïcitat. Petites variacions en el temps d'exposició, la intensitat del feix o el dany que pateix el cristall al llarg de la recollida, fan que calgui aplicar un factor d'escala entre reflexions recollides en diferents imatges. Si estem recollint un experiment de tipus MAD, durant l'escalat s'han de tenir en compte els parells de Bijvoet, ja que en aquest cas no es complirà la llei de Friedel i les reflexions oposades no tindran la mateixa amplitud.

Les dades de tots els cristalls recollits ens aquest treball van ser indexades i integrades amb el programa XDS (Kabsch 1993) i promitjades i escalades amb el programa XSCALE (Kabsch 1993) o SCALA (Evans 1993).

2. Resolució de l'estructura

Per tal de determinar l'estructura, primer cal resoldre el problema de la fase, és a dir, cal obtenir alguna informació de fase. Per a la determinació d'estructures de proteïna, això es pot fer bàsicament de dues varies maneres: pertorbant l'estructura i el patró de difracció (reemplaçament isomòrfic i dispersió anòmala) o emprant estructures conegudes similars per inferir uns valors de fase inicials (reemplaçament molecular). També es poden usar mètodes directes per estimar les fases, però això només es possible per estructures petites.

2.1 Mètodes directes

Els mètodes directes es basen en la positivitat i l'atomicitat de la densitat electrònica. Si s'assumeix que un cristall està fet per àtoms de forma similar i que tots tenen densitats electròniques positives, llavors existeixen relacions estadístiques entre els conjunts de factors d'estructura. Aquestes relacions estadístiques poden emprar-se per deduir els possibles valors de les fases.

Els mètodes directes exploten aquestes relacions, i es poden utilitzar per resoldre estructures de molècules petites. Malauradament, aquestes relacions estadístiques esdevenen més febles a mesura que el nombre d'àtoms augmenta, i els mètodes directes estan limitats a estructures amb, com a màxim, uns pocs centenars d'àtoms a la cel·la unitat. Encara que hi ha avenços que incrementen aquests límits, particularment per a cristalls que difracten fins a molt alta resolució (1.2Å o millor), els mètodes directes no són generalment aplicables a la immensa majoria d'estructures cristal·lines. De totes

maneres, són molt útils en el context dels mètodes de fasejat experimental, com ara el reemplaçament molecular i la dispersió anòmala, doncs s'empren rutinàriament per trobar la subestructura d'àtoms pesats dins de la xarxa cristal·lina.

En aquest treball s'han utilitzant els programes de mètodes directes SHAKE & BAKE (Miller *et al.* 1994), SHELXD (Uson *et al.* 1999; Schneider *et al.* 2002) i SHARP (Abrahams *et al.* 1996) per trobar les posicions dels àtoms pesats i els programes SHELXE (Sheldrick 2008) i ACORN (Yao 2002) en la subsegüent determinació de fase a partir d'aquesta subestructura.

2.2 Reemplaçament molecular:

Si es coneixen les coordenades d'una proteïna similar es pot provar de resoldre l'estructura per reemplaçament molecular (*molecular replacement*), mètode que implica rotar i traslladar el model dins del nou sistema cristal·lí fins obtenir una bon aparellament amb les dades experimentals. Com a regla general, es necessita una identitat de seqüència superior al 25% entre els aminoàcids del model i els de l'estructura a resoldre, encara que hi ha excepcions. Si es té èxit, es poden calcular llavors les amplituds i les fases a partir d'aquesta solució, el que ens permet, amb les amplituds experimentals, calcular un mapa de densitat electrònica.

Normalment, s'utilitzen mètodes de Patterson per obtenir primer l'orientació del model a la nova cel·la unitat i després la translació del model correctament orientat relativa a l'origen de la nova cel·la unitat.

En aquest treball s'utilitzaren els programes AMoRe (Navaza 1994), Phaser (McCoy *et al.* 2007) i MolRep (Vagin *et al.* 1997) per les cerques mitjançant la tècnica del reemplaçament molecular.

2.3 Reemplaçament isomòrfic

Si no es disposa d'un model inicial, es pot utilitzar el reemplaçament isomòrfic, en el qual un o més àtoms pesants són introduïts a llocs específics de la cel·la unitat sense pertorbar la xarxa cristal·lina.

El reemplaçament molecular es basa en que les proteïnes i els àcids nucleics estan formats gairebé enterament per àtoms amb nombres atòmics (Z) inferiors a 16, i majoritàriament amb $Z < 9$. Els diferents àtoms contribueixen a la intensitat dispersada en proporció al quadrat del nombre d'electrons que contenen. Per exemple, un àtom d'urani conté 15 vegades més electrons que un àtom de carboni, per tant, la seva contribució a la intensitat serà l'equivalent a 225 àtoms de carboni. Per tant, si hi ha un petit nombre d'àtoms pesats dins l'estructura cristal·lina, el patró de difracció estarà significativament pertorbat en relació amb el patró generat per l'estructura nativa, que només conté els àtoms lleugers.

Per resoldre una estructura per reemplaçament isomòrfic cal mesurar dades de difracció generades per cristalls de proteïna sense derivatitzar (dades natives) i dades amb l'àtom pesat unit a la xarxa cristal·lina. Mesurant les diferències en les intensitats dels punts per a cada reflexió del patró de difracció és possible derivar una estima de l'angle de fase emprant mètodes de sumació de vectors. Les diferències en les intensitats dispersades reflexaran majoritàriament la contribució dels àtoms pesats, i aquestes diferències poden ser emprades per, per exemple, calcular un mapa de Patterson. Com només hi ha uns pocs àtoms pesats, aquest mapa de Patterson serà relativament senzill i fàcil d'interpretar. Alternativament, es poden aplicar mètodes directes a les diferències en les intensitats. Una vegada es coneix la localització dels àtoms pesats al cristall, es pot calcular la seva contribució als factors d'estructura. Això permet deduir els possibles valors dels angles de fase de la proteïna.

Per a poder emprar aquesta tècnica, cal que la cel·la unitat corresponent al cristall natiu, així com l'estructura i l'orientació de la proteïna, romanguin inalterada per l'addició dels àtoms pesats. Això és el que anomenem isomorfisme. Si l'àtom pesat no canvia la xarxa cristal·lina ni la resta de l'estructura, llavors el factor d'estructura

estimat pel cristall de derivat (F_{PH}) és igual a la suma del factor d'estructura de la proteïna (F_P) i el factor d'estructura de l'àtom pesat (F_H). Com els factors d'estructura poden ser representats com a vectors, aquesta equació defineix un triangle. Coneixem la llargada i l'orientació d'una de les cares (F_H), i les llargades de les altres dues cares. Hi ha dues maneres de construir aquest triangle, de manera que hi ha dues fases possibles per F_P . En principi, l'ambigüitat de fase es pot desfer preparant un segon derivat amb àtoms pesats que s'uneixin a llocs diferents. En aquest cas, la tècnica emprada rep el nom de reemplaçament isomòrfic múltiple (MIR).

2.4 Dispersió anòmala (*anomalous scattering*)

La majoria dels electrons que formen un cristall interactuaran amb els raigs X de manera idèntica. Si es situen a l'origen del cristall, difractaran amb una fase relativa de zero. A causa d'això, les parelles de punts de difracció compleixen la llei de Friedel.

El camp elèctric de l'ona electromagnètica indueix una oscil·lació als electrons. Mentre la freqüència de l'oscil·lació sigui molt diferent de la freqüència natural d'oscil·lació, els electrons oscil·laran tots amb la mateixa fase. Això és cert per a la majoria dels àtoms d'un cristall. Però si és similar a la freqüència natural d'oscil·lació, llavors es produirà un petit canvi tant en l'amplitud com en la fase de l'oscil·lació induïda. Això és cert per alguns electrons de les capes més internes en alguns àtoms, on l'energia del fotó de raigs X està propera a l'energia de transició. El canvi en amplitud i fase s'anomena difusió anòmala (*anomalous scattering*).

El canvi de fase en la dispersió anòmala provoca el trencament de la llei de Friedel. El canvi de fase sempre és a $+90^\circ$ de la contribució de la dispersió anòmala, cosa que provoca el trencament de la simetria entre les amplituds dels parells de Friedel (F_{hkl} i F_{-h-k-l}). L'efecte del *scattering* anòmal depèn de que la freqüència de l'oscil·lació sigui similar a la freqüència natural de l'àtom. Per tant, clarament la magnitud de l'efecte de la dispersió anòmala depèn de la longitud d'ona dels raigs X.

2.4.1 Dispersió anòmala de múltiple longitud d'ona (MAD)

Recollint dades de difracció a varies longituds d'ona prop del cantell d'absorció d'un element del cristall, es pot obtenir informació de fase anàloga a l'obtinguda per reemplaçament isomòrfic. El cantell d'absorció és l'energia a la qual l'absorció (f'') s'incrementa dramàticament en funció de l'energia. Aquesta tècnica s'anomena dispersió anòmala de múltiple longitud d'ona (MAD) i és un mètode elegant i sovint molt efectiu que confia enterament en la mesura de les diferències anòmales produïdes per un o més àtoms de dispersió anòmala en el cristall. Les dades es recullen d'un mateix cristall normalment a tres longituds d'ona diferents per tal de maximitzar els efectes d'absorció i dispersió. Típicament, les longituds d'ona s'escullen al pic (λ_1) d'absorció, f'' , al punt d'inflexió de la corba d'absorció (λ_2), on el terme dispersiu (el qual és la derivada de la corba f'') té el seu mínim, i a una longitud d'ona remota (λ_3). Com aquest mètode requereix un font de raigs X sintonitzable, només es pot dur a terme en un sincrotró.

Els canvis en les amplituds dels factors d'estructura generats per la dispersió anòmala són generalment petits i necessiten mesures molt acurades de les intensitats. La forma real de la corba d'absorció s'ha de determinar experimentalment amb un *scan* de fluorescència del cristall al sincrotró, ja que l'entorn dels dispersors anòmals pot afectar els detalls de l'absorció. També es necessita una òptica excel·lent per donar una longitud d'ona acurada amb la mínima dispersió de longitud d'ona. Normalment, totes les dades es recullen d'un sol cristall congelat amb alta redundància per augmentar la significança estadística de les mesures i les dades es recullen amb la màxima completesa possible.

El reemplaçament isomòrfic té varis problemes: el no isomorfisme entre cristalls (canvis en la cel·la unitat, reorientacions de la proteïna, canvis conformacionals, canvis en la salt i els ions de solvent), problemes per localitzar els àtoms pesats, problemes per refinar la posició dels àtoms pesats, la seva ocupància i els paràmetres tèrmics i els errors en les mesures d'intensitats. L'ús del mètode de MAD resol els problemes de no isomorfisme, doncs totes les dades es recullen d'un mateix cristall.

2.4.2 Dispersió anòmala a una sola longitud d'ona: SAD

També es possible recollir dades només a una sola longitud d'ona, normalment el pic d'absorció, i usar els protocols de modificació de densitat per trencar l'ambigüitat de fase i donar mapes interpretables.

2.4.3 Combinar reemplaçament isomòrfic i dispersió anòmala

La dispersió anòmala pot utilitzar-se per trencar l'ambigüitat de fase en un experiment de reemplaçament isomòrfic senzill, donant lloc a un SIRAS (reemplaçament isomòrfic senzill amb dispersió anòmala). Degut a l'avançament de 90° en la fase del terme f' , la dispersió anòmala dóna informació de fase ortogonal al terme isomòrfic. L'ús de reemplaçament isomòrfic múltiple combinant amb la dispersió anòmala s'anomena MIRAS (reemplaçament isomòrfic múltiple amb dispersió anòmala).

3. Càlcul del mapa de densitat electrònica

La resolució del problema de les fases permet visualitzar el mapa de densitat electrònica de l'estructura que s'està determinant. El mapa de densitat electrònica és la transformada de Fourier dels factors d'estructura, amb les seves corresponents fases. La transformada de Fourier (que es mostra a la figura M.9) es calcula com un conjunt de valors de densitat electrònica a cada punt de la malla tridimensional.

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

Figura M.9: Funció de la densitat electrònica. Està definida a cada punt de la cel·la unitat, de coordenades x, y, z . Representa la transformada de Fourier entre l'espai real, definit per $\rho(xyz)$, i l'espai recíproc, definit pels factors d'estructura $F(hkl)$.

La densitat electrònica en un punt ve determinada per la informació que contenen tots els factors d'estructura i a la inversa, cada valor d'un factor d'estructura (mòdul i fase) ve determinat per les posicions de tots els àtoms que formen l'estructura. Igualment, la qualitat d'un mapa de densitat electrònica depèn de l'exactitud de les dades de difracció, la qualitat de les fases determinades i la resolució, de manera que al augmentar aquests paràmetres, major és el detall al qual s'observa el mapa.

4. Millora de les fases

És rar que les fases determinades experimentalment siguin suficientment acurades per donar un mapa de densitat electrònica completament interpretable. Les fases experimentals normalment són només el punt de partida per la millora de les fases emprant una gran varietat de mètodes de modificació de densitat, els quals es basen majoritàriament en la idea de que es coneixen algunes de les característiques que un bon mapa de densitat electrònica hauria de tenir i que, per tant, si es canvia el mapa per a que s'assembli més a l'ideal, les fases calculades a partir d'aquest mapa seran més acurades que les fases originals. L'aplanament del solvent (*solvent flattening*), la concordança en histograma (*histogram matching*) i el mitjanat no-cristal·logràfic (NCS *averaging*) són les principals tècniques usades per modificar la densitat electrònica i millorar les fases. Aquests mètodes es troben a dins de programes com DMMULTI (Cowtan 1994), RESOLVE (Terwilliger 2000) i CNS (Abrahams *et al.* 1996).

La modificació de densitat normalment és un procediment cíclic que comporta tornar a transformar el mapa de densitat electrònica modificat per generar fases modificades, recombinar aquestes fases amb fases experimentals (per no llençar la realitat experimental) i el càlcul d'un nou mapa que després es torna a modificar, de manera que el cicle continua fins que s'assoleix convergència.

Aquests mètodes també es poden usar per proporcionar fases més enllà de la resolució per a la qual hi havia informació de fases experimentals, assumint que es disposa de dades natives de major resolució. En aquests casos, el mapa modificat es

torna a transformar a una resolució lleugerament major a cada cicle per proporcionar noves fases per les reflexions de més alta resolució.

4.1 Aplanament de solvent

En un cristall típic de proteïna, al voltant de la meitat del volum està ocupat per molècules de proteïna ben ordenades, mentre que l'altre meitat està ocupada per solvent desordenat. El solvent desordenat hauria de tenir una densitat electrònica plana, sense cap forma, per tant, si hi ha formes a la zona del solvent segurament són degudes a errors en la fase. Si la densitat del mapa es modifica de manera que la regió del solvent s'aplana, les fases corresponents seran més acurades. El *solvent flattening* és una tècnica molt poderosa que treu la densitat electrònica negativa i posa el valor de densitat electrònica a les regions de solvent a un valor típic de $0.33\text{e}\text{\AA}^{-3}$, que contrasta amb una densitat electrònica típica per proteïna de $0.43\text{e}\text{\AA}^{-3}$. S'empren mètodes automàtics per definir la frontera entre proteïna i solvent, primer desenvolupats per Wang (1985) i després estesos a l'espai recíproc per Leslie (1988).

4.2 Mitjanat

Freqüentment les proteïnes cristal·litzen amb més d'una còpia a la unitat asimètrica de la cel·la unitat del cristall. En altres casos, les proteïnes cristal·litzen en diferents formes cristal·lines. S'espera que quan la mateixa proteïna apareix a diferents llocs en un mapa de densitat electrònica (o a mapes de diferents cristalls), la densitat sigui més o menys la mateixa a cada còpia. Tal i com passava per al *solvent flattening*, si hi ha diferències possiblement són degudes a errors a les fases. Aquest mètode realitza un mitjanat a la densitat, de manera que es cancel·len alguns dels errors aleatoris i, per tant, s'incrementa la fidelitat de les fases corresponents.

4.3 Histogram matching

L'*histogram matching* altera els valors de punts de densitat electrònica per a que concordin amb la distribució esperada de valors de densitat electrònica.

5. Construcció del model atòmic

És el procés durant el qual el mapa de densitat electrònica és interpretat en termes d'un conjunt de coordenades atòmiques. És més senzill en el reemplaçament molecular perquè ja es té un conjunt de coordenades atòmiques per treballar. En el cas del reemplaçament isomòrfic, només es té un mapa i cal construir un model.

Per construir aquest model atòmic, normalment, es col·loca un esquelet proteic a l'interior del mapa de densitat electrònica i, a mesura que la resolució ho permet, es va insertant la seqüència aminoacídica. La quantitat de detall visible depèn de la resolució i de la qualitat de les fases.

Quan es té un model preliminar, es refina contra les dades experimentals. Això millora les fases, cosa que genera mapes més clars i, per tant, models millors. Normalment, es segueix aquest cicle varies vegades fins que no es poden obtenir més millores o molt petites.

5.1 Fitejat automàtic i refinat

Es pot considerar una altra forma de modificació de la densitat. Les estructures de proteïna estan fetes d'àtoms. Si la densitat pot ser interpretada en termes d'un model atòmic (i els àtoms es col·loquen més o menys al lloc correcte), la distribució de densitat estarà més propera a la realitat i les fases corresponents seran de nou més acurades. Com es coneixen molts paràmetres que determinen com els àtoms estan col·locats uns respecte als altres en una proteïna (distàncies d'enllaç, angles d'enllaç, la connectivitat química definida per la seqüència aminoacídica), es pot explotar molt aquesta informació alhora de construir un model atòmic. A partir dels primers mapes de densitat electrònica interpretables, s'aplica iterativament la construcció del model dins d'un mapa de densitat electrònica seguit d'un refinat per assolir una major concordança del model amb les dades de difracció observades. Llavors les noves fases millorades s'empren per calcular un mapa nou i millor. Opcionalment, altres tècniques de

modificació de densitat (com ara el mitjanat i l'aplanament de solvent) es poden aplicar abans que a cada nou cicle es comenci amb la construcció d'un nou model.

5.1.1 Paràmetres estadístics

Durant el procés d'afinat, o al acabar-lo, es poden analitzar varis paràmetres estadístics per comprovar la qualitat del model atòmic. Dos d'aquests paràmetres són els valors del factor R i el factor de R_{lliure} , els quals proporcionen una bona mesura de la qualitat d'encaix entre el model i les dades experimentals. El factor R es calcula segons la següent equació $R_{\text{factor}} = \frac{\sum_{hkl} |F_{\text{obs}}| - k |F_{\text{calc}}|}{\sum_{hkl} |F_{\text{obs}}|}$, on F_{obs} és l'amplitud observada del factor d'estructura, F_{calc} l'amplitud calculada (a partir del model actual) del factor d'estructura i k és un factor d'escala. Així doncs, es tracta d'un índex de la desviació entre les amplituds observades (F_{obs}) i les calculades (F_{calc}). El factor R_{lliure} es calcula d'igual manera, però només emprant un petit percentatge de les dades, el qual es manté a part i no s'usa mai pel afinat i, conseqüentment, es considera una estima no esbiaxada de la concordança entre el model i les dades. El conjunt de reflexions que es mantindran a part i seran emprades per calcular el factor R_{lliure} es seleccionem aleatòriament entre un 5 i un 10% de les reflexions totals del conjunt de dades. El factor R_{lliure} normalment té un valor lleugerament superior al R_{factor} .

Els factors R s'utilitzen com a indicadors de qualitat durant tot el procés d'afinat i, especialment, al final d'aquest. Un augment en el factor R indica que les modificacions introduïdes al model atòmic durant el cicle d'afinat ha empitjorat la correlació d'aquest model amb les dades experimentals. El valor del factor R_{lliure} serveix de control intern.

RESULTATS

Per a una major claredat, els resultats es divideixen en quatre apartats, cadascun dels quals correspon íntegrament a una de les proteïnes estudiades en aquest treball.

La primera part correspon a la DnaB d'*Aquifex aeolicus* sencera, la segona a la DnaB de *Thermotoga maritima* sencera, la tercera al domini C-terminal d'Aq_DnaB i la quarta a la forma curta d'Aq_DnaB, Aq_DnaB[1-439].

Cada part inclou el clonatge, expressió, purificació i cristal·lització de la proteïna, així com, en el cas d'Aq_DnaB[1-439], la recollida de dades de difracció i resolució de l'estructura tridimensional.

1. DnaB d'*Aquifex aeolicus* sencera: Aq_DnaB

1.1 Clonatge

El gen complet codificant per l'helicasa replicativa DnaB es va clonar a partir de DNA genòmic d'*A. aeolicus*. El gen sencer té 1407 bases.

Mitjançant una reacció de PCR emprant els encebador Fw-Aq i Rv-Aq (descrits a la taula M.2 de materials i mètodes) es va amplificar el gen *aq_dnab* i es van incorporar les dianes de tall dels enzims de restricció NcoI i XhoI flanquejant-lo. Per a poder incorporar la diana de NcoI a l'extrem 5' del gen, es va generar una mutació a alanina del segon residu de la proteïna (Gln2). La incorporació de la diana de XhoI a l'extrem 3' no suposà cap canvi en la seqüència nucleotídica d'*aq_dnab* perquè l'encebador Rv-Aq era complementari a un regió gènica riu avall del codó de parada que marca el final del gen.

La reacció de PCR, tot i generar una banda única de la mida adequada, no tenia un rendiment gaire elevat. Això presentava problemes a l'hora de clonar aquest gen al vector d'expressió mitjançant reaccions successives de digestió i lligació, doncs la quantitat d'insert després de l'etapa de digestió era massa petita per a que la reacció de lligació funcionés correctament. Per superar aquest entrebanc, es decidí clonar el producte de PCR primer al vector de clonatge pGFPCR, que no necessita un pas previ de digestió de l'insert, doncs es tracta d'una lligació d'extrems roms, i alhora permetia una nova amplificació del gen d'interès (per transformació a una soca d'*E. coli* es poden realitzar fàcilment preparacions plasmídiques). En un segon pas, s'extreia l'insert d'aquest vector per digestió doble amb els enzims de restricció pels que s'havien incorporat dianes durant la PCR i, finalment, es lligava al vector d'expressió pET28a.

Es va comprovar mitjançant seqüenciació que el gen *aq_dnab* es trobava correctament inserit al vector d'expressió i que no presentava mutacions no desitjades generades durant l'etapa d'amplificació.

1.2 Expressió

En primer lloc es comprovà que s'obtenia sobreexpressió de la proteïna Aq_DnaB mitjançant proves d'expressió en minicultius de 4ml de medi LB en condicions astringents i emprant la soca d'expressió més comú, *E. coli* BL21(DE3). A la figura R.1 es mostra el gel SDS-PAGE, on es pot veure una banda de sobreexpressió de pes molecular aparent aproximat de més de 50 kDa, que es correspon amb l'esperat per a aquesta proteïna (53,6 kDa).

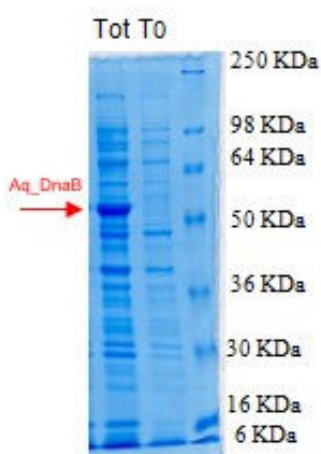


Figura R.1: Anàlisi de la sobreexpressió de la proteïna Aq_DnaB per SDS-PAGE. Les mostres corresponen a cultius d'*E. coli* BL21 (DE3) transformats amb el vector pET28[Aq_DnaB] de 4 ml de medi LB induïts a un valor d'absorbància a 600 nm de 0.5 amb 2 mM IPTG. T0: abans d'induir. Tot: mostra a les 4 hores d'haver induït el cultiu. Es marca amb una fletxa vermella la banda assignada a Aq_DnaB.

Tot seguit, es realitzaren proves pilot d'expressió per determinar la quantitat de proteïna que s'expressava de forma soluble a diferents temperatures, concentracions d'IPTG i temps d'inducció. Es determinà que la millor temperatura era 17°C i concentracions d'inductor molt baixes (entre 0.1 i 0.25 mM d'IPTG). Igualment es determinà que la durada òptima de la inducció era de 2 dies, doncs la proteïna no es degradava i, al allargar l'etapa d'inducció, s'aconseguia major quantitat final de proteïna.

També es realitzà una prova amb dues soques més, *E. coli* BL21(DE3) pLysS i *E. coli* Rosetta (DE3), per determinar si en aquestes s'incrementava l'expressió i/o la solubilitat de la proteïna. L'expressió era equivalent a totes tres soques i la solubilitat amb *E. coli* BL21(DE3) pLysS era menor i amb Rosetta similar. Es decidí treballar amb la soca normal perquè era més fàcil i ràpid créixer els cultius, doncs calien menys antibiòtics per mantenir els vectors de forma estable als cultius bacterians.

1.3 Purificació

Per al tampó de resuspensió cel·lular, ens vam basar en la literatura existent per a formular un tampó inicial. Estava basat en fosfat perquè aquest era el tampó més adequat per a l'etapa de xoc tèrmic. El pH era de 7.6 perquè es volia afavorir al màxim l'homogeneïtat de la mostra i estudis de microscòpia electrònica havien mostrat que a valors superiors a 8 o inferiors a 7 la viabilitat dels hexàmers de DnaB disminuïa considerablement (Donate *et al.* 2000).

Posteriorment, en l'etapa d'optimització es va fer un cribatge de condicions de diferents pH, concentracions de sal i presència d'agents solubilitzants. Es va decidir augmentar lleugerament la concentració de sal al tampó d'extracció i deixar tots els altres paràmetres iguals, doncs només en condicions desnaturalitzants (presència d'urea a 4-5 M) s'aconseguia recuperar més proteïna a la fracció soluble i no es va considerar oportú realitzar protocols de *refolding* donada la mida i l'estructura quaternària hexamèrica de DnaB.

Tal i com es descriu a mètodes, la mostra proteica es va extreure per sonicació i clarificar per centrifugació. S'eliminà el DNA genòmic per digestió amb DNasa i posterior precipitació amb el polímer PEI i es sotmeté la mostra a xoc tèrmic, etapa en que s'aprofita el fet de treballar amb una proteïna de termòfil expressada en un organisme mesòfil. Es varen provar diferents temperatures i temps, fins a trobar aquells que permetien recuperar la major quantitat de proteïna Aq_DnaB i el més pura possible. Tal i com es veu a la figura R.2a, aquesta etapa aconseguix eliminar moltes proteïnes contaminants i enriquir la mostra en la proteïna d'interès. Una optimització posterior del xoc tèrmic va determinar que era millor afegir MgCl₂, a una concentració final de 10 mM, abans de l'etapa de xoc tèrmic, doncs així s'aconseguia eliminar més proteïna contaminant de pes molecular aparent al voltant dels 40 kDa.

La següent etapa de purificació era una columna de bescanvi aniònic, equilibrada amb el tampó A1 (descriu a mètodes). Aq_DnaB s'el·luïa a un percentatge de tampó B1 aproximat de 57%, en un pic ample que, quan s'analitzà per SDS_PAGE es veié contenia majoritàriament la proteïna d'interès. Si aquest mateix gel es tenia amb

bromur d'etidi, es veia que la mostra es trobava contaminada amb àcids nucleics. A la figura R.2 es mostra el cromatograma (2b) i l'anàlisi de les fraccions eluïdes (2c) a l'etapa de bescanvi aniónic.

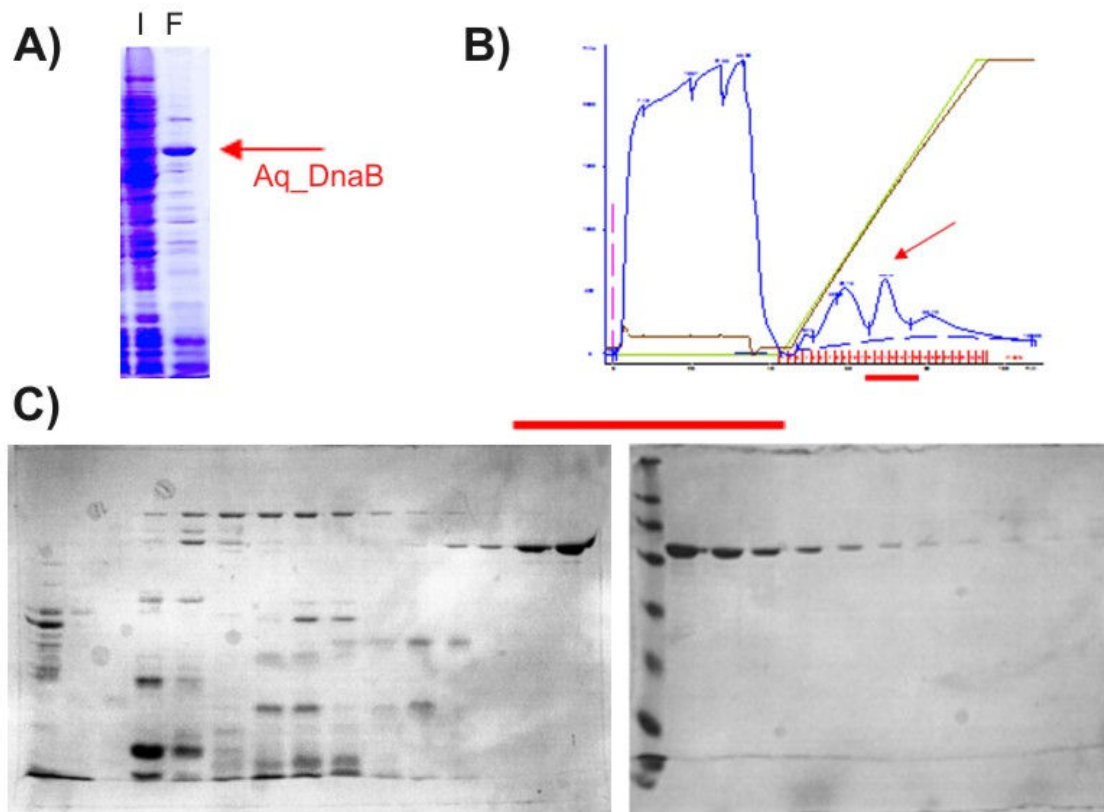


Figura R.2: Etapes de xoc tèrmic i de bescanvi aniónic de la purificació d'Aq_DnaB. **A)** Anàlisi del tractament de xoc tèrmic a 80° C durant 30'. I: extracte proteic inicial abans de sotmetre'l a xoc tèrmic. F: fracció soluble després del xoc tèrmic, aquesta mostra és la que es carrega a la columna de bescanvi aniónic. La fletxa taronja marca la banda corresponent a Aq_DnaB. **B)** Perfil cromatogràfic de l'etapa de bescanvi aniónic. Es marca amb una fletxa taronja el pic on s'elueix Aq_DnaB i les fraccions corresponents al pic estan subratllades. **C)** Anàlisi per SDS-PAGE de les fraccions eluïdes de (B). Les fraccions del pic on s'elueix Aq_DnaB estan marcades amb una franja taronja.

Es reuniren les fraccions més riques en Aq_DnaB i aquesta mostra es carregà a una columna de gel filtració Superose 6 10/300 GL. Primerament, aquesta columna estava equilibrada a una concentració de sal de 100mM NaCl i es feia al flux estàndard de 0.5 ml/min. Sota aquestes condicions, s'obtenia un pic molt ample que, si s'analitzava en detall, es veia que estava compost per un primer pic on s'eluía majoritàriament Aq_DnaB que es trobava solapat amb un segon pic molt predominant a

260 nm, on s'elueix majoritàriament DNA. Per tant, sota aquestes condicions experimentals, aquesta etapa cromatogràfica no aconseguia alliberar la mostra del DNA i, conseqüentment, no generava una mostra d'Aq_DnaB prou pura per a la següent etapa de cristal·lització.

L'etapa de cromatografia d'exclusió molecular es va optimitzar augmentant la concentració de sal del tampó a 200mM i baixant el flux a que es realitzava la columna a la meitat. Sota aquestes noves condicions, s'aconseguí un pic aïllat per Aq_DnaB i que el DNA contaminant s'elueix en un segon pic més retardat. Es comprovà que la mostra d'Aq_DnaB purificada es trobava lliure d'àcids nucleics per tinció amb bromur d'etidi. A la figura R.3 es mostra el cromatograma i l'anàlisi de les fraccions d'aquesta etapa.

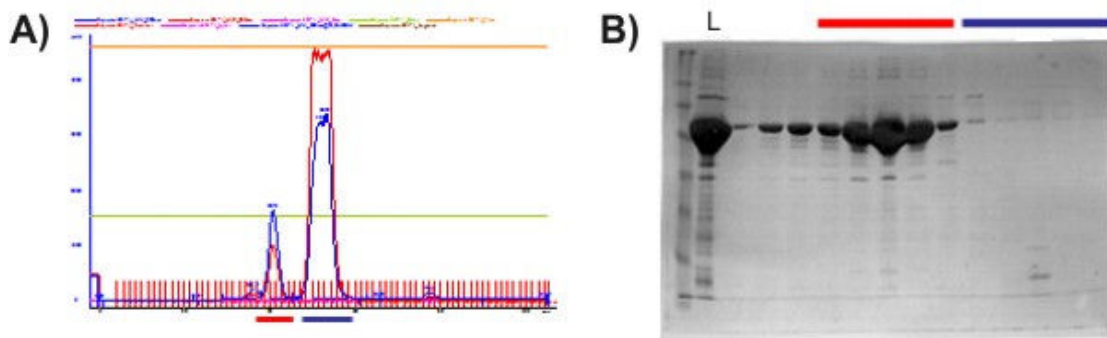


Figura R.3: Etapa de cromatografia d'exclusió molecular de la purificació d'Aq_DnaB. **A)** Perfil cromatogràfic de les fraccions eluïdes de la Superose 6 10/300 GL. La línia blava és l'absorbància a 280 nm i la vermella a 260 nm. El primer pic, subratllat en taronja, correspon a Aq_DnaB. El segon pic, subratllat en blau, és majoritàriament DNA. **B)** Anàlisi per SDS-PAGE de les fraccions eluïdes de (a). L: mostra carregada inicialment a la columna. Els carrils corresponents a les fraccions del pic on s'elueix Aq_DnaB estan marcades amb una franja taronja. Les del pic de DNA amb una franja blava.

Finalment, com a pas previ a l'etapa de cristal·lització, la mostra d'Aq_DnaB es va transferir a un tampó mínim (20mM Tris pH 7.6 100 mM NaCl), es va concentrar a 9 mg/ml i es va filtrar per eliminar impureses particulades.

La identitat de la mostra es va comprovar mitjançant espectrometria de masses, que indicava un sol pic de la massa esperada per a la proteïna.

El rendiment final de les etapes d'expressió i purificació per a la proteïna Aq_DnaB és d'aproximadament 2 mg per a litre de cultiu bacterià inicial.

1.3 Cristal·lització

El cribatge de les solucions de cristal·lització per a Aq-DnaB es va realitzar íntegrament en microgotetes i manualment. Això comportà la utilització d'una gran quantitat de proteïna i també que s'empressin diferents rondes de purificació, amb concentracions lleugerament diferents que comprenien valors entre 4 i 9 mg/ml.

Es van obtenir cristalls petits directament de la condició 12 del *screen* Clear Strategy 1, el qual té la següent formulació 25% (p/v) PEG 2000 MME 0.2 mM KBr. Posteriors optimitzacions d'aquesta condició van determinar que les millors concentracions de precipitant eren entre 18 i 20 % (p/v) PEG 2000 MME, la millor sal acetat de magnesi a 0,25 M i era important tamponar la solució de cristal·lització entre valors de pH de 7.6 i 7.9. Totes les optimitzacions es varen realitzar en plaques de gota suspesa, per superar la tendència d'aquests cristalls a créixer enganxats al plàstic de la base del pou, d'on era molt complicat recuperar-los. A la figura R.4 es mostren fotografies dels cristalls finals.

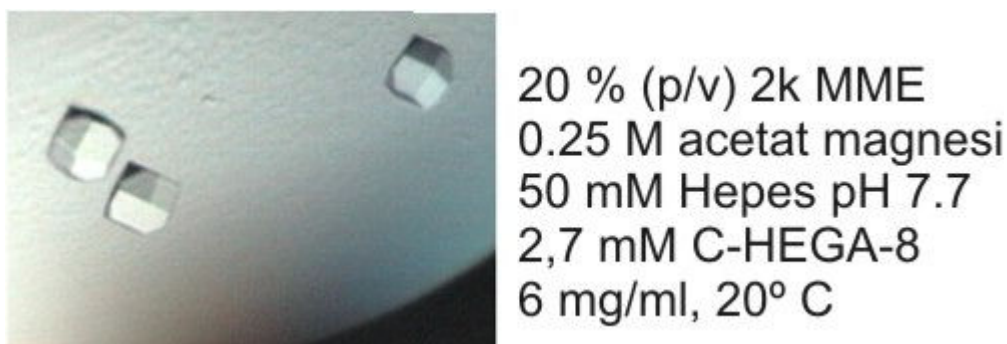


Figura R.4: Cristalls d'Aq_DnaB. S'indiquen les condicions de cristal·lització. La tècnica utilitzada fou la difusió de vapor per gota penjant. Els cristalls que es mostren a la imatge s'obtingueren en presència del detergent C-HEGA-8, que afavoria la formació de menys cristalls, més grans.

Pel que fa al pH, es important destacar que la condició inicial no estava tamponada, així doncs, hi afectava el tampó que portava la mostra proteica. Quan es varen repassar els resultats de les restants proves de cristal·lització, efectivament aquelles condicions amb pH superiors a 8 donaven precipitat abundant.

Tot i el seu bon aspecte visual, aquests cristalls eren extremadament fràgils i la seva manipulació per a crioprotegir-los i muntar-los en *loops* era molt difícil, doncs sovint es desintegren al menor contacte amb el *loop* o, fins i tot, al provocar corrents a la gota on havien crescut. Per tal que els cristalls suportessin millor l'estrès físic durant la seva crioprotecció i muntatge, eren rutinàriament sotmesos a un procés d'entrecruament amb glutaraldehid mitjançant difusió de vapor.

Durant les optimitzacions s'aconseguien cristalls molt grans i tridimensionals, però, malauradament, la qualitat de les dades de difracció generades a partir d'ells no va aconseguir millorar-se. Tots els cristalls d'Aq_DnaB provats difractaven com a màxim a 7 Å, però normalment a una resolució menor que rondava els 10 Å. La recollida d'imatges de difracció a temperatura ambient no mostrà major resolució, indicant que no es tractava d'un problema de congelat sinó que era degut al desordre intrínsec d'aquests cristalls.

S'abandonà aquesta proteïna després de provar més de 250 cristalls diferents, crescuts en condicions diferents, optimitzats amb múltiples tècniques de millora i crioprotegits amb diverses combinacions de crioprotectors i tècniques.

També es va produir i cristal·litzar proteïna derivatitzada amb seleno-metionina, amb la intenció de testar si la qualitat de les dades de difracció millorava. No va ser així.

2. DnaB de *Thermotoga maritima*: Th_DnaB

2.1 Clonatge

El gen codificant per a la proteïna DnaB sencera, de 1355 pb de llargària, fou amplificat mitjançant PCR del genoma de *Thermotoga maritima* gràcies als encebadors Fw-Th i Rv-Th (descrits a materials i mètodes). La introducció de les dianes de tall per als enzims de restricció NdeI i XmaIII flanquejant el gen *th_dnab* no suposà la generació de cap mutació a la seqüència gènica original.

El fragment amplificat fou clonat directament als vectors d'expressió pET27b i pET29a, en ambdós casos entre les dianes de NdeI i XmaIII.

La seqüenciació dels clons positius obtinguts per ambdós vectors d'expressió mostrà que s'havia aconseguit clonar la proteïna sencera i que aquesta no presentava cap mutació.

2.2 Expressió

Les primeres proves de sobreexpressió de la proteïna Th_DnaB a petita escala permeteren observar que el grau de sobreexpressió era molt elevat, major que per a Aq_DnaB. Malauradament, la major part de la proteïna s'expressava de forma insoluble, tal i com es mostra a la figura R.5.

Els nivells de sobreexpressió de Th_DnaB, així com el percentatge de proteïna que es recuperava de forma soluble, eren equivalents per als dos vectors d'expressió en que s'havia clonat la seqüència codificant per aquesta proteïna. Es va decidir treballar amb el constructe pET27[Th_DnaB], per al qual s'havien obtingut més clons positius.

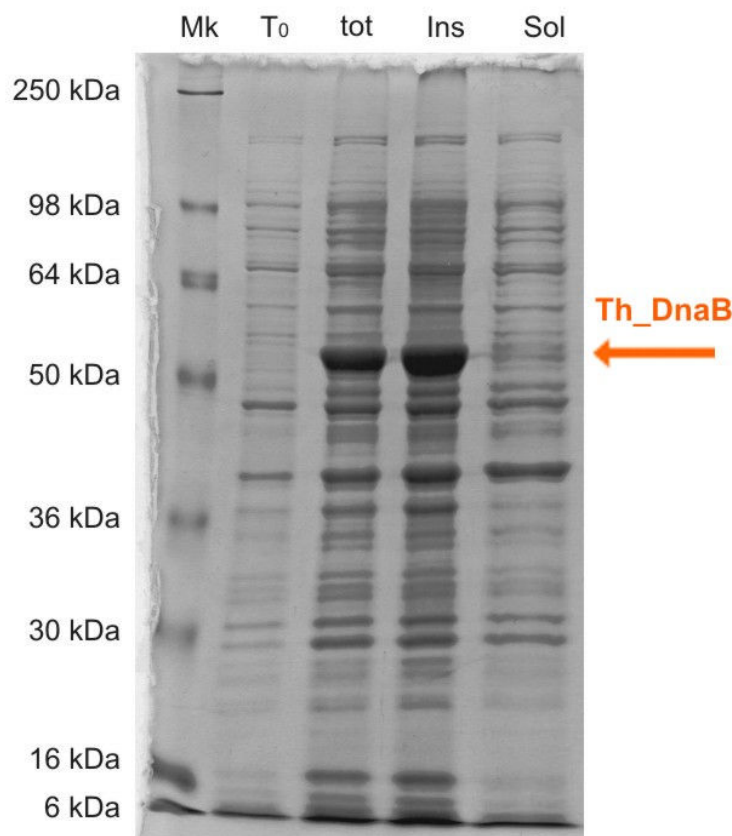


Figura R.5: Proves d'expressió a petita escala de Th_DnaB. Proves d'expressió realitzades en cultius de 4 ml de medi LB, a 37°C i 2mM IPTG. Es mostra l'anàlisi per SDS-PAGE, on els carrils corresponen a: **Mk**: marcador de pes molecular; **T₀**: mostra abans d'induir; **Tot**: fracció total a les 4 hores d'inducció; **Sol**: fracció expressada de manera soluble; **Ins**: fracció insoluble. La fletxa taronja marca la banda corresponent a Th_DnaB de pes molecular aparent coincident amb l'esperat (52.4 kDa).

Per tal d'intentar optimitzar l'expressió soluble d'aquesta proteïna, es van provar múltiples combinacions de soques d'expressió, medis de cultiu, temperatures i temps d'inducció i concentracions d'inductor. La taula R.1 detalla les variables testades.

| | |
|------------------------------|---|
| Soques | <i>E.coli</i> BL21 (DE3), BL21 (DE3) pLys S, Origami (DE3), BL21 Star (DE3), Codon Plus, Rosetta (DE·3), BL21 pT-GroE, Tuner (DE3). |
| Medis | LB, TB, Superior Broth™, medi d'autoinducció Overnight Express™ |
| T^a i temps | 37, 30, 25, 20, 17°C. De 2h a dos dies (monitoritzant cada 2-6h) |
| [IPTG] | 0.1, 0.2, 0.5, 1 mM |

Taula R.1: Variables de sobreexpressió provades per a Th_DnaB. Les característiques de les soques d'expressió, així com la composició dels medis de cultius estan explicades a materials i mètodes. LB: medi Lúria Bernani, TB: medi Terrific Broth.

A la figura es R.6 mostra el resultat de l'optimització final, comparant l'expressió soluble a *E. coli* BL21 (DE3) i a *E. coli* BL21 Tuner (DE3).

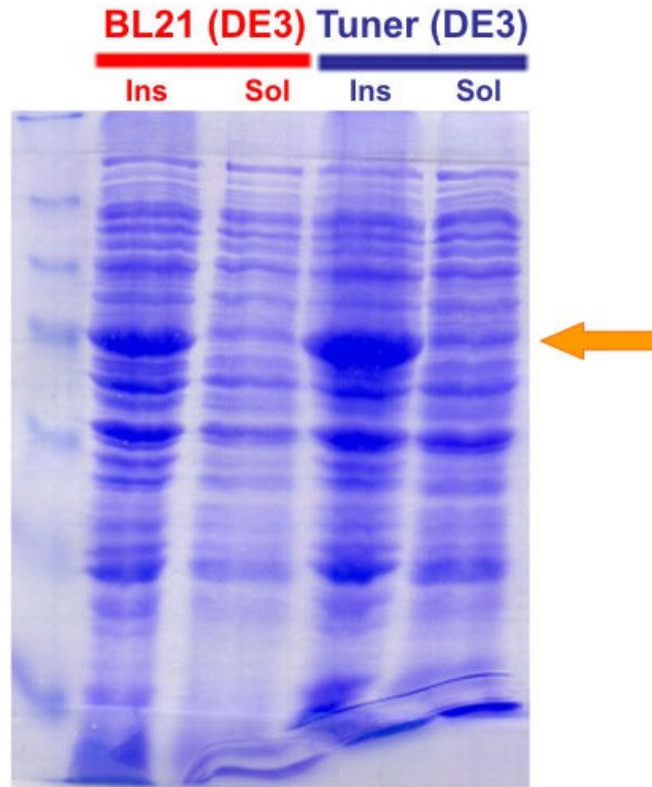


Figura R.6: Optimització de l'expressió de Th_DnaB. Es comparen els nivells de solubilitat obtinguts emprant: **Vermell)** la soca normal d' *E. coli* BL21(DE3), medi LB, 25°C, 0.1 mM IPTG i 20 hores d'inducció. **Blau)** la soca *E. coli* BL21 Tuner (DE3), medi Superior Broth™, 25° C i 0.1 mM durant 20 hores. Per a cada soca es mostren els carrils corresponents a les fraccions insolubles (Ins) i solubles (Sol). La fletxa taronja indica la banda corresponent a Th_DnaB.

Així doncs, les condicions d'expressió optimitzades implicaven emprar la soca *E. coli* Tuner (DE3), medi de cultiu Superior Broth™, IPTG a una concentració final de 0.1 mM i realitzar la sobreexpressió a 25° C durant 20 hores. Tot i que amb aquest protocol s'aconseguí augmentar considerablement el percentatge de proteïna Th_DnaB recuperada de forma soluble, la major part encara continuava expressant-se de manera insoluble en cossos d'inclusió.

2.3 Purificació

Per a l'etapa d'extracció i clarificació es provaren diferents tampons de resuspensió cel·lular, seguint una aproximació semblant a l'emprada per a Aq_DnaB sencera. Es determinà que el millor tampó de resuspensió cel·lular era 50 mM Tris pH 7.6 10% (p/v) sacarosa, 300 mM NaCl, 5 mM DTT, 1 mM EDTA (tampó R2).

Quan es realitzava el tractament de xoc tèrmic després d'aquesta etapa, la major part de la proteïna coprecipitava juntament amb la resta de proteïnes expressades per l'hoste, fins i tot quan s'empraven tractaments suaus. Això podria ser degut a que *T. maritima* no és un termòfil extrem com és *A. aeolicus*.

Es decidí carregar directament l'extracte cel·lular clarificat, prèvia precipitació del DNA alliberat durant la sonicació, a una columna de bescanvi aniònic, equilibrada amb el tampó B1. Th_DnaB eluïa a un percentatge de tampó B2 del 62 %, formant un pic aïllat. L'anàlisi de les fraccions corresponents a aquest pic mostrà que coeluien dues contaminants majoritàries, una d'alt pes molecular (> 98 KDa) i una altra de baix pes molecular (propera als 40 kDa). La figura R.7 mostra aquesta etapa cromatogràfica i l'anàlisi de les fraccions resultants.

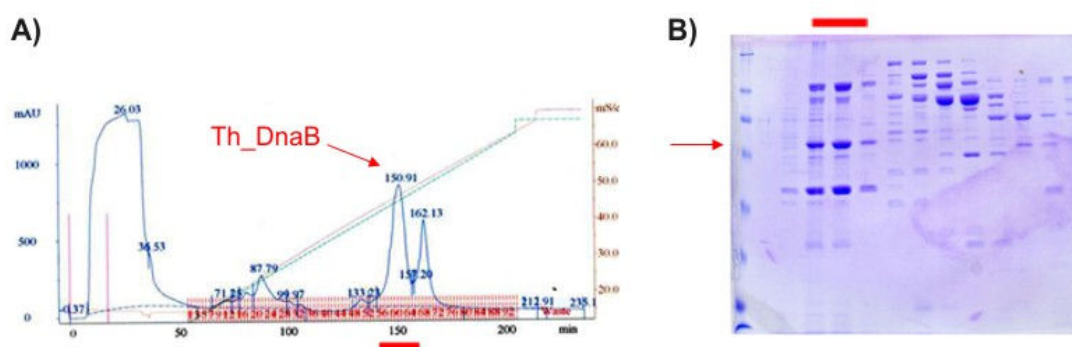


Figura R.7: Etapa de bescanvi aniònic de la purificació de Th_DnaB. A) Perfil cromatogràfic de la columna Hitrap Q XL. Es marca amb una fletxa taronja el pic d'absorbància corresponent a Th_DnaB i les fraccions d'elució que el conformen es subratllen en taronja. **B)** Anàlisi per SDS-PAGE de les fraccions cromatogràfiques corresponents al pic de Th_DnaB (marcades amb una línia taronja gruixuda). La fletxa taronja indica la banda corresponent a Th_DnaB.

Per eliminar aquestes dues contaminants, es carregà la mostra a una columna d'exclusió molecular. Th_DnaB eluí a un volum corresponent a un hexàmer de la proteïna, formant un pic aïllat però que tenia lleugers colzes a ambdues bandes. A la figura R.8 es representa el perfil cromatogràfic i l'anàlisi de les fraccions eluïdes de la columna Superose 6 10/300 GL.

En aquest punt, la mostra encara presentava algunes proteïnes contaminants. Es decidí sotmetre la mostra eluïda de la filtració en gel a un pas de xoc tèrmic suau (10' a 65° C), que va permetre precipitar les contaminants.

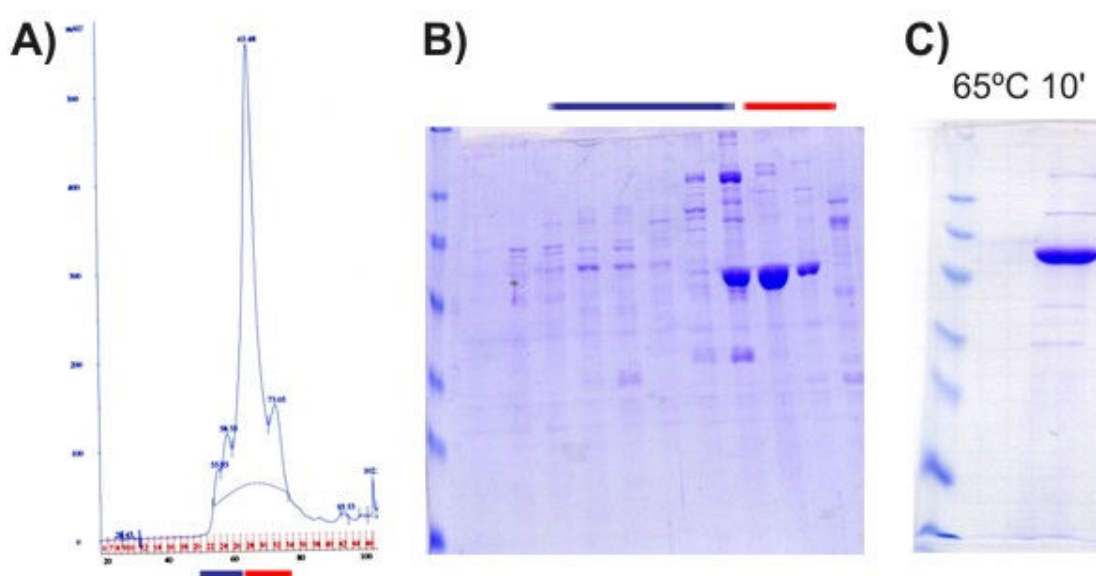


Figura R.8: Etapes de gel filtració i xoc tèrmic de la purificació de Th_DnaB. **A)** Cromatograma de la columna Superose 6 10/300 GL. Les fraccions corresponents al pic d'absorbància es marquen en blau (colze anterior) i en taronja (pic majoritari). **B)** Anàlisi per SDS-PAGE de les fraccions cromatogràfiques. Es segueix el mateix codi de colors que a (A). **C)** Anàlisi de la mostra proteica després de l'etapa final de xoc tèrmic.

La puresa final de la mostra, comprovada per SDS-PAGE tenyit amb blau de Coomassie (veure figura R.8c), es considerà suficient per a l'etapa de cristal·lització. Finalment, la identitat de la proteïna es comprovà per espectrometria de masses.

El rendiment final de les etapes d'expressió i purificació de Th_DnaB és aproximadament de 2.3 mg de proteïna pura per litre de cultiu bacterià inicial

2.4 Cristal·lització

Les primeres proves de cristal·lització per a Th_DnaB es van fer manualment en microgotetes. Posteriorment, es van repetir tots els *screens* en nanogotes amb el robot de cristal·lització Cartesian Microsys 4000 XL.

Es varen provar totes les condicions disponibles a la plataforma automatitzada de cristal·lografia (PAC) a 4° i 20° i a tres concentracions diferents de proteïna 4, 8 i 11 mg/ml. A la concentració més baixa, la majoria de les condicions testades resultaven en gota transparent. En canvi, tant a 8 com a 11 mg/ml s'obtenien tant condicions amb precipitat cristal·lí com transparents.

Només s'obtingueren “eriçons” (veure figura R.9) en un parell de condicions de 20°C que tenien com a precipitant MPD. Els esforços per intentar millorar aquests cristalls foren infructuosos, tot i provar múltiples combinacions de precipitants, sals, tampons, concentracions i volums de gota.

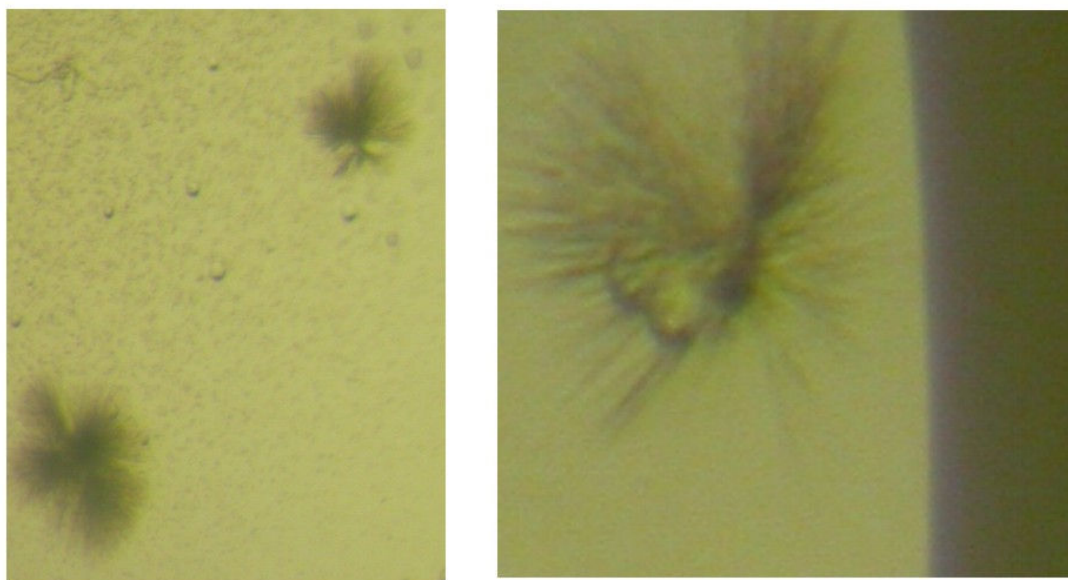


Figura R.9: Cristall en eriçó de la DnaB de *Thermotoga maritima*.

Finalment, es va provar el *cross-seeding* amb cristalls d'Aq_DnaB esmicolats. Es varen repetir els *screens* corresponents a LMB1, LMB2, LMB12 i LMB13, sense èxit. Tampoc funcionà la metilació de les lisines exposades de Th_DnaB, doncs la mostra metilada no generà noves condicions de cristallització positives.

3. Domini C-terminal de la DnaB d' *A. aeolicus*: Aq_DnaB-CT

3.1 Estratègia de clonatge

Donat que els cristalls obtinguts amb la proteïna Aq_DnaB sencera no eren de prou qualitat i no es van poder millorar, es va decidir provar de cristal·litzar només el domini C-terminal de DnaB.

El punt d'inici d'aquest nou constructe es determinà en relació a l'alineament de seqüència amb l'estructura coneguda del domini N-terminal de la DnaB d'*E. coli* (Fass *et al.* 1999; Weigelt *et al.* 1999) i les proves biofísiques de definició de dominis descrites a la bibliografia i que s'han detallat a la introducció d'aquest treball (Nakayama *et al.* 1984; Biswas *et al.* 1994; Miles *et al.* 1997). A la figura R.10 es mostra l'alineament de seqüència entre les DnaB d'*E. coli* i *A. aeolicus* superposat amb l'estructura secundària determinada pel domini N-terminal de la proteïna d' *E. coli*. Igualment, s'indiquen els límits entre dominis proposats a la bibliografia.

Paral·lelament, es decidí escurçar també la proteïna pel seu extrem C-terminal per tal d'obtenir un domini més compacte, segons les prediccions de desordre i els alineaments de seqüència amb la resta d'helicases DnaB bacterianes descrites (veure estratègia de clonatge d'Aq_DnaB [1-439]).

Amb aquest doble objectiu es dissenyaren uns encebadors de PCR que permetien incorporar un codó d'iniciació a la posició corresponent al residu T151 d'Aq_DnaB i un codó de parada després del residu E438. Aquest fragment amplificat es clonà directament al vector d'expressió pET29a. Mitjançant seqüenciació es comprovà que s'havien generat correctament els codons d'iniciació i parada a les posicions escollides, que la seqüència d'interès es trobava correctament inserida al vector pET29a entre les dianes per a NdeI i XhoI, i que aquesta no presentava cap mutació.

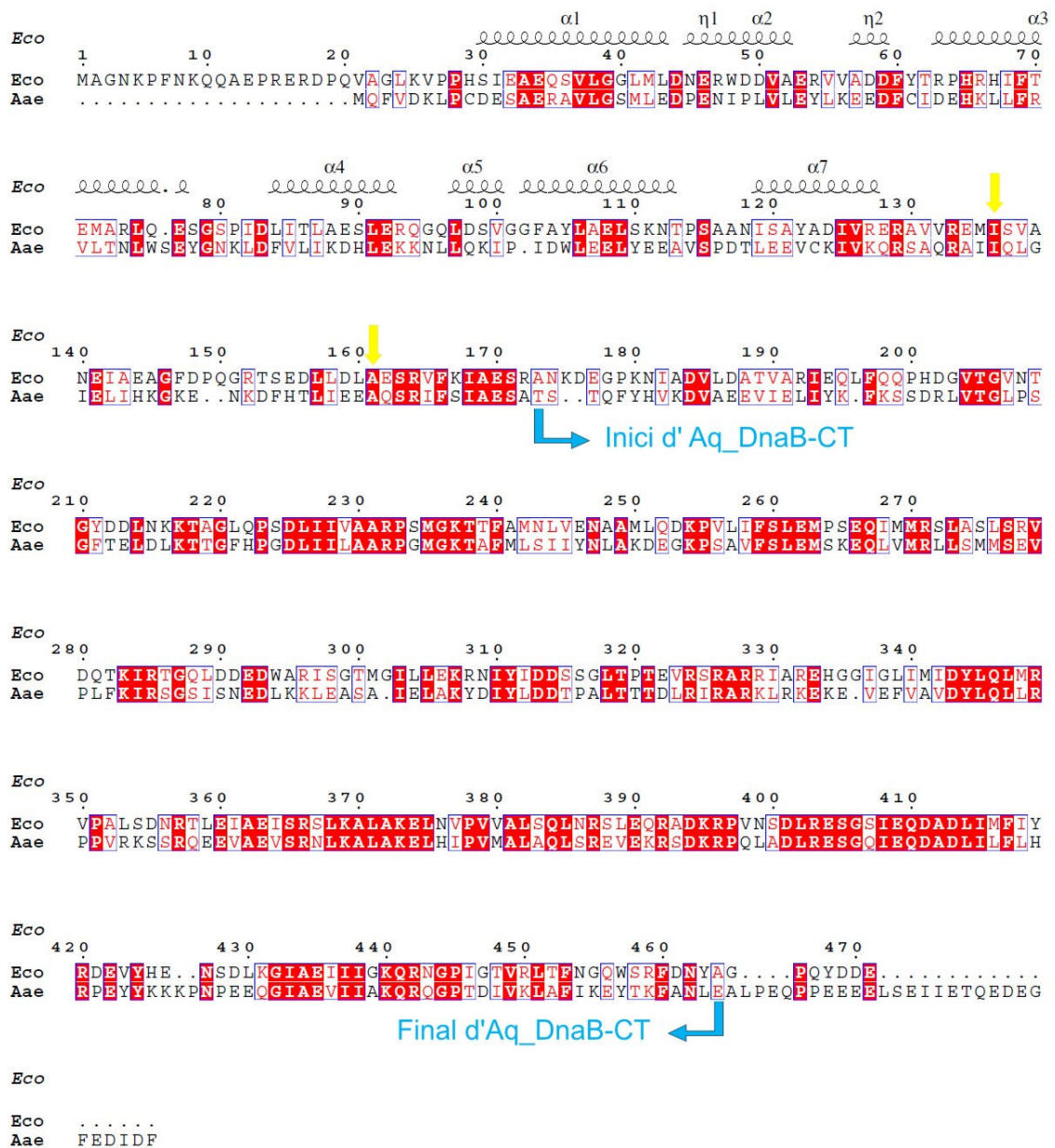


Figura R.10: Alineament de seqüència d'Aq_DnaB enfront la DnaB d'*E. coli*. Es mostra l'estructura secundària del domini N-terminal de la DnaB d'*E. coli* (codi PDB: 1B79) superposada a l'alineament. Les fletxes grogues indiquen la frontera del domini N-terminal definida, respectivament, per proteòlisi limitada (primera fletxa) i ressonància magnètica nuclear (segona fletxa). La regió d'enllaç entre dominis s'havia estimat d'uns 45 residus de llargària. Les fletxes blaves indiquen l'aminoàcid d'inici i de final del constructe Aq_DnaB-CT.

3.2 Expressió

El protocol d'expressió d'Aq_DnaB-CT es va basar en el prèviament determinat per a la proteïna Aq_DnaB sencera. S'emprà la soca *E. coli* BL21 (DE3), LB com a medi de cultiu, una temperatura d'expressió de 17°C, una concentració final d'inductor de 0.1 mM i un temps d'expressió total de 3 dies.

3.3 Purificació

Les etapes d'extracció i clarificació es van fer igual que per a la proteïna sencera, adaptant l'etapa de xoc tèrmic a la sensibilitat del nou constructe proteic: 75°C durant 15'.

Ara bé, calgué modificar el protocol de purificació cromatogràfica degut a que el pH emprat per l'extracció i el xoc tèrmic (pH 7.6) es trobava molt proper al punt isoelèctric teòric d'Aq_DnaB-CT, que és de 7.74. Això presentava problemes a l'etapa de bescanvi iònic, doncs per a que aquesta estratègia de purificació funcioni correctament cal que la proteïna presenti una càrrega neta total no propera a la neutralitat, cosa que implica una diferència mínima de 0.5-1 unitats entre el pH de treball i el pI de la proteïna que es vol separar.

Per a superar aquest entrebanc, es dissenyà un protocol d'optimització en que es provaren tampons amb vuit valors de pH diferents (de 5.5 a 9) i, conseqüentment, bescanviadors catiònics i aniònics per a la primera etapa cromatogràfica. Durant aquesta optimització, el tampó en que s'havia realitzat el xoc tèrmic era substituït mitjançant diàlisi seqüencial per l'adequat per aquesta etapa cromatogràfica. Això provocà la precipitació de part de la mostra en alguns dels canvis de tampó, especialment quan s'empraven tampons amb valors de pH molt bàsics.

A la figura R.11 s'esquematitzen els passos de diàlisi seguits per aconseguir portar la mostra inicial als vuits valors de pHs que es provaren, alhora que s'indica si es produí precipitació de la mostra en algun d'ells.

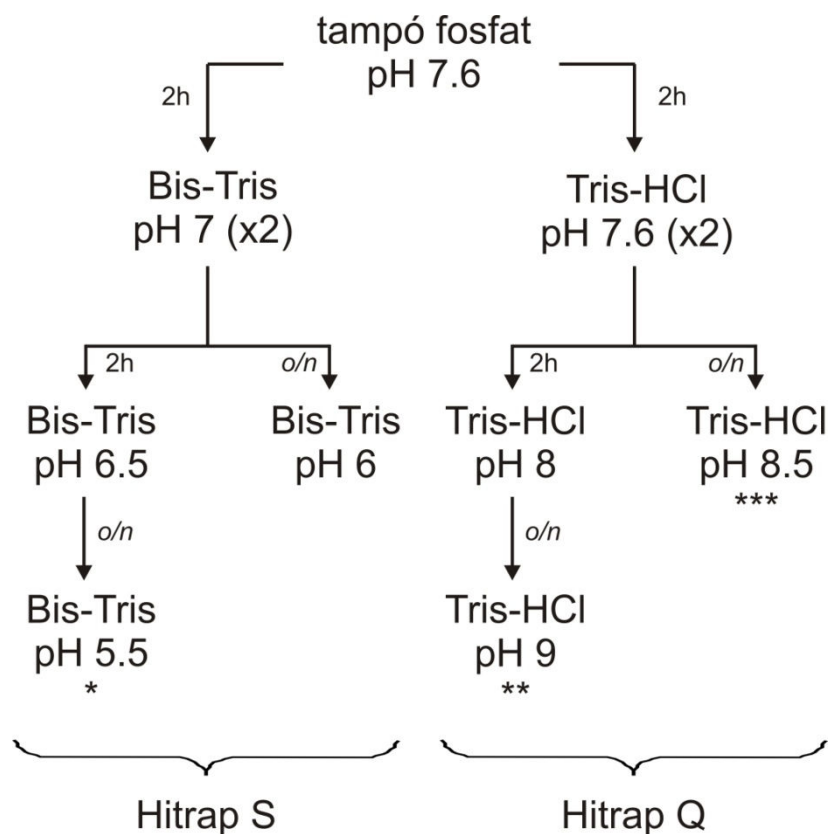


Figura R.11: Esquema de les etapes de diàlisi prèvies a la columna de bescanvi iònic. Les fletxes representen passos de diàlisi, s'indica en cada cas la durada d'aquests. Tot el procés es realitzà a 4°C i en agitació suau. Es provaren tampons a 8 valors de pH diferents. Els de la part esquerra del gràfic (pHs 7, 6.5, 6 i 5.5) eren adequats per a carregar la mostra a un columna de bescanvi catiònic. Els de la part dreta (pHs 7.6, 8, 8.5 i 9), per a bescanvi aniònic. Per a la composició detallada dels tampons, consultar l'apartat de materials i mètodes El símbol (x2) indica que, per a les mostres a pH final de 7 i 7.6, es realitzaren dos canvis de diàlisi de 2h cadascun. S'indiquen els canvis de diàlisi durant els quals es produí precipitació de la mostra proteïca amb els símbols: * lleugera, ** abundant, *** molt abundant.

Cal remarcar que també es provà de fer directament l'extracció de la proteïna al pH que s'empraria per a la columna de bescanvi, però això no donà bons resultats, doncs la quantitat d'Aq_DnaB-CT soluble que es recuperava era menor. Igualment, a l'etapa de xoc tèrmic es perdia més proteïna, particularment a valors de pH superiors a 8, on no es pot emprar fosfat com agent tamponador.

Quan s'empraven intercanviadors aniònics, la proteïna gairebé no s'unia a la reïna, sinó que s'eluí a l'etapa de rentat o molt aviat durant el gradient de força iònica. A més, no ho feia en un pic aïllat, sinó amb moltes proteïnes contaminants. Per tant, es

descartà aquest mètode de purificació, doncs no era prou resolutiu per separar la proteïna d'interès de les contaminants.

En canvi, la purificació a valors de pH més àcids i emprant intercanviadors catiónics donà molt bons resultats, tal i com es mostra a la figura R.12. A 3 dels 4 valors de pH provats (5.5, 6 i 6.5) Aq_DnaB-CT s'unia al bescanviador catiónic, mentre que la majoria de les contaminants presents a la mostra inicial no eren retingudes i s'elueïen a l'etapa de rentat. A pH 7, el més proper al pI teòric, la proteïna s'unia molt dèbilment a la reïna i començava a eluir-se a la part final de l'etapa de rentat de la columna.

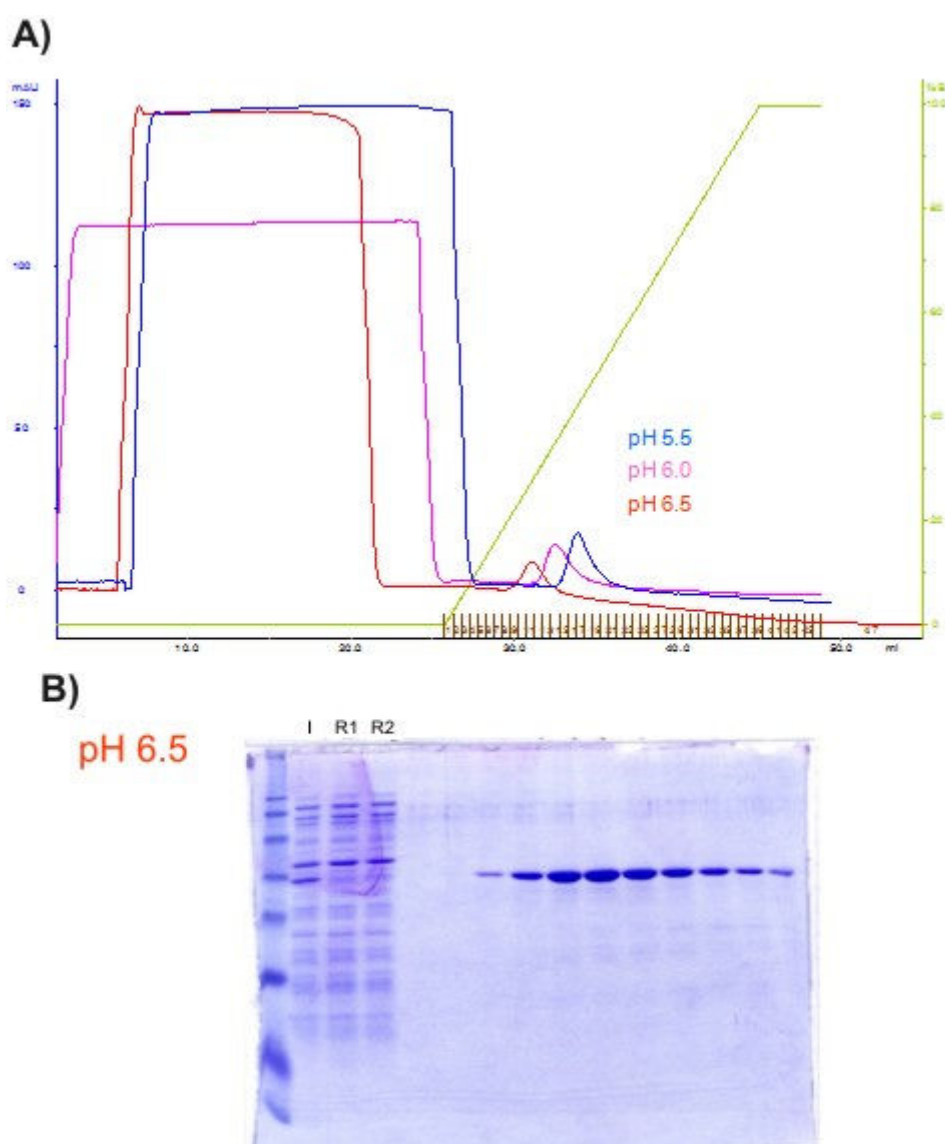


Figura R.12: Optimització de l'etapa de bescanvi catiónic per a Aq_DnaB-CT. A) Superposició dels perfils cromatogràfics obtinguts sota les mateixes condicions de flux i gradient iònic emprant valors de pH diferents: en blau a pH 5.5, en lila a 6 i en vermell a 6.5. **B)**

Anàlisi per SDS-PAGE de les fraccions cromatogràfiques eluïdes de la columna equilibrada a pH 6.5. I: mostra carregada a la columna, R1 i R2: etapes de *flow-through*.

A la figura R.12 es pot observar com el pic corresponent a Aq_DnaB-CT es desplaça cap a la dreta a mesura que es disminueix el pH de treball. Això indica que la proteïna s'uneix cada vegada amb més força a la columna, doncs cal més força iònica per eluir-la. La proteïna s'elueix en tots els casos testats en un sol pic aïllat que, quan s'analitzà amb SDS-PAGE, es comprovà que presentava molt poques contaminants. Es va escollir treballar a pH 6 perquè la precipitació de la mostra durant l'etapa de canvi de tampó per diàlisi era menor i, conseqüentment, la recuperació major.

Després d'aquesta etapa, i per tal de comprovar en quin estat oligomèric es trobava la proteïna, es realitzà un cromatografia de gel filtració (Superdex200). La proteïna eluïa majoritàriament a un pic corresponent a un dímer, per bé que també hi havia una petita fracció de proteïna que formava agregats d'alt pes molecular. Si l'etapa de gel filtració es realitzava a pH al voltant de 7.6, s'eliminava aquest pic d'agregats, es reduïa molt el pic corresponent a dímers d'Aq_DnaB-CT i apareixia un pic majoritari corresponent a la forma hexamèrica de la proteïna.

El rendiment final mitjà de les etapes d'expressió i purificació d'Aq_DnaB-CT és de 3 mg per litre de cultiu inicial. La identitat de la mostra purificada es comprovà per *molecular fingerprinting*.

3.4 Cristal·lització

Es realitzaren proves de cristal·lització en nanogotes amb les fraccions tant de dímer a pH 6.5 com d'hexàmer a pH 7.6. Es provaren totes les condicions disponibles a la plataforma automatitzada de cristal·lografia (unes 1500) a dues temperatures diferents i varies concentracions proteiques (4 i 8 mg/ml per al dímer i 4, 7 i 11 mg/ml per a l'hexàmer). Per a ambdues formes quaternàries, la majoria de les condicions de cristal·lització testades eren gota transparent, tot i incrementar la concentració fins al

valor màxim a que es podia concentrar la proteïna sense que comences a precipitar (8 mg/ml per la forma dimèrica i 11 mg/ml per l'hexamèrica).

Finalment, es provà també *cross-seeding*, emprant com a llavors cristalls de la proteïna Aq_DnaB sencera esmicolats, per als *screens* de cristal·lització més generals, però sense èxit.

4. Forma curta de la DnaB d'*A. aeolicus*: Aq_DnaB [1-439]

4.1 Estratègia de clonatge

Amb l'objectiu de dissenyar un nou constructe de la proteïna Aq_DnaB es va realitzar una alineament de seqüència enfront la resta de proteïnes DnaB bacterianes descrites, així com una predicció de desordre i d'estructura secundària de la proteïna sencera. Es pretenia treballar amb una forma més compacte d'aquesta helicasa, eliminant les possibles zones no ordenades o altament flexibles, que podrien ser la causa del baix poder de difracció dels raigs X que mostraven els cristalls obtinguts amb la proteïna sencera.

A la figura R.13 es mostra el resultat dels alineaments de seqüència i la predicció d'estructura secundària per Aq_DnaB.

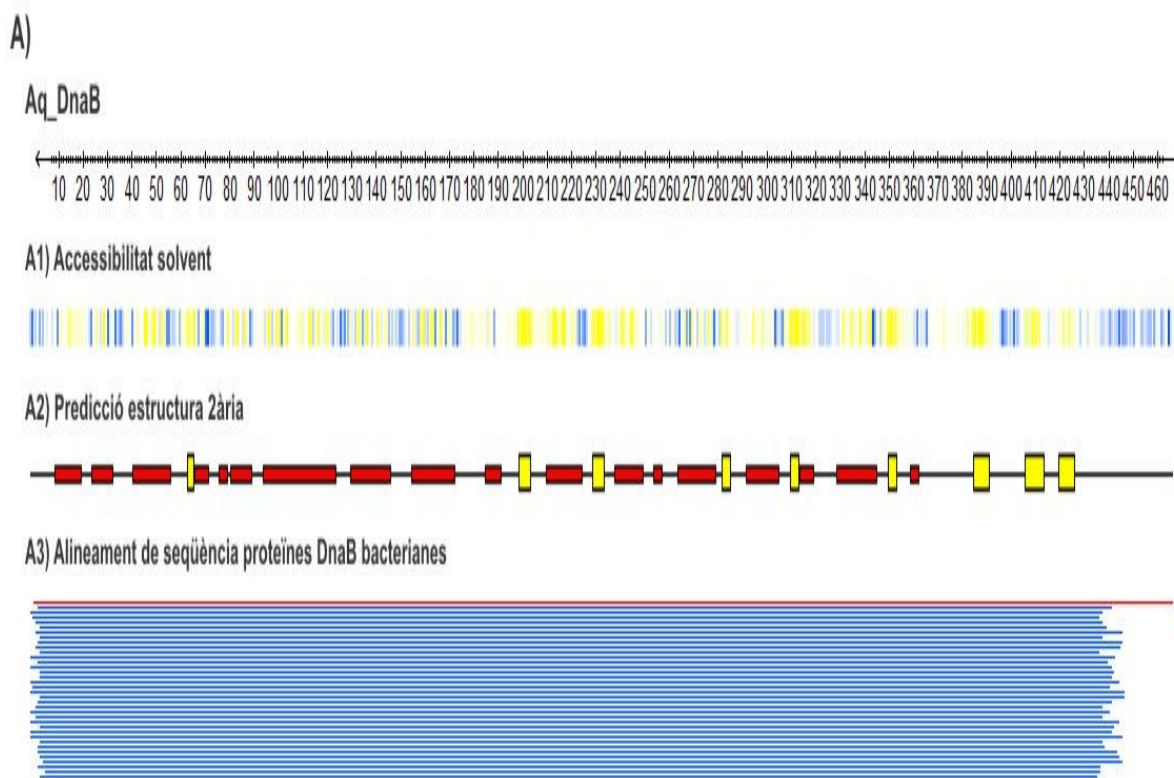


Figura R.13a: Anàlisi bioinformàtica d'Aq_DnaB (veure llegenda a la pàgina següent).

B)

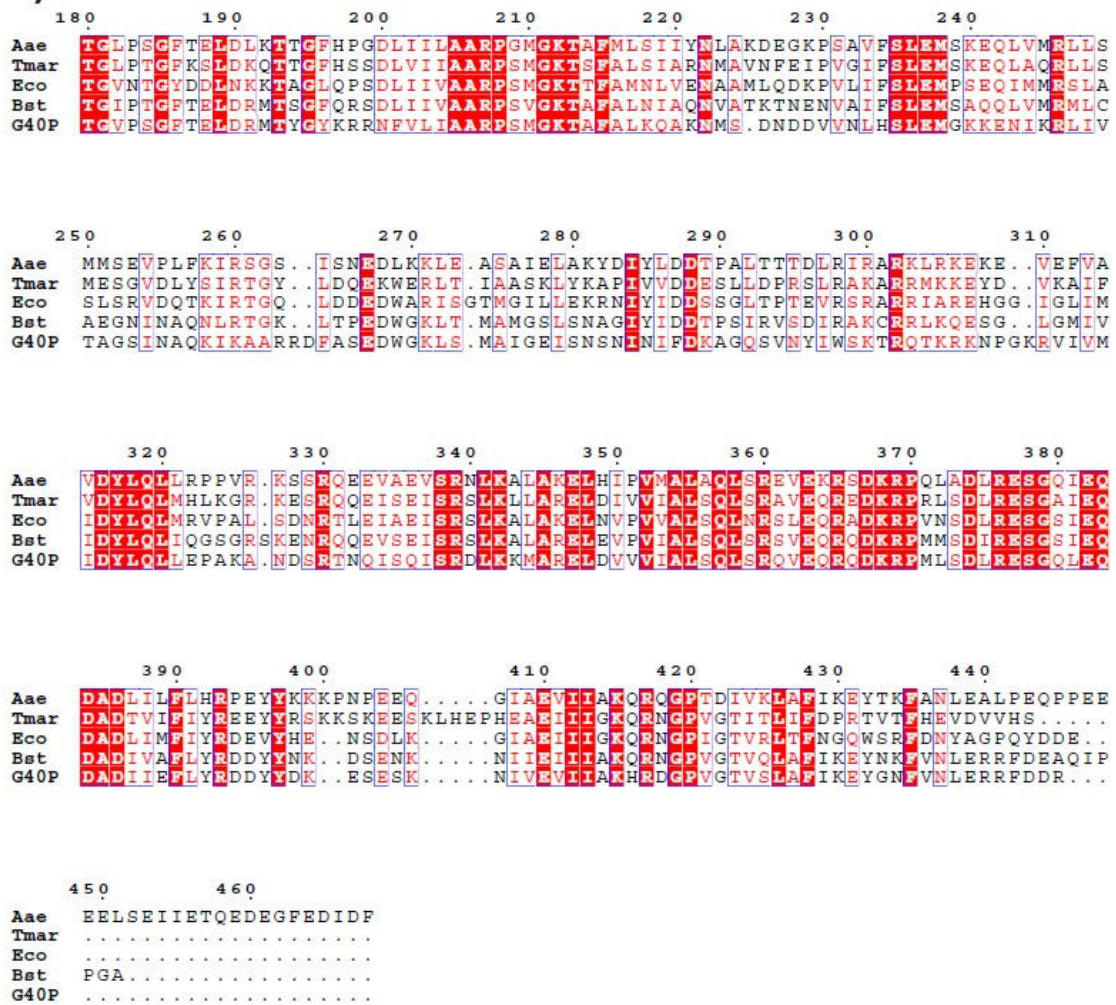


Figura R.13b: Anàlisi bioinformàtica d'Aq_DnaB. A) Esquema general de la predicció d'estructura secundària, d'accessibilitat al solvent i de l'alineament. A la part superior es representa la seqüència aminoacídica d'Aq_DnaB, en relació de la qual estan dibuixades la resta de figures. **A1)** Representació de l'accessibilitat al solvent dels diferents residus, com més blau més accessible. **A2)** Representació de l'estructura secundària on són les hèlix α es representen com un rectangle vermell, les fulles β com un rectangle groc i la línia senzilla representa les zones de *coiled coil*. **A3)** Alineament global enfront de la resta de DnaB bacterianes descrites (més de 80 en el moment de fer l'anàlisi), on la línia taronja superior representa Aq_DnaB i cada línia blava representa una espècie bacteriana diferent. Figura realitzada amb el programa Predict, que empra PSI-BLAST per l'alineament i PSPRED per a la predicció d'estructura secundària. **B)** Detall de l'alineament de la regió C-terminal de les proteïnes DnaB de diferents espècies respecte a Aq_DnaB, mostrant la conservació de seqüència entre elles. Els organismes representats són els següents: Aae: *Aquifex aeolicus*, Bst: *Bacillus stearothermophilus*, Tma: *Thermotoga maritima*, Eco: *Escheria coli*, SPP1: bacteriòfag SPP1, Figura realitzada amb el programa ESPrit 2.2.

D'aquest anàlisi bioinformàtic es desprèn que la majoria d'helicases replicatives bacterianes són més curtes que Aq_DnaB, així com que la part C-terminal d'aquesta

proteïna és una zona molt exposada al solvent i, segons la predicció d'estructura secundària, es tracta d'una regió molt flexible, doncs es troba en *coiled coil*.

A la figura R.14 es mostra un resum amb la localització dins la seqüència aminoacídica de la proteïna de les zones de desordre predites per a la DnaB d' *A. aeolicus* sencera.

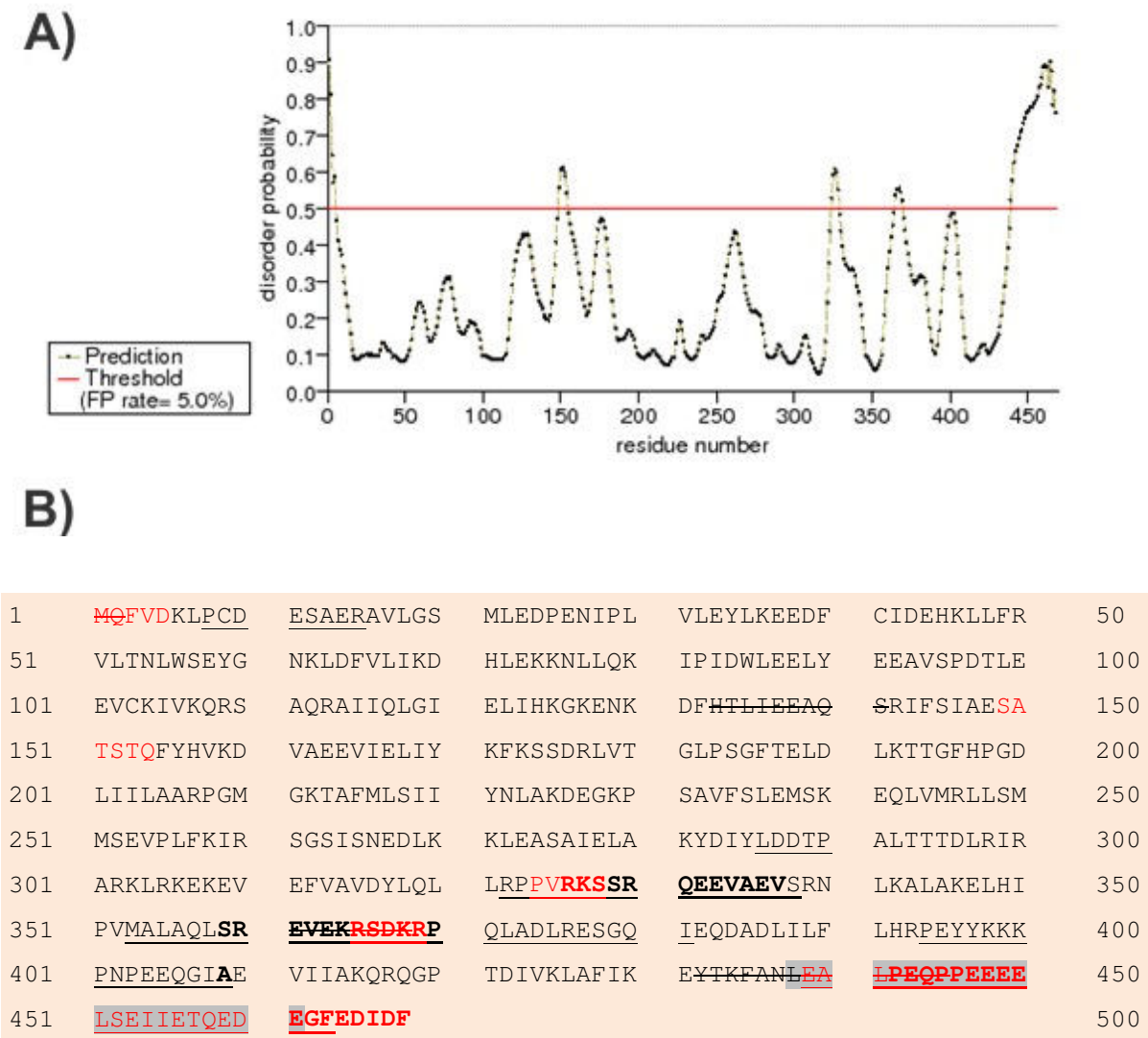


Figura R.14: Predicció de les zones desordenades d'Aq_DnaB. **A)** Gràfica generada amb el programa PrDOS on es mostra la probabilitat de desordre al llarg de la cadena aminoacídica d'Aq_DnaB. La zona més C-terminal de la proteïna, aproximadament a partir del residu 439, es troba clarament per sobre del llindar de significació. **B)** Detall de la localització de les zones desordenades a la seqüència d'Aq_DnaB. Es mostren en vermell els residus desordenats predits pel programa PrDOS, subratllats els predits pel programa RONN, en negreta els de DisEMBL, ressaltats en gris els de Predict SEG *low complexity regions* i taxats els predits per DISOPRED.

Tots els programes emprats prediuen que la regió més C-terminal de l'helicasa està desordenada. Varia lleugerament la llargada d'aquesta zona, però la seva localització és al voltant del residu 440. La resta de zones desordenades predites per a Aq_DnaB són més petites i la seva localització varia significativament depenent del programa emprat, cosa que indica que són menys probables i depenen del mètode d'anàlisi utilitzat.

Així doncs, d'aquests dos tipus d'anàlisis bioinformàtiques es poden inferir bàsicament dues conclusions de la DnaB d'*A. aeolicus*:

- que la regió més C-terminal de la proteïna sencera està poc conservada entre les diferents espècies bacterianes, així com també a bacteriòfags, i a la majoria no hi és present,
- que aquesta zona C-terminal de la proteïna, no conservada entre espècies, té una alta propensió a estar desordenada i no sembla tenir estructura secundària definida.

Tenint en compte aquests resultats i per tal d'obtenir una nova forma proteica amb més propensió a formar cristalls de bona qualitat, es decidí eliminar aquesta zona C-terminal desordenada i no conservada. S'incorporà un codó de parada darrera del residu Glu439 de la proteïna sencera, generant així la proteïna Aq_DnaB [1-439], a la qual li manquen els darrers 29 residus a l'extrem C-terminal. S'escollí específicament aquest residu perquè, segons les prediccions d'estructura secundària, sembla ser l'últim d'una hèlix α i, a partir d'aquest, la zona està més desestructurada, tal i com mostren les prediccions de desordre.

La seqüenciació de dos dels clons positius obtinguts durant l'etapa de clonatge mostrà que en ambdós casos s'havia clonat correctament el gen codificant per a Aq_DnaB [1-439] al vector d'expressió pET29a, entre les dianes de NdeI i XhoI, i que la seqüència de la proteïna d'interès no presentava cap mutació (llevat del codó *stop* introduït durant el procés de clonatge).

4.2 Expressió

El nivell d'expressió dels diferents clons d'Aq_DnaB[1-439] obtinguts es va testar per escollir aquells que presentessin els nivells de sobreexpressió de la proteïna d'interès més elevats.

A continuació, es va procedir a la sobreexpressió d'Aq_DnaB[1-439] a gran escala, modificant lleugerament les condicions optimitzades prèviament per a la proteïna sencera. El protocol final emprava la soca *E. coli* BL21(DE3), una concentració d'IPTG de 0.1 mM, una temperatura d'expressió de 15°C i un temps total d'inducció de 3 dies.

4.3 Purificació

El protocol de purificació també es va basar en el prèviament determinat per a la proteïna sencera, canviant, però, la temperatura i la durada del xoc tèrmic a 75° C i 15', doncs aquest constructe presenta una resistència lleugerament menor a les temperatures elevades.

La fracció soluble després de la precipitació per tractament tèrmic es va sotmetre a una cromatografia de bescanvi aniònic emprant la columna HitrapQ XL de 5ml, equilibrada en front el tampó A1 i eluïda mitjançant un gradient lineal de 22 volums de columna fins assolir la força iònica del tampó B1.

A la figura R.15 es mostra el perfil cromatogràfic d'aquesta etapa de purificació, així com l'anàlisi per SDS-PAGE de l'eficiència del tractament de xoc tèrmic previ i de les fraccions eluïdes d'aquest primer pas cromatogràfic.

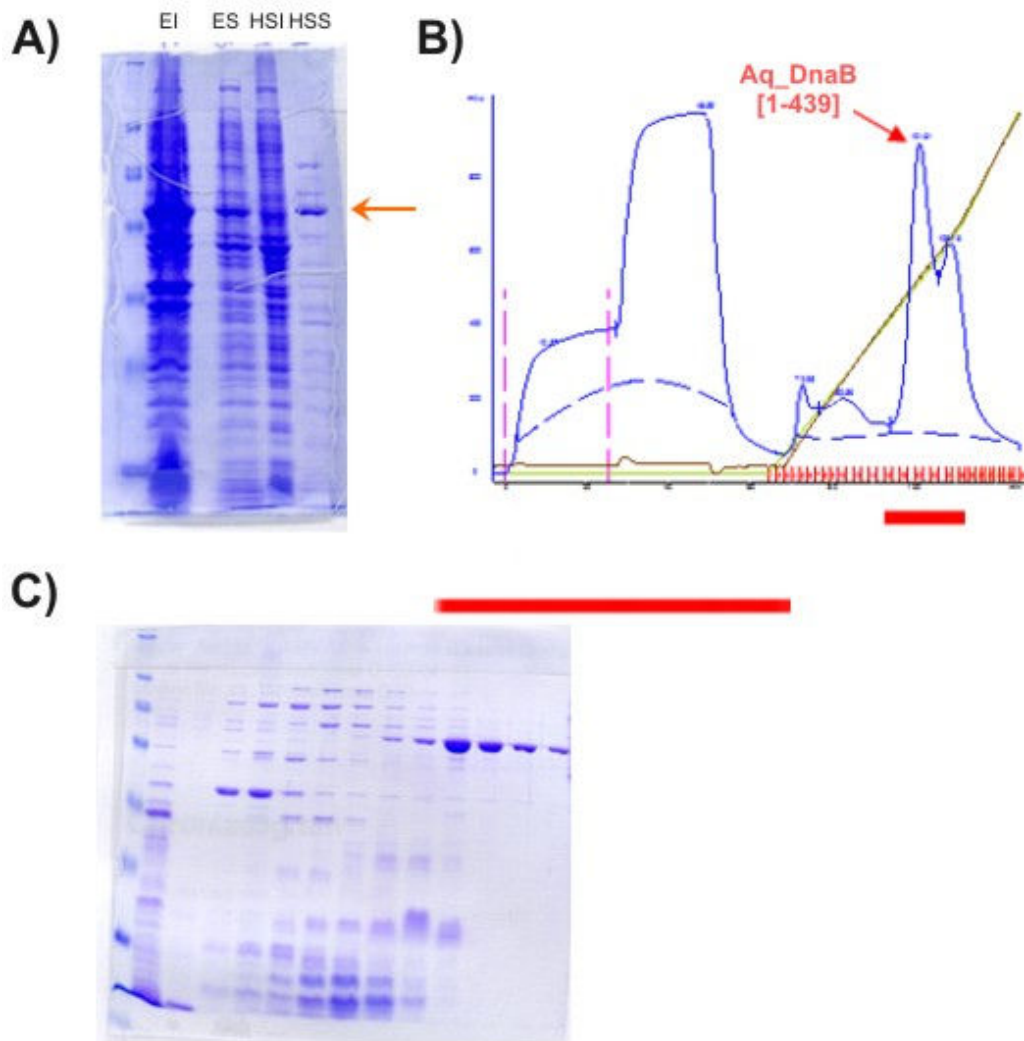


Figura R.15: Etapes de xoc tèrmic i de bescanvi aniònic de la purificació d'Aq_DnaB[1-439]. **A)** Anàlisi per SDS-PAGE de la fracció expressada de forma soluble i de l'etapa de xoc tèrmic. EI: fracció expressada de manera soluble; ES: fracció expressada de forma soluble; HSI: fracció insoluble de l'etapa de xoc tèrmic; HSS: fracció soluble de l'etapa de xoc tèrmic. **B)** Cromatograma de la columna Hitrap Q XL 5ml. La línia blava és el valor d'absorbància de les fraccions eluïdes a 280 nm i la línia marró és el percentatge de tampó B1 al llarg del gradient de força iònica. Es marca amb una fletxa taronja el pic d'absorbància corresponent a Aq_DnaB[1-439] i les fraccions que el conformen es subratllen també en taronja. **C)** Anàlisi de les fraccions corresponents al pic d'Aq_DnaB[1-439], marcades en taronja com a (B).

Aq_DnaB[1-439] s'elueix al pic d'absorbància comprès entre les fraccions corresponents a un 60 % de gradient del tampó B1 i n'és la proteïna més abundant. Cal destacar que la relació entre les absorbàncies a 260 i 280 nm indica que la mostra conté contaminació d'àcids nucleics, cosa que també es comprovà per tinció del gel amb bromur d'etidi.

Per tal d'eliminar les darreres contaminacions, tant proteiques com d'àcids nucleics, i comprovar que la mostra no estava agregada, es realitzà una darrera etapa de purificació mitjançant una columna de gel filtració, el perfil cromatogràfic i l'anàlisi de les fraccions eluïdes de la qual es mostra a la figura R.16.

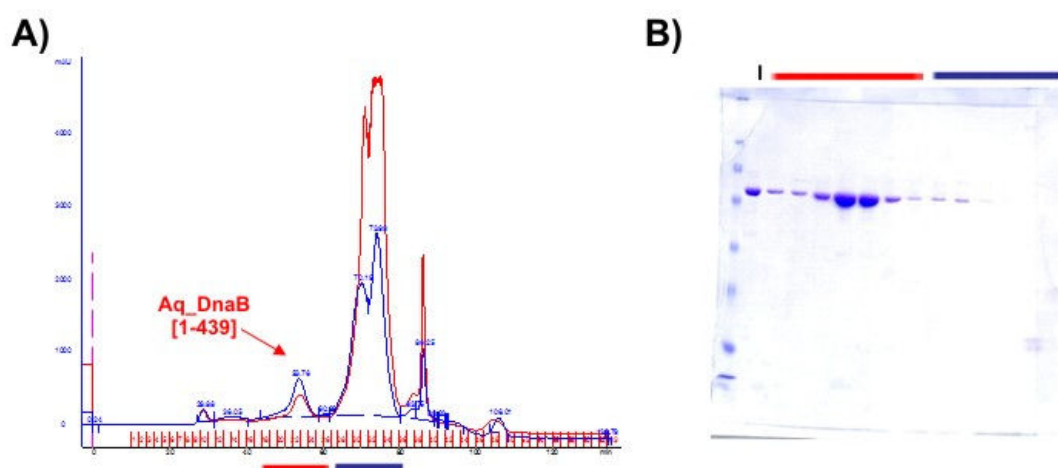


Figura R.16: Etapa de cromatografia de gel filtració de la purificació d'Aq_DnaB[1-439]. **A)** Perfil cromatogràfic de les fraccions eluïdes de la Superose 6 10/300 GL. La línia blava és l'absorbància a 280nm i la vermella a 260 nm. El primer pic, subratllat en taronja, correspon a Aq_DnaB. El segon pic, subratllat en blau, és majoritàriament DNA. **B)** Anàlisi per SDS-PAGE de les fraccions eluïdes de (A). I: mostra carregada a la columna. Els carrils corresponents a les fraccions del pic on s'elueix Aq_DnaB[1-439] estan marcades amb una franja taronja i les del pic de DNA blava. Les franges són equivalent a les de (A).

La proteïna s'elueix en un sol pic d'absorbància al volum d'elució esperat per a un hexàmer d'aquesta proteïna. Hi ha un segon pic d'absorbància, particularment alt per a 260 nm, corresponent al DNA.

Després d'aquesta etapa de purificació, la puresa de la mostra es considerà suficient per iniciar l'etapa de cristal·lització. Com a pas final, es comprovà la identitat de la mostra mitjançant la tècnica de *molecular fingerprinting*, que donà un *score* de 998 quan, per aquest cas, un *score* major de 42 ja identifica la proteïna amb una probabilitat mínima del 95%.

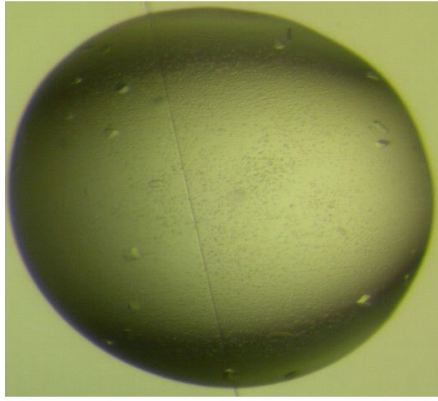
El rendiment mitjà de les etapes d'expressió i purificació és de 1.5 mg de proteïna Aq_DnaB [1-439] pura per litre de cultiu bacterià inicial.

Cal destacar que, per poder produir proteïna Aq_DnaB[1-439] derivatitzada amb seleno-metionina, es van haver de suplementar els cultius amb glucosa i vitamines aproximadament cada 24h, així com augmentar el temps total de sobreexpressió a quatre dies. Si no es feia així, no s'aconseguia sobreexpressió d'Aq_DnaB[1-439]. El rendiment mitjà de l'expressió i la purificació de la proteïna derivatitzada és molt similar al de la proteïna nativa, aproximadament 1.3 mg de proteïna amb seleno-metionina per litre de cultiu bacterià inicial.

4.4 Cristal·lització

Durant l'etapa de cribatge es van provar un total de 1530 solucions de cristal·lització diferents, a dues temperatures (4° i 20°C) i dues concentracions de proteïna (3.7 i 7.4 mg/ml) diferents. Això representa més de 6000 proves de cristal·lització en nanogotes, de les quals es van obtenir només tres condicions que generaven microcristalls. A la figura R.17 es mostren les fotografies corresponents a aquestes tres condicions, així com la composició de la solució de cristal·lització, la concentració de proteïna i la temperatura de la prova. Cal destacar que totes eren proves de cristal·lització a 4° C, tenien una sal de magnesi al tampó de cristal·lització i PEG com a precipitant. Cap altra condició generà resultats positius, susceptibles de ser optimitzats.

Es varen escalar aquestes tres condicions positives a microgota, alhora que s'optimitzaven les concentracions dels agents precipitants i de la proteïna. És necessari escalar la prova de cristal·lització a volums majors per obtenir cristalls de mida suficient per a ser sotmesos a difracció de raigs X. Malauradament, la condició C no es va poder escalar satisfactòriament, doncs les plaques de microgota només produïren precipitat cristal·lí. Amb les dues altres condicions (A i B) sí que s'aconseguí reproduir els cristalls en microgota.



A)

Condicció cristal·lització:

Crystal-Screen Lite- n° 18

10% (v/v) PEG 8000

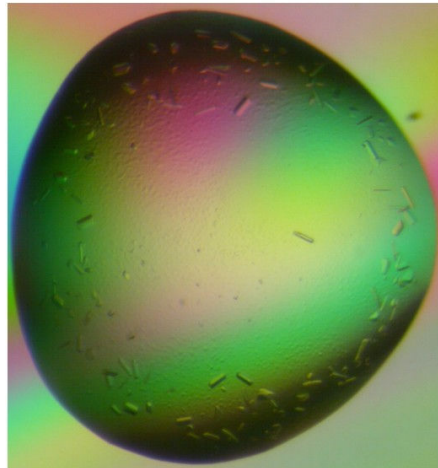
0.2 M acetat de magnesi

0.1 M cacodilat sòdic pH 6.5

Aq_DnaB [1-439] a 3.7 mg/ml

Temperatura: 4°C

Cristalls al cap de 8-9 dies.



B)

Condicció cristal·lització:

JBS4 - n° 22

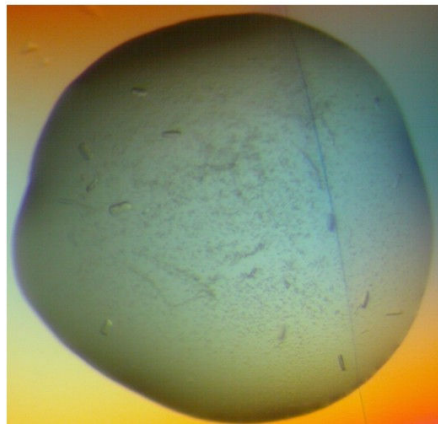
10% (v/v) PEG 8000

0.2 M acetat de magnesi

Aq_DnaB [1-439] a 3.7 mg/ml

Temperatura: 4°C

Cristalls al cap de 10 dies.



C)

Condicció cristal·lització:

Index - n° 22

15% (v/v) PEG 3350

0.1 M format de magnesi

Aq_DnaB [1-439] a 3.7 mg/ml

Temperatura: 4°C

Cristalls al cap de 7 dies.

Figura R.17: Condicions de cristal·lització positives en el procés de cribatge inicial. Les gotes de cristal·lització tenien un volum de 200 nl i contenien una barreja 1:1 de solució de precipitant i solució de proteïna Aq_DnaB [1-439]. La tècnica de cristal·lització emprada fou la difusió de vapor per gota seguda.

Les condicions A i B són pràcticament idèntiques, doncs només difereixen en que la primera està tamponada a pH 6.5 i la segona no està tamponada. Per això les millores d'ambdues condicions es feren de manera conjunta, variant la concentració de

precipitant i sal i provant diferents valors de pH (de 6 a 8.5). També es va provar de canviar el tipus de sal i el tipus de PEG.

Quan no es tamponava la condició de cristal·lització, s'obtenien cristalls en forma de barra. Inicialment, aquest cristalls en forma de barra no estaven complerts, sinó que es veien buits al seu interior. Amb un procés de *seeding* serial de 4-5 cicles es van aconseguir barres compactes amb aquesta condició. La figura R.18 mostra els cristalls inicials i els finals.

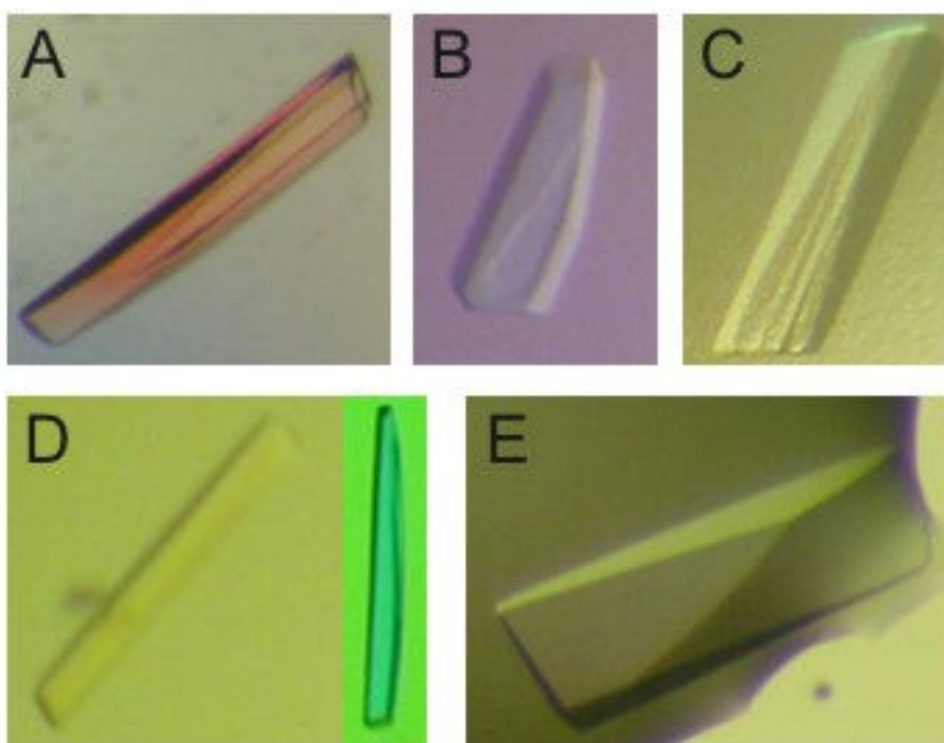


Figura R.18: Cristalls d'Aq_DnaB [1-439] a condicions sense tamponar. **A)** Primers cristalls obtinguts, amb la part interior buida. S'utilitzaren com a llavors per a la primera etapa de *seeding*. **B i C)** Cristalls obtinguts, respectivament, a la segona i tercera etapa del *seeding* serial. Les barres esdevenen progressivament més grans i compactes, per bé que encara presenten irregularitats en alguna de les seves cares. **D)** Cristalls en forma de barra compacta obtinguts a les darreres etapes del procés de millora. **E)** Cristall en forma de barra tridimensional obtingut després de 5 passos de *seeding* serial. A part de la millora en tridimensionalitat, també són de major mida que les primeres barres obtingudes.

Quan es tamponaven les condicions de cristal·lització, s'obtenien cristalls més geomètrics. A un pH entre 6.5 i 7.2 s'obtenien cristalls en forma de placa quadrada. A

valors de pH entre 7.4 i 7.6, s'obtenien cristalls més tridimensionals. La figura R.19 mostra ambdós tipus de cristalls.

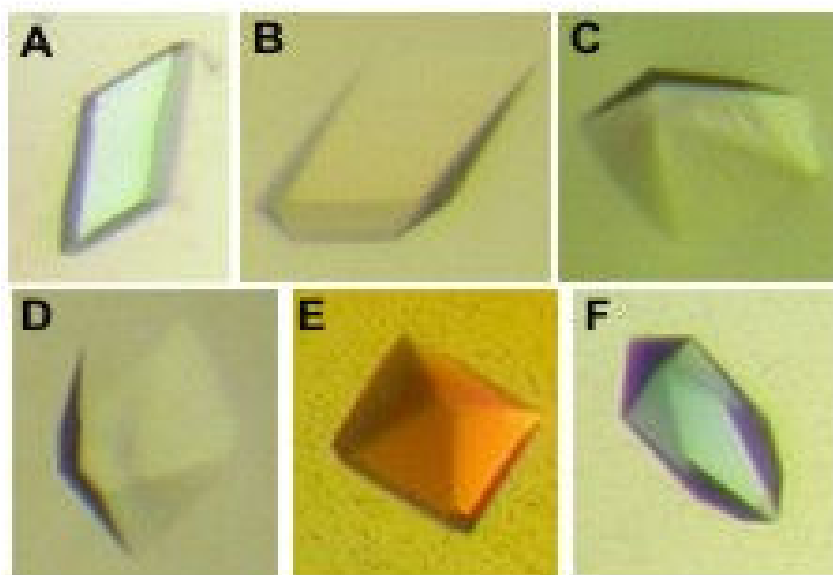


Figura R.19: Cristalls d'Aq_DnaB [1-439] crescuts en condicions de cristal·lització tamponades i cocrystals amb nucleòtids. A i B) Cristalls obtinguts a condicions tamponades a pH 6.5. C i D) Cristalls a condicions tamponades a pH 7.6. E) Cocrystal amb ATP- γ -S. F) Cocrystal amb ADP.

Tots aquests tipus de cristalls apareixien al cap de més de dues setmanes d'incubació a les condicions de cristal·lització, alguns fins i tot al cap d'un mes. Per comprovar que no cristal·litzava una forma degradada de la proteïna inicial, es dissolgueren varis cristalls en aigua i es carregà la mostra resultant a un gel SDS-PAGE, per tal de poder comparar-lo amb la mostra inicial. Es veié que no s'havien generat bandes addicionals.

També es van realitzar cocrystals amb ADP i dos anàlegs no hidrolitzables d'ATP: l'AMP-PNP i l'ATP- γ -S. A la figura R.20 es mostren fotografies d'aquests cristalls. S'obtenien resultats més reproduïbles quan la proteïna emprada per a les proves de cristal·lització era prèviament incubada amb el nucleòtid (a la concentració final desitjada) durant un mínim de 1h a temperatura ambient. L'addició directe del nucleòtid a la solució de cristal·lització mentre es posaven les gotes sovint provocava que aquest precipités.

Alternativament, es provà de metilar els residus de lisina exposats de la proteïna Aq_DnaB [1-439]. Aquesta tècnica de millora precristal·lització es basa en canviar les propietats de la superfície de la proteïna amb la intenció de generar una nova forma que pugui establir noves xarxes cristal·lines i/o cristalls més ben empaquetats. Es feren de nou tots els *screens*, però no s'obtingueren condicions de cristal·lització noves. Els cristalls que s'obtenien en condicions similars a les de la proteïna no metilada no generaven dades de difracció de raigs X de major qualitat.

Per últim, es va provar de cristal·litzar Aq_DnaB [1-439] en complex amb un DNA monocadena de 21 bases. Tal i com quedà palès durant la purificació, DnaB uneix DNA de forma inespecífica i s'utilitzà aquesta característica per generar el complex. S'escollí un ssDNA de la llargada que protegeix l'helicasa en experiments d'unió a ssDNA (Bujalowski *et al.* 1995; Jezewska *et al.* 1996), amb una seqüència que no mostres propensió a formar forquetes. Tal i com s'ha descrit a mètodes, es barrejà en calent (55° C) a una relació 1:1 el ssDNA i l'hexàmer de DnaB i es purificà el complex per gel filtració. Els cristalls apareixeren en condicions molt semblant als de la proteïna sense DNA. S'optimitzaren afegint una capa d'oli sobre el reservori que contenia la solució de cristal·lització per evitar la sobrenucleació, que era molt habitual. Els cristalls finals optimitzats tenien una mida superior als de proteïna sola.

4.4.1 Cristalls amb àtoms pesants

A part dels múltiples *soakings* amb àtoms pesants de cristalls ja formats, que es descriuran més endavant, per a obtenir un derivat útil pel fasejat experimental, es va cocrystal·litzar Aq_DnaB[1-439] amb nucleòtids halogenats, que incorporen un àtom de brom o iode, seguint les recomanacions descrites anteriorment per a la formació de cocrystals amb nucleòtids. Es provaren 4 nucleòtids halogenats diferents i, per bé que s'obtingueren cristalls per a tots ells, els generats amb iode eren molt petits i no s'aconseguí augmentar la seva mida per *seeding*, dilució o emprant olis.

Per a la proteïna derivatitzada amb seleno-metionina es realitzà un *screening* directament al voltant de les condicions de cristal·lització de la proteïna nativa, amb el

qual s'obtingueren ja els primers cristalls. La condició de cristal·lització final tenia un percentatge d'agent precipitant lleugerament inferior a les de la proteïna nativa.

4.5 Recollida i processament de dades difracció

4.5.1 Qualitat de les dades de difracció de les diferents formes cristal·lines

Els cristalls en forma de barra presentaven un poder de difracció dels raigs X molt feble, amb una resolució al voltant dels 6 Å, que no millorava al aconseguir barres més compactes. Els diferents protocols de crioprotecció testats no van millorar la qualitat de les dades de difracció.

Els cristalls en forma de placa, procedents de condicions de cristal·lització amb valors de pH entre 6.5 i 7, generaven patrons de difracció de major resolució, entre 4 i 5 Å, però les dades eren molt mosaïques. El millor crioprotector per aquest tipus de cristall era l'etilenglicol afegit gradualment.

Només els cristalls formats a valors de pH al voltant de 7.6 generaven patrons de difracció de qualitat i resolució acceptable, tot i que sovint calia provar molts cristalls abans de trobar-ne un que difractes al voltant dels 4 Å. Un altre problema de les dades recollides a partir d'aquests cristalls era que presentaven una severa anisotropia, difractant fins a 4 Å en la direcció millor, però només fins a uns 6-7 Å en l'altre. Es va decidir treballar només amb aquesta forma cristal·lina, doncs era la que donava millors resultats i intentar optimitzar-la. Es varen provar diferents protocols d'optimització, que han estat descrits a la secció de materials i mètodes.

Tot i que alguns cristalls d'Aq_DnaB[1-438] crescuts a pH 7.6 i sotmesos a deshidratació van presentar millor resolució, al voltant de 3.7 Å, no es tractava d'un tractament reproducible i s'abandonà, doncs també es malmetien nombrosos cristalls intentant deshidratar-los.

De totes les tècniques de millora provades, la que donà resultats més espectaculars va ser l'addició de nucleòtids durant l'etapa de cristallització. En concret, quan s'afegia ADP per cocrystal·lització, s'obtenien cristalls que podien assolir major mida, generaven patrons de difracció de resolució mínima a 4 Å i que no eren tant anisotròpics. A més, al tractar-se de cristalls més grans, no patien tant severament els danys per radiació i es podien recollir conjunts de dades més complets. Posteriors tractaments de millora d'aquests cristalls no donaren resultat.

4.5.1 Millor conjunt de dades natiu

El millor conjunt de dades per a un cristall natiu es recollí a la línia ID23eh1 del sincrotró de l'ESRF de Grenoble a partir d'un cocrystal amb 0.5 mM ADP crioprotegit amb etilenglicol al 24 % (p/v) afegit seqüencialment mitjançant equilibrat per pressió de vapor en canvis de 6 hores. Els paràmetres de recollida es mostren a la figura R.20, juntament amb una imatge de difracció i una fotografia del cristall presa durant la recollida de dades.

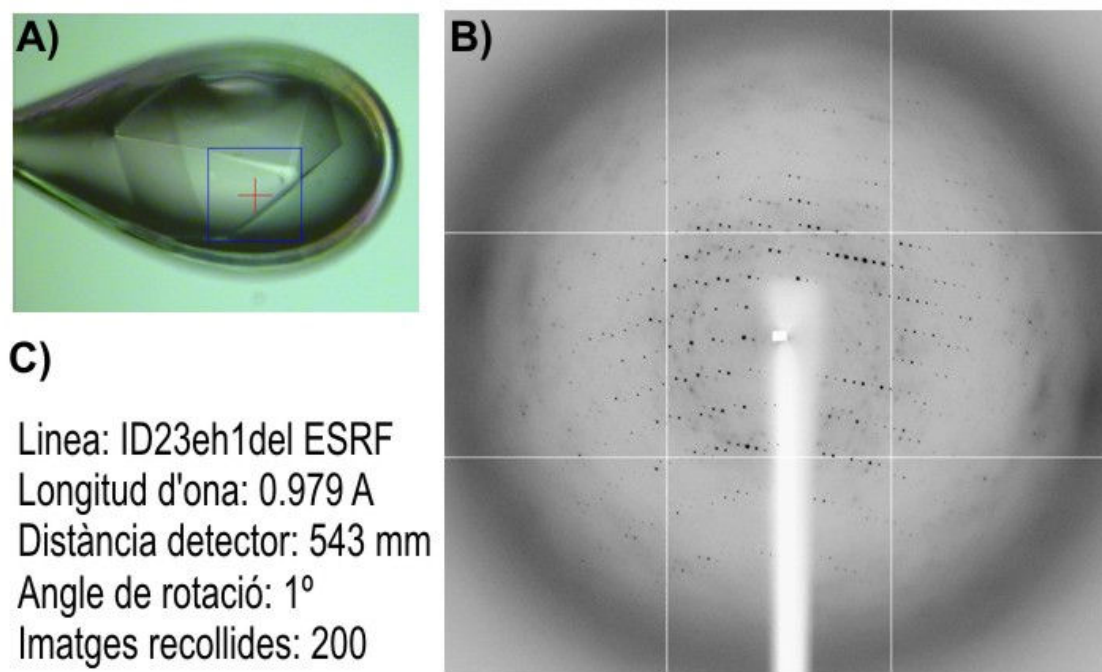


Figura R.20: Recollida de dades de difracció d'Aq_DnaB[1-439] cocrystal·litzada amb ADP. A) Cocrystal d'Aq_DnaB[1-439] amb ADP durant la recollida de dades de difracció. **B)** Imatge de difracció de raigs X generada per aquest cristall. **C)** Paràmetres de recollida.

Les dades varen ser indexades amb el programa XDS i escalades amb XSCALE. Les estadístiques de processament es mostren a la taula R.2. El cristall pertany al grup espacial C2, amb uns valors de cel·la unitat afinada de: $a = 282.51 \text{ \AA}$, $b = 103.31 \text{ \AA}$, $c = 125.25 \text{ \AA}$ i $\beta = 102.43^\circ$.

SUBSET OF INTENSITY DATA WITH SIGNAL/NOISE ≥ 1.0 AS FUNCTION OF RESOLUTION

| RESOLUTION LIMIT | NUMBER OF REFLECTIONS | | | COMPLETENESS OF DATA | R-FACTOR | | R-FACTOR COMPARED | I/SIGMA | R-meas | Rmrgd-F |
|---------------------|-----------------------|--------------|--------------|-------------------------|-------------|--------------|-------------------|-------------|--------------|-------------|
| | OBSERVED | UNIQUE | POSSIBLE | | observed | expected | | | | |
| 10.00 | 10153 | 1615 | 2007 | 80.5% | 5.6% | 7.4% | 10153 | 23.94 | 6.1% | 2.8% |
| 6.00 | 50388 | 6980 | 7006 | 99.6% | 6.2% | 7.8% | 50388 | 21.13 | 6.7% | 3.4% |
| 5.00 | 47972 | 6373 | 6463 | 98.6% | 11.9% | 12.1% | 47972 | 13.63 | 12.8% | 7.0% |
| 4.80 | 15003 | 1980 | 2008 | 98.6% | 13.3% | 13.7% | 15003 | 12.03 | 14.3% | 8.0% |
| 4.60 | 17593 | 2316 | 2367 | 97.8% | 14.8% | 15.1% | 17593 | 10.81 | 15.8% | 9.0% |
| 4.40 | 20597 | 2708 | 2787 | 97.2% | 18.9% | 19.2% | 20597 | 8.96 | 20.3% | 11.2% |
| 4.20 | 24132 | 3160 | 3338 | 94.7% | 27.5% | 28.1% | 24132 | 6.71 | 29.5% | 15.5% |
| 4.00 | 28931 | 3792 | 4070 | 93.2% | 40.9% | 40.7% | 28931 | 4.91 | 43.9% | 22.2% |
| 3.80 | 34933 | 4564 | 4932 | 92.5% | 55.4% | 54.4% | 34933 | 3.78 | 59.5% | 23.1% |
| 3.60 | 41940 | 5482 | 6095 | 89.9% | 69.4% | 70.0% | 41940 | 2.95 | 64.4% | 28.9% |
| 3.40 | 37868 | 5377 | 7619 | 80.6% | 86.7% | 87.9% | 37868 | 2.25 | 83.6% | 36.6% |
| total | 329510 | 44347 | 48692 | 95.1% | 9.7% | 11.2% | 329510 | 9.73 | 10.5% | 9.9% |

Taula R.2: Estadístiques de processament de les dades de difracció d'Aq_DnaB[1-439] cocrystal·litzada amb ADP. Dades processades amb XDS i XSCALE fins a una resolució de 3.4 Å en el grup espacial C2.

El càlcul del paràmetre de Matthews (Matthews 1968) donà un valor de $V_m = 2.8 \text{ \AA}^3 \text{ Da}^{-1}$ per a 6 molècules per unitat asimètrica, suggerint la presència d'un hexàmer de DnaB i un percentatge de contingut de solvent del 65,2 %.

La funció d'autorotació (*self-rotation*) d'aquest conjunt de dades, que es mostra a la figura R.21, suggereix que la partícula presenta simetria tres no cristal·logràfica, és a dir, l'hexàmer de DnaB de la cel·la unitat es trobaria en simetria rotacional C3.

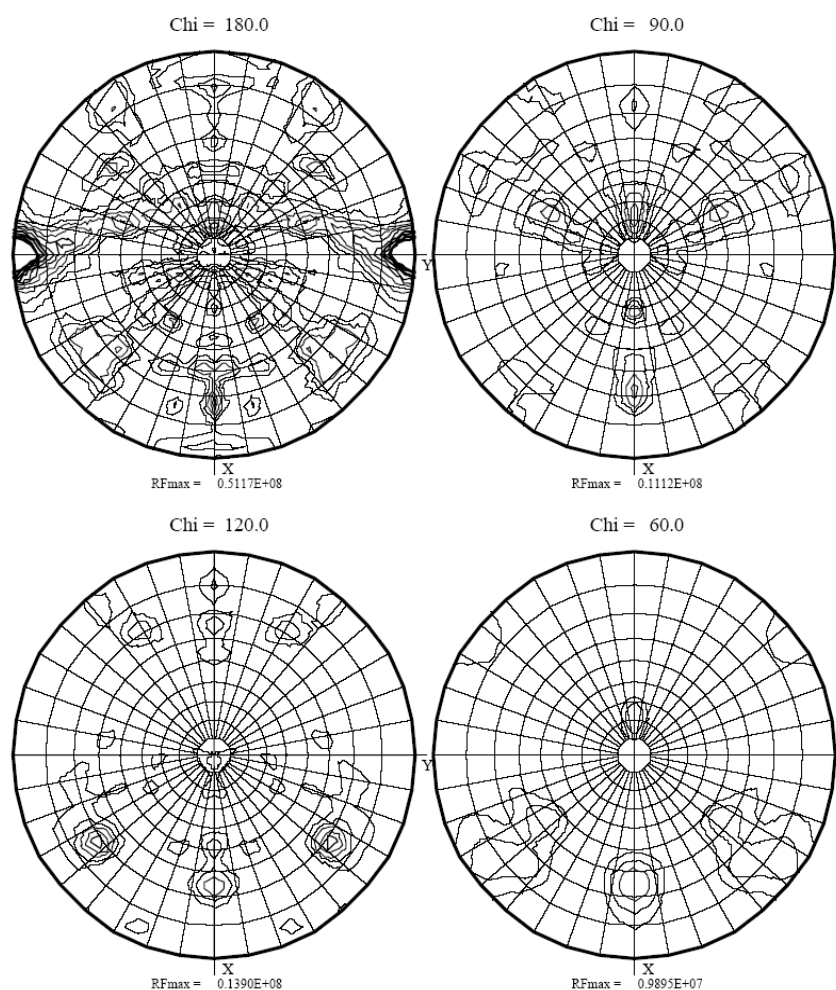


Figura R.21: Funció d'autorotació del millor conjunt de dades natiu. Figura generada amb el programa MolRep amb els següents paràmetres: resolució màxima : 3.00 Å; radi: 30.00 Å. Màxim de la funció de rotació (theta,phi,chi) : 0.5117E+08 i rms : 0.4386E+07.

4.5.2 Cocristalls amb ssDNA

Els cocristalls d'Aq_DnaB[1-439] amb el DNA de cadena senzilla de 21 pb presenten una resolució i qualitat de difracció similars als de proteïna sola.

A la taula R.3 es mostren els paràmetres de recollida i processament del conjunt de dades de difracció del millor cocristall amb ssDNA recollit, el qual pertany també al grup espacial C2 i té els següents paràmetres de cel·la unitat afinada: $a = 280.12 \text{ \AA}$, $b = 103.21 \text{ \AA}$, $c = 124.70 \text{ \AA}$ i $\beta = 102.38^\circ$.

SUBSET OF INTENSITY DATA WITH SIGNAL/NOISE \geq 1.0 AS FUNCTION OF RESOLUTION

| RESOLUTION LIMIT | NUMBER OF REFLECTIONS | | | COMPLETENESS OF DATA | R-FACTOR | | R-FACTOR COMPARED expected | I/SIGMA | R-meas | Rmrgd-F |
|---------------------|-----------------------|--------------|--------------|-------------------------|-------------|-------------|-------------------------------|-------------|-------------|-------------|
| | OBSERVED | UNIQUE | POSSIBLE | | observed | | | | | |
| 10.00 | 10153 | 1615 | 2007 | 80.5% | 5.6% | 7.4% | 10153 | 23.94 | 6.1% | 2.8% |
| 6.00 | 50388 | 6980 | 7006 | 99.6% | 6.2% | 7.8% | 50388 | 21.13 | 6.7% | 3.4% |
| 5.00 | 47972 | 6373 | 6463 | 98.6% | 11.9% | 12.1% | 47972 | 13.63 | 12.8% | 7.0% |
| 4.80 | 15003 | 1980 | 2008 | 98.6% | 13.3% | 13.7% | 15003 | 12.03 | 14.3% | 8.0% |
| 4.60 | 17593 | 2316 | 2367 | 97.8% | 14.8% | 15.1% | 17593 | 10.81 | 15.8% | 9.0% |
| 4.40 | 20597 | 2708 | 2787 | 97.2% | 18.9% | 19.2% | 20597 | 8.96 | 20.3% | 11.2% |
| 4.20 | 24132 | 3160 | 3338 | 94.7% | 27.5% | 28.1% | 24132 | 6.71 | 29.5% | 15.5% |
| 4.00 | 28931 | 3792 | 4070 | 93.2% | 40.9% | 40.7% | 28931 | 4.91 | 43.9% | 22.2% |
| 3.80 | 34933 | 4564 | 4932 | 92.5% | 55.4% | 54.4% | 34933 | 3.78 | 59.5% | 23.1% |
| 3.60 | 41940 | 5482 | 6095 | 90.9% | 69.4% | 70.0% | 41940 | 2.95 | 64.4% | 28.9% |
| 3.40 | 37868 | 5377 | 7619 | 83.6% | 86.7% | 77.9% | 37868 | 2.25 | 83.6% | 36.6% |
| total | 166589 | 39986 | 47939 | 83.4% | 8.0% | 8.5% | 166589 | 9.20 | 9.2% | 9.9% |

Taula R.3: Estadístiques de processament de les dades de difracció d'Aq_DnaB[1-439] cocrystal·litzada amb ssDNA. Dades processades fins a una resolució de 3.4 Å en el grup espacial C2, recollides a la línia ID23eh1 del ESRF a una longitud d'ona de 0.8731 Å, distància de detector de 552 mm, angle de rotació de 1^o i 360^o recollits.

4.5.3 Cristalls derivats amb àtoms pesant

Es van provar múltiples compostos d'àtoms pesants per intentar obtenir cristalls derivats que difractessin i fossin útils per resoldre l'estructura bé per reemplaçament isomòrfic o diferències anòmals, o combinacions d'ambdues tècniques. La taula R.4 resumeix el resultat dels diferents compostos provats, en quan a si produïren o no cristalls que generessin patrons de difracció acceptables.

Es provaren diversos compostos de lantànids, doncs aquest solen substituir als ions divalents en proteïnes que uneixen Mg²⁺ o Ca²⁺ i, per tant, hi havia possibilitats que s'incorporessin ordenadament a la xarxa cristal·lina. Malauradament, totes les solucions de lantànids provades (fins i tot a concentracions i temps molt baixos) provocaren que els cristalls s'esquerdessin de manera visible gairebé immediatament.

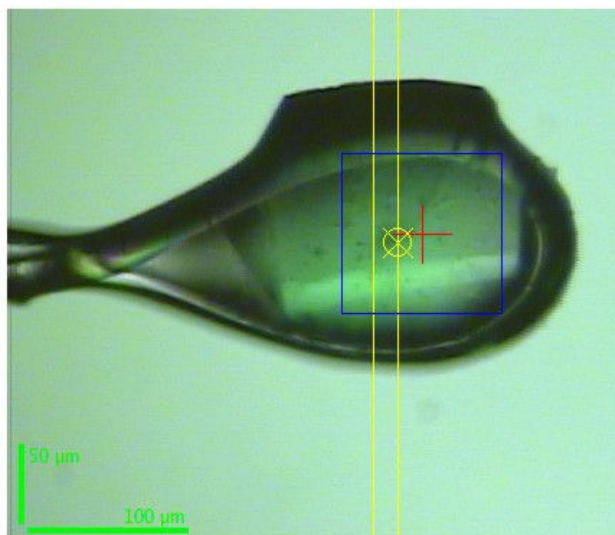
Els cristalls submergits en solucions de timerosal (Hg) no patien deformacions visuals ni una disminució important del seu poder de difracció de raigs X. Ara bé, tot i que donaven un pic de fluorescència molt clar al cantell d'absorció del Hg, segurament aquest no s'incorporava ordenadament a l'estructura cristal·lina, doncs no es pogueren localitzar els àtoms pesants.

| Compost | Tipus | Resultat |
|-------------------------------------|----------|---|
| Acetat monohidrat de samari (III) | Lantànid | Cristalls s'esquerden a totes les conc. |
| Nitrat pentahidrat d'ytterbi | Lantànid | Cristalls s'esquerden a alta conc. |
| Clorur d'hexahidrat de terbi (III) | Lantànid | Cristalls s'esquerden a totes conc. |
| Clorur monohidrat de gadolini (III) | Lantànid | Resolució < 8 Å |
| Trietil acetat de plom | Pb | Resolució < 6 Å |
| Nitrat de plom (II) | Pb | Resolució ≈ 5 Å però no completa |
| Hexacloroosmat (IV) de potassi | Os | Resolució < 6 Å |
| Clorur d'or (III) | Au | Precipita en les condicions de cristal·lització |
| Tetracloroaurat de sodi | Au | No s'incorpora |
| Tetracloroplatinat de potassi | Pt | Resolució < 6 Å |
| Clorur d'etil mercuri | Hg | Cristalls s'esquerden a totes les conc. |
| Timerosal | Hg | Resolució ≈ 4.5 Å però no completa |
| Tetracloroplatinat (II) d'amoní | Pt | Resolució < 7 Å |
| Tetracionoplatinat (II) de potasi | Pt | Resolució < 6 Å |
| Clorur d'urani | U | Resolució ≈ 5.5 Å però no completa |
| Hexabromur de tàntal | Clúster | Alguns cristalls a resolució ≈ Å |
| Compost de wolfrani | Clúster | Resolució < 7 Å |

Taula R.4: Compostos d'àtoms pesants provats amb els cristalls d'Aq_DnaB[1-439].

Pel que fa al clúster d'hexabromur de tàntal, quan es realitzaven *soakings* curts a concentracions altes, s'obtenien cristalls que presentaven una forta senyal de fluorescència al cantell d'absorció del tàntal, però que tenien una qualitat de difracció molt minsa, poca resolució (al voltant dels 8 Å) i patien un dany per radiació molt sever. Els tractaments curts a concentracions baixes o intermèdies de la solució de tàntal no permetien la incorporació de l'àtom pesant. Els *soakings* llargs a concentracions mitjanes generaven cristalls que difractaven els raigs X amb una resolució al voltant dels 6 Å, però que continuaven patint molt dany per radiació. Això impossibilitava la recollida de conjunts de dades complets a partir d'aquests cristalls, doncs, degut a la poca simetria del grup espacial C₂, es necessiten un mínim de 180° recollits i aquests cristalls deixaven de generar dades de difracció als aproximadament 20° (unes 10 exposicions). Finalment, s'aconseguí un cristall sotmès a *soaking* en una solució molt diluïda de Ta₆ Br₁₄ durant 2 dies que difractà fins a una resolució de 4.6 Å. A la figura

R.22 es mostra una fotografia d'aquest cristall presa durant la recollida de dades de difracció, així com els paràmetres d'aquesta.



Paràmetres de recollida:

Linea: ID23eh1
Distància detector: 474 mm
Angle de rotació: 1.5°
Graus recollits: 210°
Longitud d'ona:
- pic: 1.2557 Å
- inflexió: 1.2561 Å

Figura R.22: Cristall derivatitzat amb bromur de tàntal. Es mostra una imatge del cristall muntat al difractòmetre i els paràmetres de recollida de les dades de difracció. Els valors de f' i f'' són, respectivament, de -18.5 i 25.1 per al pic, i de -29.5 i 14.5 per a la inflexió.

A la taula R.5 es mostren les estadístiques de processament del pic i la inflexió. No es va poder recollir una tercera longitud d'ona degut al fort dany per radiació que havia patit el cristall. De fet, el conjunt de dades corresponent a la inflexió, que es va recollir en segon lloc, ja té una qualitat global molt pobre i es va tallar a una resolució de 6 Å durant el processament.

Les dades pertanyen al grup espacial C2 i tenen els següents paràmetres de cel·la unitat afinada: $a = 280.68 \text{ \AA}$, $b = 102.98 \text{ \AA}$, $c = 124.24 \text{ \AA}$ i $\beta = 101.95^\circ$.

| Estadístiques de processament de les dades del cristall de tàntal recollides al pic d'absorció | | | | | | | | |
|--|-----------------------|--------------------|----------------------|---------------|-------------------|-------------|---------------|----------------|
| Resolució (Å) | Reflexions observades | Reflexions úniques | Reflexions possibles | Completesa | R-factor observat | I/sigma | Rmerge | Senyal anòmala |
| 10 | 8872 | 3173 | 3672 | 86.4% | 3.7 % | 21.21 | 3.3 % | 85 % |
| 8 | 9988 | 3500 | 3538 | 98.9 % | 4.5 % | 18.14 | 4.3 % | 63 % |
| 7 | 9907 | 3428 | 3545 | 96.7 % | 7.1 % | 12.60 | 8.0 % | 44 % |
| 6.5 | 7023 | 2390 | 2635 | 90.7 % | 12.8 % | 8.55 | 14.8 % | 25 % |
| 6 | 9156 | 3119 | 3621 | 86.1 % | 18.2 % | 6.28 | 21.7 % | 12 % |
| 5.8 | 4409 | 1507 | 1862 | 80.9 % | 23.4 % | 5.45 | 26.0 % | 4 % |
| 5.5 | 7406 | 2506 | 3268 | 76.7 % | 27.2 % | 4.62 | 29.7 % | 3 % |
| 5.3 | 5494 | 1862 | 2590 | 71.9 % | 33.4 % | 3.85 | 35.2 % | 3 % |
| 5 | 9687 | 3269 | 4728 | 69.1 % | 32.7 % | 3.79 | 34.7 % | 6 % |
| 4.8 | 7604 | 2597 | 3857 | 67.3 % | 36.1 % | 3.63 | 39.2 % | 2 % |
| 4.6 | 8427 | 2870 | 4524 | 63.4 % | 37.7 % | 3.40 | 39.6 % | 5 % |
| TOTAL | 87973 | 30221 | 37840 | 79.9 % | 6.4 % | 9.02 | 12.8 % | 44 % |
| Estadístiques de processament de les dades del cristall de tàntal recollides a la inflexió | | | | | | | | |
| Resolució (Å) | Reflexions observades | Reflexions úniques | Reflexions possibles | Completesa | R-factor observat | I/sigma | Rmerge | Senyal anòmala |
| 10 | 6732 | 3473 | 3672 | 93.5% | 5.1 % | 12.49 | 5.7 % | 44 % |
| 8 | 6566 | 3300 | 3538 | 96.4 % | 7.5 % | 9.14 | 10.3 % | 1 % |
| 7 | 6149 | 3028 | 3545 | 87.7 % | 18.1 % | 5.60 | 26.0 % | 0 % |
| 6.5 | 3817 | 1903 | 2635 | 72.7 % | 32.8 % | 3.22 | 42.3 % | 1 % |
| 6 | 4417 | 2216 | 3621 | 64.2 % | 44.2 % | 2.28 | 54.5 % | 4 % |
| TOTAL | 27681 | 13966 | 17011 | 82.3 % | 9.6 % | 3.72 | 15.6 % | 30 % |

Taula R.5: Estadístiques de processament de les dades recollides al pic d'absorció del cantell del tàntal per al cocristalls amb ADP derivatitzat amb Ta₆ Br₁₄. Dades processades fins a una resolució de 4.6 Å en el grup espacial C2. També es mostren les estadístiques de processament de les dades recollides a la inflexió, tot i que la qualitat d'aquestes és molt baixa.

4.6 Resolució de l'estructura tridimensional

4.6.1 Mètodes no experimentals: reemplaçament molecular

4.6.1.1 Amb models atòmics

Tal i com s'ha descrit a l'apartat de mètodes de cristal·lografia de raigs X, la resolució del les fases mitjançant la tècnica del reemplaçament molecular consisteix en

inferir un valor inicial per a les fases de l'estructura d'interès a partir d'una estructura similar coneguda, que és correctament orientada i posicionada a la cel·la unitat de la nova estructura. Conseqüentment, per a utilitzar el reemplaçament molecular cal disposar d'un model, com més similar a l'estructura que es vol resoldre, millor.

Quan es van obtenir els primers conjunts de dades nadius per a Aq_DnaB[1-439], només es disposava d'un model atòmic parcial corresponent al domini N-terminal de la proteïna DnaB d'*E. coli*. Aquest havia estat resolt per cristal·lografia de raigs X i ressonància magnètica nuclear en forma de dímer (Fass *et al.* 1999; Weigelt *et al.* 1999). El domini N-terminal de DnaB representa només un 27% de l'estructura total de la proteïna. La identitat de seqüència entre els dominis N-terminals de la DnaB d'*E. coli* i d'*A. aeolicus* és del 29 %, doncs aquesta és la zona menys conservada entre les diferents DnaB bacterianes.

Per a la regió C-terminal de la proteïna, el model relacionat més proper del qual es disposava era l'estructura cristal·lina de la regió helicasa de la proteïna del gen 4 (gp4) del fag de T7, membre de la mateixa família d'helicases (SF3) que DnaB. Es va prendre com a model l'estructura de gp4 que havia cristal·litzat en forma d'anell hexamèric (figura I.6 de la introducció de (Singleton *et al.* 2000)). La identitat de seqüència d'ambdues proteïnes és només del 17%, valor que puja al 24% si s'alinea només la regió C-terminal d'Aq_DnaB.

Es varen utilitzar ambdós models, junts i per separat, per intentar resoldre l'estructura mitjançant la tècnica del reemplaçament molecular. Els models es van millorar amb el programa Chainsaw, que elimina les possibles insercions del model i muta els residus no conservats a alanina. També es provà retallant les zones corresponents a llaços exposats a solvent d'ambdues estructures. No s'obtingué cap solució en cap de les múltiples cerques, emprant els programes Phaser, MolRep i AMoRe.

A la part final d'aquest treball, va publicar-se l'estructura cristal·lina del monòmer de DnaB de *Thermus aquaticus* (Bailey *et al.* 2007a) i, al cap d'uns mesos, de l'hexàmer de DnaB de *Bacillus stearothermophilus* (Bailey *et al.* 2007b) i de la proteïna

GP40, homòloga a DnaB del bacteriòfag SPP1 de *Bacillus sp* (Wang *et al.* 2008). Les identitats de seqüència d'aquestes proteïnes entre elles és de 43.7 %, i amb Aq_DnaB [1-439] de 37 % per la DnaB de *T. aquaticus* i de 36.1 % per a la G40P.

Es varem emprar múltiples models d'aquestes estructures, senceres, retallades i combinades sense èxit. Tots els models s'havien millorats amb el programa Chainsaw, eliminant insercions i mutant a alanina els residus no conservats. A continuació es descriuen els diferents models provats:

- monòmer sencer de *T. aquaticus* i dominis per separat. Domini C-terminal retallant les zones de llaços, utilitzant només la part central més globular.
- hexàmer sencer de *B. stearothermophilus* i hexàmer només de C-terminals i només de N-terminals. Els mateixos models però retallant els llaços i les zones menys conservades respecte a *A. aeolicus*. Només monòmer (emprant els dos monòmers diferents que hi ha a l'estructura) i també els seus dominis per separat. El dímer de la unitat asimètrica, sencer i per dominis.
- igual que amb el model anterior amb l'hexàmer de G40P, però en aquest cas, com hi ha un hexàmer sencer per unitat asimètrica, provant amb cadascun dels monòmers per separat.
- combinacions amb l'hexàmer C-terminal de *Bacillus* i l'hexàmer N-terminal de SPP1, i al revés.

En cap cas s'obtingué una solució de reemplaçament molecular satisfactòria, tot i emprar varis programes, Phaser i MolRep, i les opcions més innovadores dintre d'aquests. Per exemple, a Phaser es va ampliar el nombre de rotacions i, particularment, translacions que s'evaluen per evitar que es perdi alguna solució correcta durant el càlcul. A MolRep, es van fer cerques forçant que les solucions tinguessin una simetria concreta respecte al model que entrem, per exemple, si busquem emprant monòmers, els que vagi col·locant el programa haurien d'estar disposats formant un hexàmer.

Igualment, es van realitzar les cerques a varies resolucions i en tots els grups espacials possibles.

4.6.1.2 Amb models de microscòpia electrònica

Com inicialment no teníem models atòmics per a la proteïna sencera, es va decidir provar la utilització de models procedents de reconstruccions de criomicroscòpia electrònica d'Aq_DnaB[1-439] com a models de cerca per al reemplaçament molecular. A part d'aquests models, per a dur a terme aquesta estratègia de fasejat també necessitàvem recollir noves dades de difracció a molt baixa resolució, per tal que es solapessin bé amb la resolució obtinguda a les reconstruccions de criomicroscòpia.

4.6.1.2.1 Reconstruccions de criomicroscòpia electrònica

Per obtenir els models de microscòpia electrònica es va col·laborar amb el laboratori del Professor José María Carazo, el qual ha treballat extensament en la caracterització estructural mitjançant tècniques de microscòpia electrònica de diverses helicases hexamèriques i entre aquestes especialment amb la proteïna DnaB d'*E. coli* (San Martín *et al.* 1995; San Martín *et al.* 1998; Donate *et al.* 2000; Valle *et al.* 2000; Barcena *et al.* 2001; Gomez-Llorente *et al.* 2005; Nunez-Ramirez *et al.* 2006; Nunez-Ramirez *et al.* 2007).

Les imatges de tinció negativa així com les reconstruccions de criomicroscòpia electrònica que s'empraren en aquest treball les realitzà el Dr. Rafael Nuñez-Ramírez a partir de la mostra d'Aq_DnaB[1-439] purificada i en el mateix tampó que s'emprava a les proves de cristal·lització (20 mM Tris pH 7.6, 100 mM NaCl i 5 mM MgCl₂)

A la figura R.23 es mostren les imatges de microscòpia electrònica de l'anàlisi preliminar per tinció negativa de la mostra, les quals indicaven clarament que l'hexàmer d'Aq_DnaB[1-439] es troba en simetria tres i no presenta heterogeneïtat quaternària sota les condicions estudiades. S'observen sis masses clares que probablement corresponen als sis monòmers que conformen l'hexàmer. Igualment, es poden

diferenciar dos tipus de masses que difereixen en la seva mida i podrien correspondre als dos dominis de DnaB.

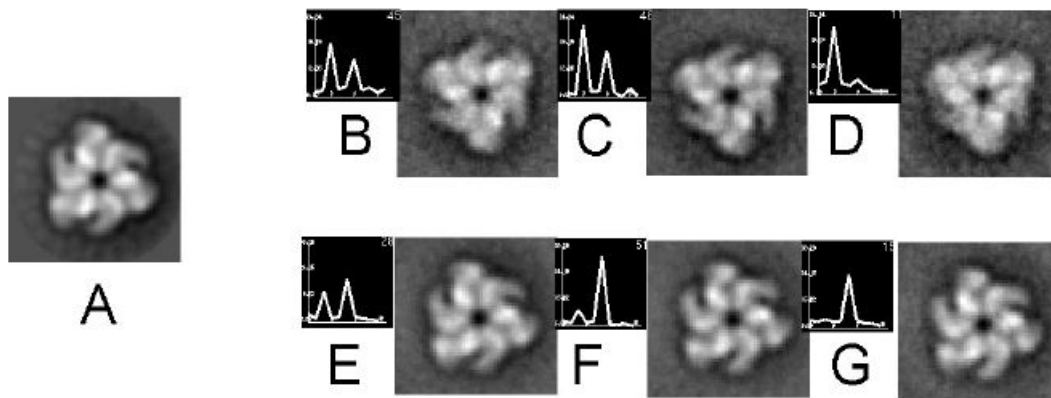


Figura R.23: Projeccions bidimensionals d'imatges per microscòpia electrònica per tinció negativa d'Aq_DnaB[1-439]. A) Imatge promig de totes les poblacions presents a la mostra analitzada per microscòpia electrònica. Dades obtingudes amb l'alineament d'unes 24000 partícules individuals. B, C, D, E, F, G) Representats de cadascuna de les sis poblacions diferenciades segons el seu espectre rotacional. Encara que aquestes poblacions presenten valors d'espectre rotacional diferents, la seva imatge promig es molt similar a la de la resta. Les variacions es centren en la longitud de la massa de menor mida. Imatges i processament realitzats pel Dr. Rafael Nuñez-Ramírez.

A la figura R.24 es mostren les imatges de criomicroscòpia electrònica d'Aq_DnaB[1-439], les qual presenten una relació senyal/soroll baixa, fet que en dificulta el seu anàlisi.

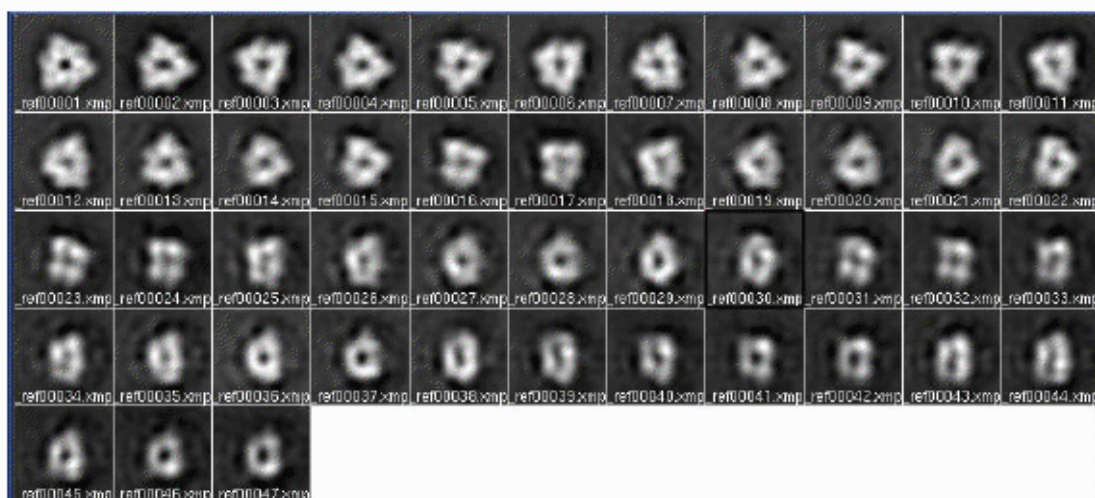


Figura R.24: Imatges de criomicroscòpia electrònica d'Aq_DnaB[1-439]. Mostra d'algunes de les imatges emprades per generar les reconstruccions tridimensionals. Corresponen a l'aspecte que té l'hexàmer des de diferents perspectives. La primera seria la vista frontal i a

mesura que s'avança d'esquerra a dreta i de dalt cap a baix ens acostem a les vistes laterals. La vista frontal s'assembla molt a la imatge de tinció negativa. Són imatges en dos pisos, algunes amb quatre masses predominants i la majoria amb una zona interna de menor densitat. Imatges recollides i processades pel Dr. Rafael Nunez-Ramírez.

Partint de diferents models inicials de referència, es va tractar de generar una reconstrucció tridimensional que encaixes amb les vistes bidimensionals, mitjançant un procés d'afinament angular. Malauradament, no es va aconseguir que les dades de criomicroscòpia convergissin en una sola reconstrucció tridimensional. Podria ser bé perquè el contrast de les imatge vers el soroll no fos prou bo i/o perquè hi havia certa heterogeneïtat a la mostra, segurament deguda a la flexibilitat intrínseca de la proteïna.

Així doncs, es van obtenir múltiples models, a una resolució mitjana de 25 Å, que es poden separar bàsicament en dos grans grups: els que presenten una simetria global C3 i els de simetria global C6. A les figures R.25 i R.26 es mostren tots aquests models de criomicroscòpia d'Aq_DnaB[1-439], els dos amb simetria C3 i els tres amb simetria C6. Es va decidir provar-los tots com a models de cerca per al reemplaçament molecular, començant pels models amb simetria C3, que semblava que serien els més semblants a l'estructura de l'hexàmer a l'interior del cristall, doncs la funció d'autorotació indicava que aquest tenia simetria 3.

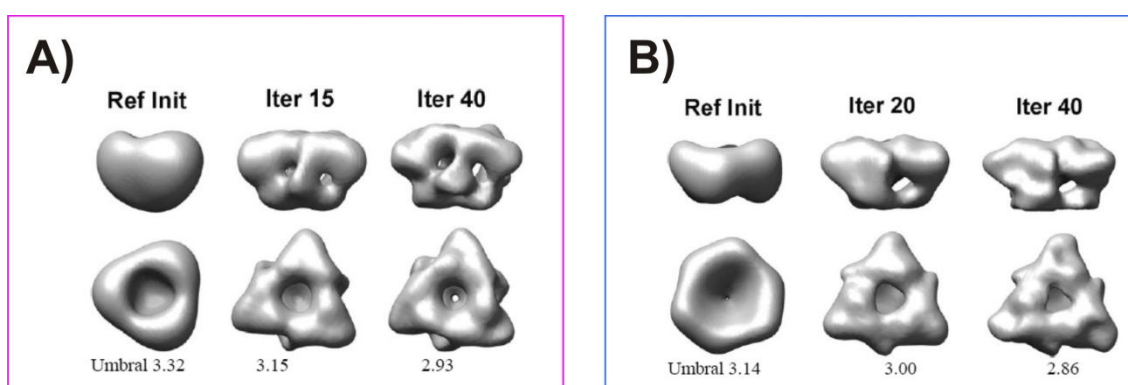


Figura R.25: Reconstruccions tridimensionals amb simetria global C3 generades per criomicroscòpia electrònica d'Aq_DnaB[1-439]. A cadascun es representa: a la dreta el model inicial, al centre el model obtingut en iteracions intermèdies i a l'esquerra el model final al qual convergeix el procés d'afinament angular. S'indica el llindar emprat per generar les figures. Reconstruccions realitzades pel Dr. Rafael Nuñez-Ramírez.

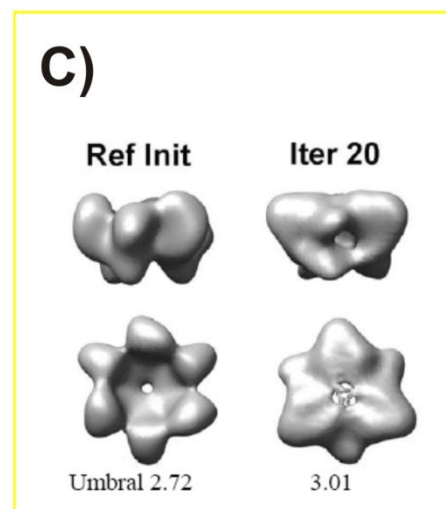
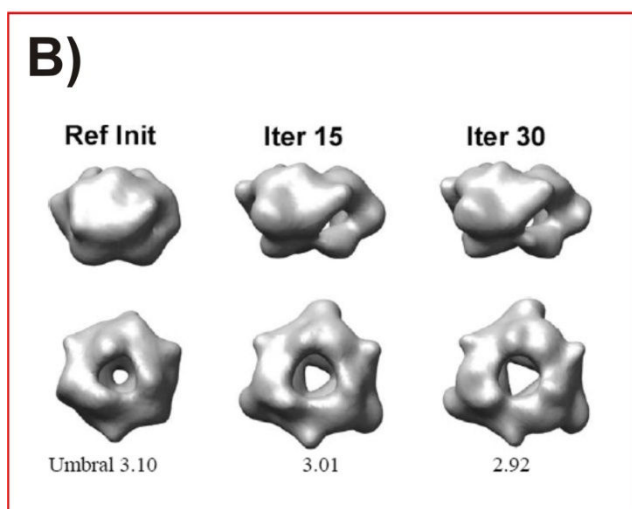
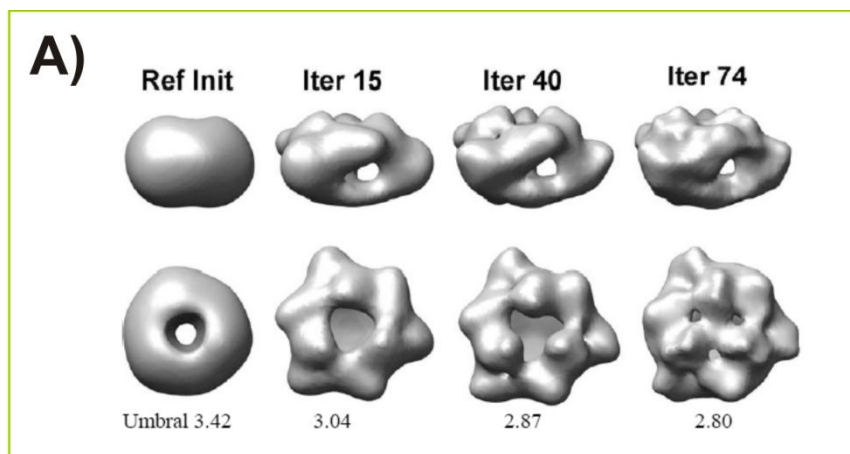


Figura R.26: Reconstruccions tridimensionals amb simetria global C6 generades per criomicroscòpia electrònica d'Aq_DnaB[1-439]. A cadascun es representa: a la dreta el model inicial, al centre el model obtingut en iteracions intermèdies i a l'esquerra el model final al qual convergeix la cerca. S'indica el llindar emprat per generar les figures, que és el corresponent a un volum capaç de contenir una massa de 320 kDa, la corresponent per a un hexàmer d'Aq_DnaB[1-439]. Reconstruccions realitzades pel Dr. Rafael Nuñez-Ramírez.

4.6.1.2.2 Dades de difracció de molt baixa resolució

Igualment, calia disposar de dades de difracció a molt baixa resolució, cosa que no es podia obtenir a les línies de sincrotró de l'ESRF, on la resolució mínima de les dades recollides era d'uns 20 Å i les capes de més baixa resolució tenien poca completesa. Per això, es va demanar temps de sincrotró a les instal·lacions del SLS, concretament a la línia XS60A, especialment dissenyada per a la difracció de cristalls amb cel·les grans i per a l'obtenció de dades de molt baixa resolució. A la figura R.27

es mostra una comparació dels límits de resolució inferiors dels conjunts de dades recollits a ID23eh1 i XS60A.

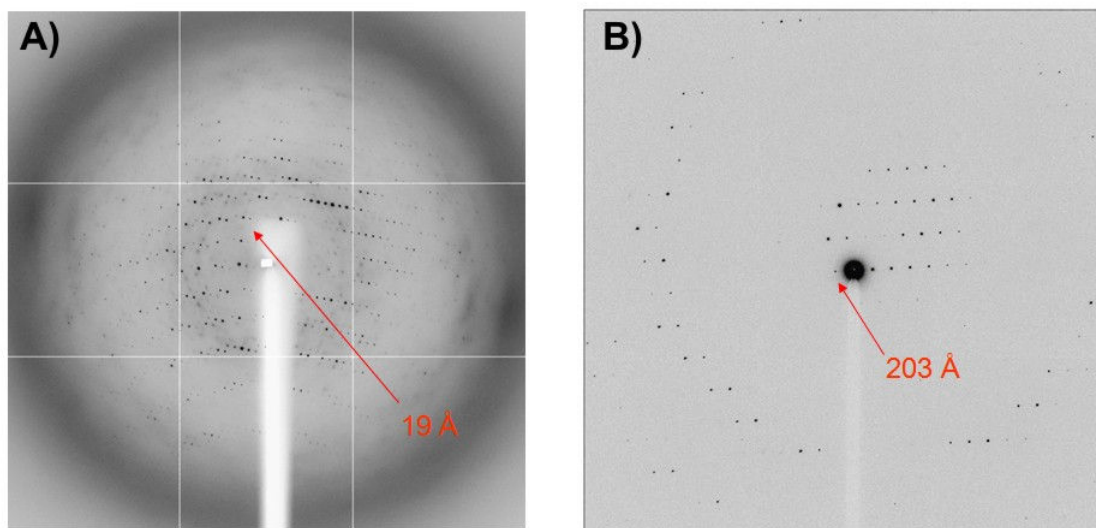


Figura R.27: Comparació de les resolucions mínimes obtingudes al ESRF i al SLS. A) Imatge de difracció obtinguda a la línia ID23eh1 del ESRF. **B)** Imatge obtinguda a la línia XS60A del SLS. S'indica en cada cas amb una fletxa taronja la resolució de les taques de més baixa resolució.

Es varen recollir diversos conjunts de dades de difracció d'un mateix cocristall d'Aq_DnaB[1-439] amb ADP, a diferents graus d'atenuació i distàncies de detector (que defineix la resolució màxima de les dades que es poden recollir) per tal d'assegurar-nos de no saturar els punts corresponents a les reflexions de més baixa resolució. Es varen indexar els diferents conjunts de dades amb el programa XDS i es varen combinar en el seu escalat amb XSCALE fins obtenir un grau de completesa màxim, de manera que al processat final només faltava una reflexió a la capa de resolució compresa entre 100 i 70 Å, mentre que la completesa de 100 a 20 Å era del 99.6%. La taula R.6 mostra els estadístics de processat d'aquestes dades, i la figura R.7 els paràmetres de recollida dels diferents conjunts de dades que es varen combinar per obtenir-lo.

Les dades pertanyien al grup espacial C2 i tenien uns paràmetres de cel·la unitat afinada de $a = 280.55 \text{ \AA}$, $b = 103.21 \text{ \AA}$, $c = 124.7 \text{ \AA}$ i $\beta = 102.384$.

SUBSET OF INTENSITY DATA WITH SIGNAL/NOISE \geq 1.0 AS FUNCTION OF RESOLUTION

| RESOLUTION | NUMBER OF REFLECTIONS | | COMPLETENESS | | R-FACTOR | R-FACTOR COMPARED | I/SIGMA | | R-meas | Rmrgd-F |
|--------------|-----------------------|--------------|--------------|--------------|--------------|-------------------|---------------|-------------|--------------|-------------|
| | LIMIT | OBSERVED | UNIQUE | POSSIBLE | | | OF DATA | observed | | |
| 100.00 | 10 | 3 | 3 | 100.0% | 4.3% | 8.5% | 10 | 13.80 | 5.0% | 1.7% |
| 70.00 | 33 | 4 | 5 | 80.0% | 7.4% | 8.6% | 33 | 22.67 | 8.3% | 3.2% |
| 40.00 | 339 | 31 | 31 | 100.0% | 9.4% | 10.0% | 339 | 24.50 | 10.0% | 3.3% |
| 20.00 | 3508 | 229 | 230 | 99.6% | 10.9% | 10.8% | 3508 | 26.43 | 11.3% | 3.0% |
| 10.00 | 32193 | 1715 | 1716 | 99.9% | 15.0% | 14.4% | 32193 | 25.03 | 15.5% | 3.1% |
| 9.00 | 12192 | 713 | 719 | 99.2% | 18.8% | 17.2% | 12192 | 22.06 | 19.4% | 3.7% |
| 8.00 | 15928 | 1121 | 1127 | 99.5% | 23.5% | 21.5% | 15928 | 17.09 | 24.4% | 5.1% |
| 7.00 | 19418 | 1825 | 1837 | 99.3% | 26.3% | 24.7% | 19418 | 12.69 | 27.7% | 7.0% |
| 6.00 | 33689 | 3212 | 3243 | 99.0% | 28.9% | 27.6% | 33689 | 9.73 | 30.4% | 8.8% |
| 5.00 | 43626 | 6244 | 6412 | 97.4% | 22.6% | 29.3% | 43625 | 6.81 | 24.5% | 11.9% |
| 4.50 | 37544 | 5364 | 5616 | 95.5% | 20.1% | 29.3% | 37543 | 6.18 | 21.8% | 13.3% |
| 4.00 | 54473 | 7731 | 8789 | 88.0% | 30.3% | 35.1% | 54472 | 4.59 | 32.8% | 20.8% |
| 3.90 | 12841 | 1860 | 2314 | 80.4% | 56.0% | 50.6% | 12840 | 3.29 | 60.7% | 34.2% |
| total | 265794 | 30053 | 32042 | 93.8% | 17.1% | 16.7% | 265790 | 8.53 | 17.8% | 8.8% |

Taula R.6: Estadístiques de processament de les dades de molt baixa resolució. Dades processades amb XDS i XSCALE fins a una resolució de 3.9 Å en el grup espacial C2.

| | Recollida 1 | Recollida 2 | Recollida 3 | Recollida 4 | Recollida 5 |
|---------------------------|-------------|-------------|-------------|-------------|-------------|
| Longitud d'ona | 0.9795 Å | 0.9795 Å | 0.9795 Å | 0.9795 Å | 0.9795 Å |
| Distància detector | 400 mm | 1000 mm | 1000 mm | 700 mm | 300 mm |
| Atenuació feix | No | 0.071 | 0.012 | 0.172 | No |
| Exposició | 1'' | 0.5'' | 0.5'' | 0.5'' | 1'' |
| Angle de rotació | 1° | 1° | 1° | 1° | 1° |
| Graus recollits | 180° | 180° | 200° | 180° | 200° |
| Resolució | 100 – 3.5 Å | 250 – 8 Å | 250 – 8 Å | 200 – 6 Å | 100 – 3 Å |
| Escalat conjunt | 70 – 3.9 Å | 250 – 7 Å | 250 – 10 Å | 70 – 6 Å | 50 – 3.9 Å |

Taula R.7: Paràmetres de les diferents recollides del conjunt de dades a molt baixa resolució. Es van recollir 5 conjunts de dades del mateix cristall, a cadascun dels quals es variaven lleugerament els paràmetres de recollida. L'última fila indica els rangs de resolució finals de cada recollida emprats per a l'escalat conjunt final, els quals es van escollir per maximitzar la completa de les diferents capes de resolució sense augmentar considerablement els valors de Rmerge.

Aquest conjunt de dades a molt baixa resolució va ser l'utilitzat per a les cerques amb els models de microscòpia electrònica a diferents resolucions de tall

(normalment 35, 20, 15 i 10 Å) i emprant els programes de reemplaçament molecular Phaser i AMoRe.

Phaser es va usar per introduir directament els mapes de criomicroscòpia d'Aq_DnaB[1-439] com a models de cerca. Malauradament, aquesta opció no funcionà satisfactòriament. Malgrat provar diversos l·lindars de densitat, semblava difícil ajustar bé les dimensions dels mapes de criomicroscòpia per tal de fer-los totalment compatibles amb el programa de reemplaçament molecular, optimitzat per treballar amb dades atòmiques.

L'estratègia que va donar millors resultats va ser generar models de pseudoàtoms a partir dels mapes de criomicroscòpia i emprar-los com a models de cerca atòmics als programes Phaser i AMoRe. Aquests models de pseudoàtoms es generaren gràcies al programa Map2Atom creat pel Dr. Joan Pous Ramos, el qual col·loca un pseudoàtom a cada punt del mapa de microscòpia que està pel cim d'una determinada densitat l·lindar. El l·lindar de densitat amb que es treballa és aquell que permet encaixar dins el mapa de criomicroscòpia la massa molecular corresponent a l'hexàmer de DnaB (uns 320 kDa) i s'ha d'anar afinant iterativament.

Amb aquests models de pseudoàtoms, concretament amb les procedents de la reconstrucció A iteracions 30 i 40 que s'han mostrat a la figura R.29, s'aconseguí tronar una solució a la funció de rotació que es discriminava amb certa claredat de la resta. Ara bé, les posteriors funcions de translació d'aquesta solució parcial no donaren cap solució satisfactòria, tot i realitzar cerques molt extenses.

Finalment, amb un model de pseudoàtoms que combinava tots sis mapes de criomicroscòpia en un de sol (veure figura R.28), s'aconseguí una solució de reemplaçament molecular comuna tant amb el programa Phaser com AMoRe, però que era impossible empaquetar dins la cel·la unitat del cristall sense que es produïssin xocs estèrics severos.

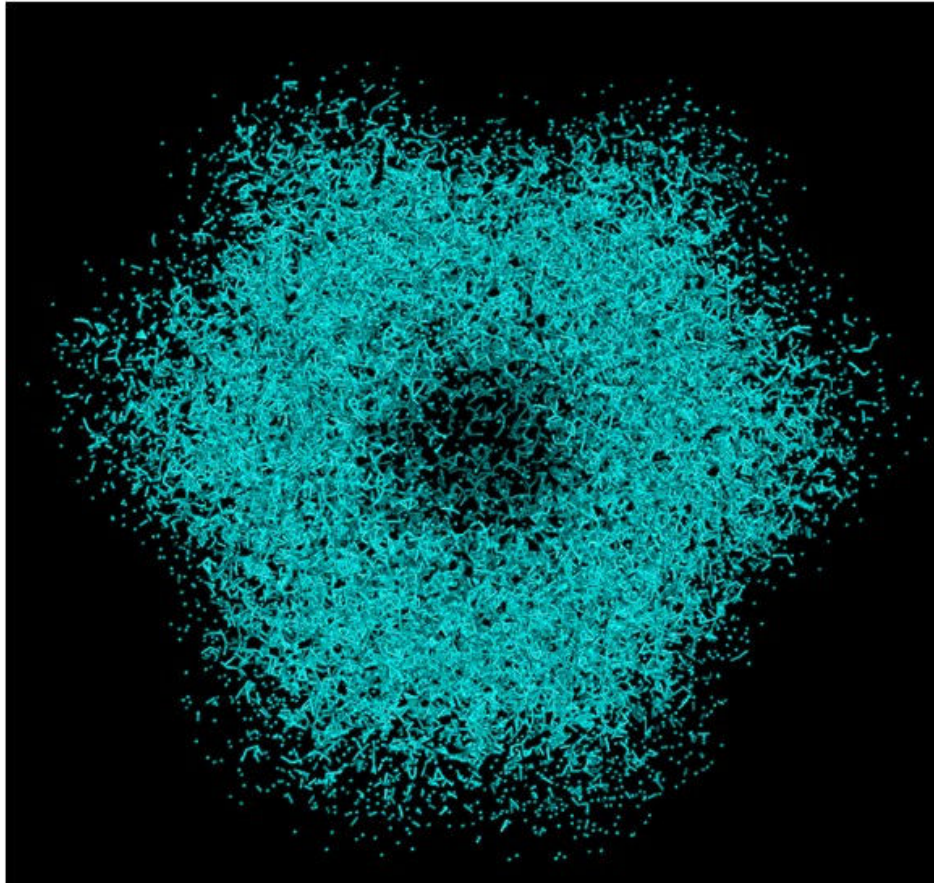


Figura R.28: Model de pseudoàtoms generat combinant els cinc models de criomicroscòpia d'Aq_DnaB[1-439]. La densitat de pseudoàtoms és major a les zones comunes entre les cinc reconstruccions tridimensionals de criomicroscòpia i menor a les zones que només es troben presents a un dels models. Aquest model ha estat generat amb el programa Map2Atom.

4.6.2 Tècniques experimentals

Aquesta secció es centrarà en els compostos d'àtoms pesants per als quals es generaren dades de difracció. El total de compostos provats ja ha estat descrit a la secció de cristal·lització.

Els cristalls de proteïna derivatitzada amb selenometionina tenien un límit de difracció similar als nadius, però eren molt més sensibles al dany per radiació, fet que impossibilita recollir conjunts de dades complets.

Els nucleòtids halogenats produïren cristalls molt grans, que generaven un pic de fluorescència clar al cantell d'absorció del brom o el iode, i que generaven patrons de

difracció que no patien tant severament els danys per radiació com els altres cristalls de derivats. Es varen obtenir varis conjunts de dades a una resolució al voltant de 3.9 Å, però que no tenien prou senyal anòmala per a emprar-los amb èxit en la resolució de l'estructura. Segurament, degut a l'elevada intensitat de feix de raigs X necessària per generar difracció a partir dels cristalls d'Aq_DnaB[1-439], els àtoms de Br i I units als nucleòtids es dissociaven d'aquests després d'haver recollit només unes poques imatges i, conseqüentment, es perdia la senyal anòmala procedent d'aquests àtoms.

Així doncs, dels múltiples conjunts de dades recollits amb cristalls derivatitzats amb diferents àtoms pesants, només foren útils per a la resolució de l'estructura d'Aq_DnaB[1-439] els de bromur de tàntal i, concretament, només un dels conjunts de dades recollits per a aquest compost.

4.6.2.1 Resolució de l'estructura per SIRAS

S'emprà el programa XPREP per analitzar les dades de difracció del cristall de tàntal i escalar-les conjuntament amb els diferents conjunts de dades nadius dels quals es disposava (cocristall amb ADP original, cocristall amb ADP i ssDNA i conjunt de dades a molt baixa resolució), així com per analitzar la resolució fins a la qual les diferències anòmales eren significatives i crear un arxiu d'entrada per al paquet SHELX (Sheldrick 2008), que seria posteriorment utilitzat per resoldre l'estructura.

Després d'analitzar les diferents parelles de conjunts de dades en quan a isomorfisme i qualitat, es decidí emprar el conjunt de dades recollit al pic d'absorció del tàntal per al cristall derivat i el conjunt de dades corresponent al cocristall amb ADP original com a dades de difracció natives. Igualment, es decidí utilitzar per a la cerca de les posicions dels àtoms pesants les dades només fins a una resolució de 6 Å, doncs per sota d'aquest límit el ràtio senyal/soroll disminueix dràsticament i, conseqüentment, les diferències ja no eren significatives. A la figura R.29 es mostra el mapa de Patterson de diferències anòmales generat amb aquestes dades.

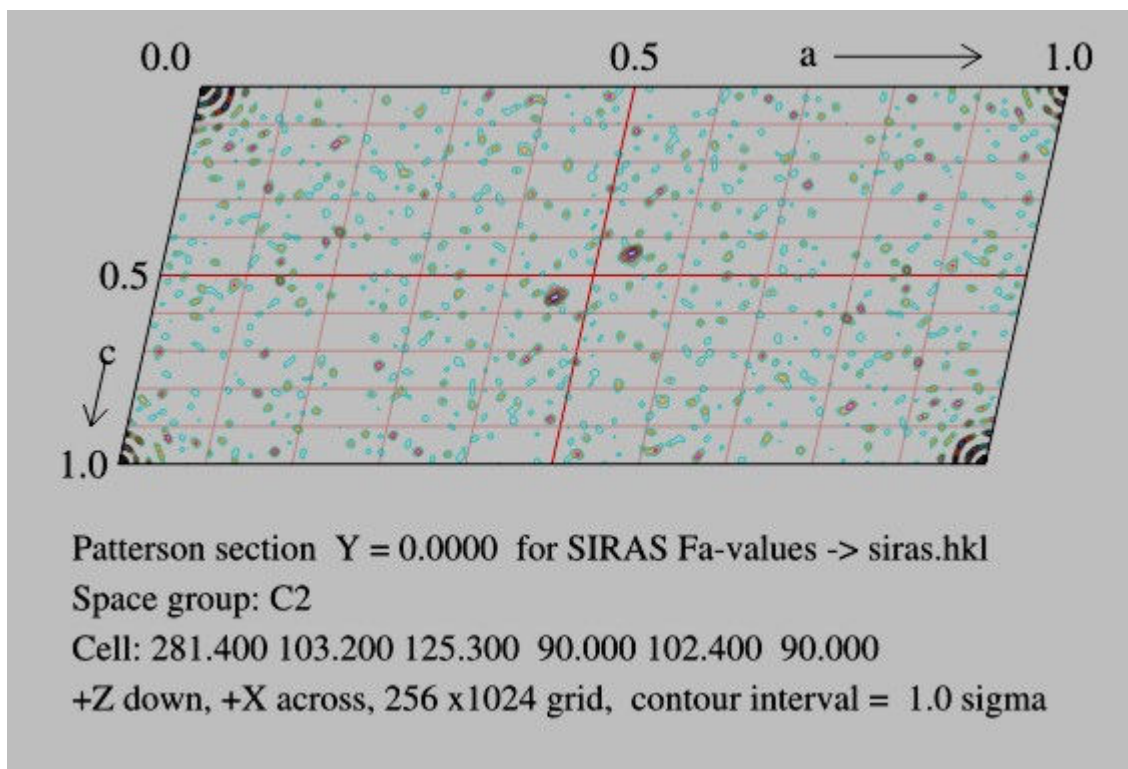


Figura R.29: Mapa de Patterson de diferències anòmales. Creat amb el programa XPREP per un fasejat tipus SIRAS emprant com a conjunt de dades natiu el cocristall amb ADP i com a conjunt de dades de derivat el cristall de bromur de tàntal recollit al pic d'absorció d'aquest àtom pesant.

L'estructura d'Aq_DnaB[1-439] es va resoldre mitjançant el mètode de reemplaçament isomòrfic senzill amb diferències anòmales (SIRAS) amb el paquet SHELX. Amb el programa SHELXD es van localitzar 6 àtoms de tàntal, les posicions dels quals es van refinar amb el programa SHELXE, que també es va emprar per calcular les primeres fases de l'estructura completa.

Al mapa de densitat electrònica generat amb SHELXE a 5 Å es podien distingir alguns elements d'estructura secundària. Això va permetre encaixar manualment el model atòmic per a un monòmer del domini C-terminal de G40P a aquest mapa de densitat electrònica inicial, tal i com es mostra a la figura R.30. Posteriorment, es va anar col·locant un a un cadascun dels dominis C-terminals dels sis monòmers que conformen l'hexàmer dins del mapa de densitat, intercalant etapes de refinat en l'espai real entre cada pas. Tot aquest procés d'encaix d'un model atòmic a la densitat es realitzà amb el programa Xfit.

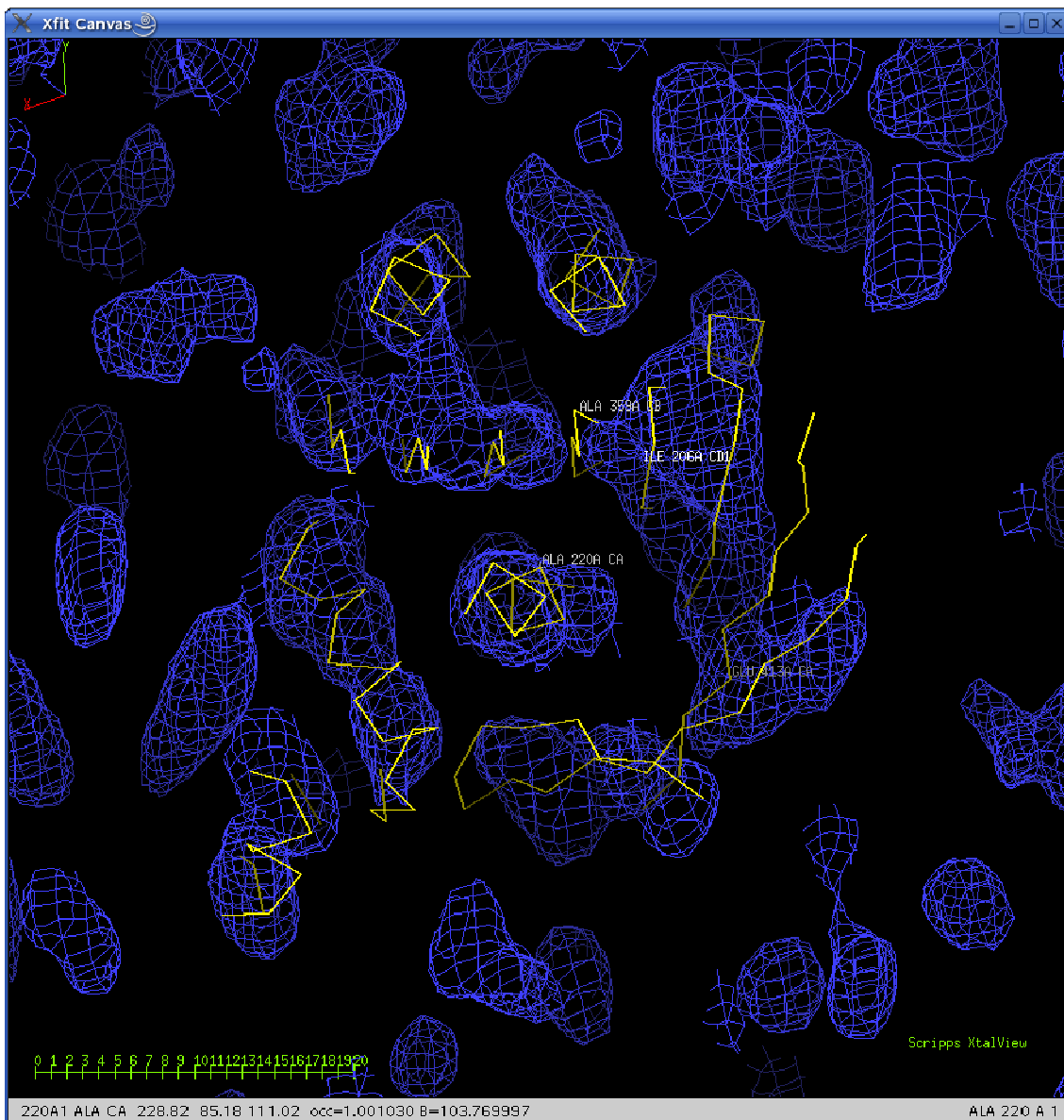


Figura R.30: Encaix inicial del model atòmic per al domini C-terminal de G40P al mapa de densitat electrònica experimental per a Aq_DnaB[1-439]. Fases inicials calculades amb el dels programa SHELXE a una resolució de 5 Å.

Un cop col·locats els dominis C-terminals, es procedí a encaixar dins la densitat el model de G40P per al domini N-terminal. Ara bé, degut a que la densitat electrònica del mapa experimental en aquesta zona era especialment pobre, no s'aconseguí col·locar cap dels monòmer N-terminals sencers. Donat que l'element d'estructura secundària més clarament distingible a baixa resolució són les hèlixs α , es decidí canviar d'estratègia i intentar localitzar al mapa de densitat electrònica l'estructura en forma de triangle que formen les parelles d'hèlixs α dels subdominis α -hairpin N-terminals de

DnaB (veure figura I.16 de la introducció). Aquesta estratègia funcionà i s'encaixaren satisfactòriament els tres dímers. Per contra, els subdominis globulars N-terminals no es varen poder encaixar totalment, doncs a alguns dels monòmers no hi havia densitat electrònica a la zona on haurien d'aparèixer.

Es varen analitzar la resta de conjunts de dades natives dels quals es disposava per intentar trobar algun en el qual aquesta zona estigués més ben definida. Quan s'empraven les reflexions del cocristall amb ADP i ssDNA, la densitat a la regió N-terminal era més clara i això va permetre acabar d'encaixar temptativament la resta de subdominis globulars N-terminals, de manera que finalment ja s'havia col·locat un model atòmic per a tot l'hexàmer dins el mapa de densitat experimental. Cal destacar que la posició dels subdominis globulars N-terminals d'Aq_DnaB[1-439] semblava variar lleugerament entre els diferents conjunts de dades recollits, indicant que aquesta zona presenta flexibilitat dins l'estructura global de l'hexàmer.

Per tal de millorar les fases inicials i, consegüentment, obtenir mapes més resolutius, es va intentar sense èxit aplicar tècniques de mitjanat entre els diferents monòmers, forçant simetria 6 no cristal·logràfica per a la regió C-terminal de l'hexàmer i simetria 3 no cristal·logràfica per a la regió N-terminal. S'emprà el programa DMMULTI (Cowtan 1994) per aquestes etapes de modificació de densitat i les màscares que diferenciaven ambdues regions es construïren amb el programa NCSMASK en base al model atòmic construït per la molècula sencera. Cap de les proves realitzades, aplicant el mitjanat a totes dues capes de l'hexàmer o només a una i amb diferents màscares inicials, aconseguí millorar el mapa de densitat electrònica inicial. Actualment, s'està intentant estendre la resolució de les fases inicials aplicant protocols de modificació de densitat amb altres programes.

Alternativament, si s'utilitza el model per a l'hexàmer d'Aq_DnaB[1-439] que s'ha construït sobre el mapa experimental per realitzar una cerca per reemplaçament molecular amb el programa Phaser, s'obté una sola solució molt clara amb un valor de Z de 16.3 per a la funció de rotació i de 14.3 per a la funció de translació, i un valor de LLG final de 242. Així doncs, combinant les fases experimentals amb les obtingudes

per reemplaçament molecular, s'intentarà millorar la resolució actual i acabar de refinar el model preliminar.

4.7 Descripció preliminar de l'estructura

4.7.1 Estructura global de l'helicasa hexamèrica

Tot i que el procés d'afinament no està acabat i, per tant, encara són dades preliminars, està clar que trobem un hexàmer d'Aq_DnaB[1-439] a la unitat asimètrica i que aquest es disposa en forma d'anell.

El model preliminar per a Aq_DnaB[1-439] actual es representa a la figura R.31. Al tractar-se d'un model no afinat i construït a baixa resolució, al voltant dels 4.5 Å, només es descriuen la forma global que prenen les diferents capes que conformen l'hexàmer.

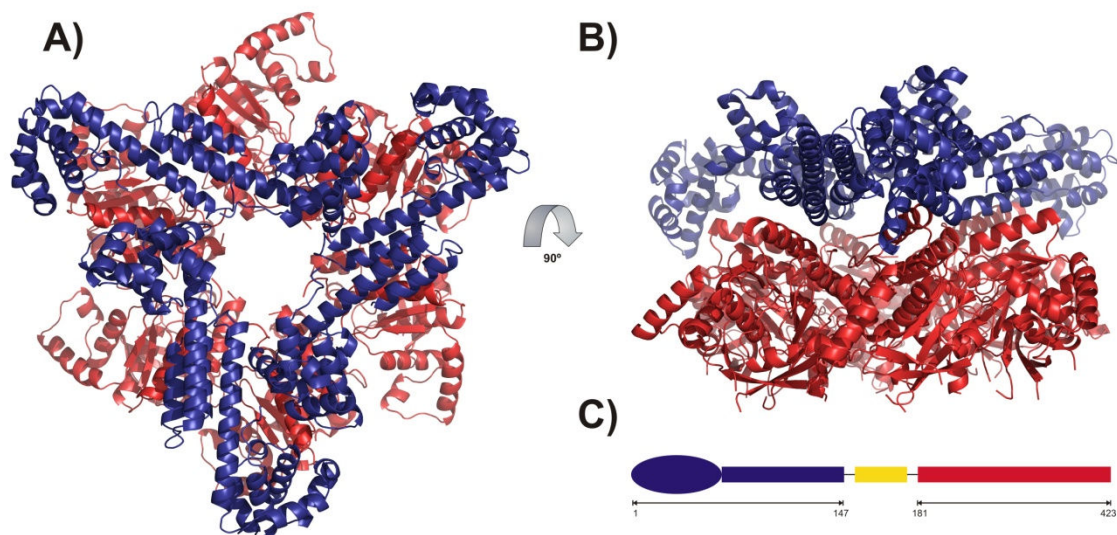


Figura R.31: Representació en cintes de l'estructura d'Aq_DnaB[1-439]. A) Vista frontal. **B)** Vista lateral. **C)** Representació esquemàtica dels dominis i les seves fronteres. Es representa en blau el domini N-terminal, format pel subdomini globular i el domini α -hairpin. En vermell, el domini C-terminal. La zona d'enllaç, en groc, encara no està localitzada a l'estructura resolta en aquest treball.

Tal i com ja s'havia observat a les dues estructures cristal·logràfiques d'hexàmers de la proteïna DnaB resoltes (Bailey *et al.* 2007b; Wang *et al.* 2008), l'hexàmer de la DnaB d'*Aquifex aeolicus* es disposa en dues capes de simetria diferent. La regió N-terminal pren una simetria rotacional tres clara, mentre que la capa formada pels dominis C-terminals té una disposició més irregular, però propra a simetria sis.

Les dimensions globals de l'hexàmer són d'uns 80 Å d'alçada i 100 Å d'amplada, amb un diàmetre del canal interior central d'uns 25 Å tant al llarg de la capa formada pels dominis N-terminals com al llarg de la capa formada pels dominis C-terminals.

4.7.2 Estructura del monòmer

El monòmer d'Aq_DnaB[1-439] està format per dos dominis: el domini N-terminal i el domini C-terminal, separats per una regió d'enllaç helicoïdal. Aquest plegament ja havia estat observat tant a l'estructura del monòmer de la DnaB de *Thermus aquaticus* (Bailey *et al.* 2007a) com a les estructures hexamèriques de DnaB (Bailey *et al.* 2007b; Wang *et al.* 2008).

El domini N-terminal compren els residus del 1 al 147 i està format exclusivament per hèlixs α , les quals s'organitzen en dos subdominis de plegament diferenciat. A la regió més N-terminal es situa el subdomini globular, format per un llaç de cinc hèlixs α , quatre de les quals s'embofallen al voltant d'una hèlix α central ($\alpha 1$). A continuació es troba el segon subdomini N-terminal, format per un α -hairpin estès entre la llarga l'hèlix $\alpha 6$, que es projecta des del subdomini globular, i l'hèlix $\alpha 7$, més curta.

El domini C-terminal, que compren els residus entre el 181 i el 423, té un plegament tipus RecA, format per un fulla β paral·lela de nou cadenes, embolcallada per tres hèlixs α a cada banda.

A la figura R.32 es mostra la representació en cintes del monòmer d'Aq_DnaB[1-439].

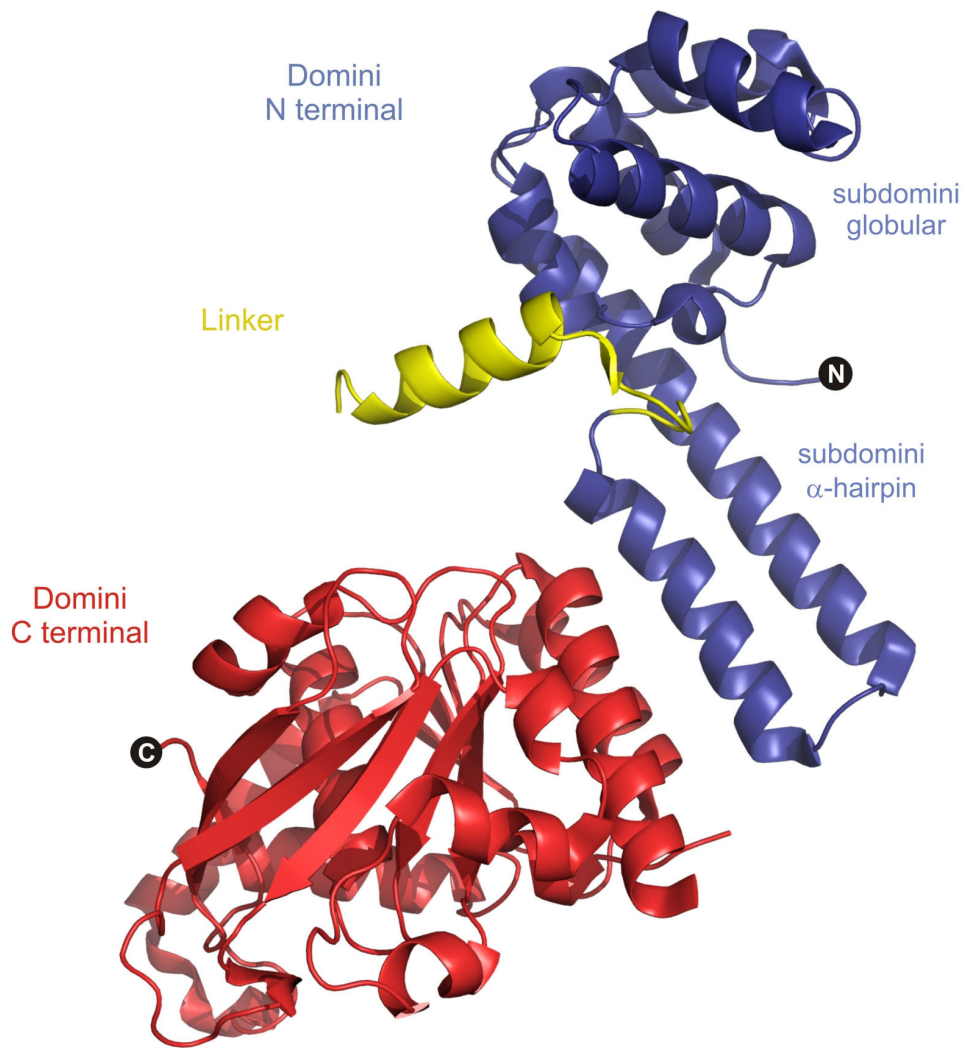


Figura R.32: Representació en cintes dels dominis del monòmer d'Aq_DnaB[1-439]. En blau es representa el domini N-terminal, del qual s'indiquen els dos subdominis. En groc es representa l'hèlix α que constitueix la regió d'enllaç. En vermell es representa el domini C-terminal.

DISCUSSIÓ

1. Producció de les proteïnes i obtenció de cristalls de qualitat

Donades les característiques de flexibilitat i els importants canvis conformacionals descrits per a les helicases replicatives, es va decidir treballar amb DnaB de bacteries termòfiles perquè les proteïnes procedents d'aquests organismes són més estables i menys flexibles que les proteïnes relacionades d'organismes mesòfils (Daniel *et al.* 2000), característiques que els confereixen una major propensió a formar cristalls ordenats. De fet, les estructures tridimensionals resoltes fins l'actualitat d'helicases hexamèriques són majoritàriament de termòfils o de virus, que també es caracteritzen per tenir proteïnes més compactes.

Primerament es treballà amb la DnaB sencera d'*Aquifex aeolicus*, de la qual es varen obtenir cristalls aïllats i tridimensionals que, tot i presentar un bon aspecte visual, no generaven dades de difracció de raigs X més enllà dels 7 Å de resolució. Es va invertir gran quantitat de temps i esforç per intentar millorar l'ordre intern d'aquests cristalls emprant múltiples tècniques de millora tant prèvies com posteriors a la seva formació. Malauradament, no s'aconseguí millorar la qualitat de les dades de difracció obtingudes a partir d'aquests cristalls.

Conseqüentment, es decidí provar la cristal·lització d'una DnaB homòloga d'una altra bactèria termòfila, *Thermotoga maritima*, així com modificar el constructe emprat originalment per a la DnaB d'*A. aeolicus* per generar una forma proteica més compacte, eliminant les regions no conservades i predites com a desordenades. Igualment, com DnaB és una proteïna organitzada en dos dominis separats per una llarga regió d'enllaç, que podria dificultar la cristal·lització de la proteïna sencera, es decidí provar només la

crystal·lització del domini C-terminal de la proteïna (el domini N-terminal de la DnaB d' *E. coli* ja havia estat resolt (Fass *et al.* 1999)).

Tot i que la DnaB de *T. maritima* s'expressà i purificà sense dificultats, totes les proves de crystal·lització realitzades amb aquesta proteïna foren infructuoses. Per tant, s'abandonà aquest constructe.

Pel que fa a les formes proteïques modificades de la DnaB d'*A. aeolicus*, només s'obtingueren cristalls per a la forma escurçada en 29 aminoàcids a la regió C-terminal (Aq_DnaB[1-439]), però no per a la forma corresponent al domini C-terminal de la proteïna (Aq_DnaB-CT).

Si comparem el constructe Aq_DnaB-CT amb les estructures actualment conegudes de DnaB, tant del monòmer com les de l'hexàmer complet, observem que el domini C-terminal de la DnaB d'*A. aeolicus* no estava correctament definit, sinó que també comprenia l'hèlix $\alpha 8$, que serveix d'enllaç entre els dominis N-terminal i C-terminal, així com uns cinc aminoàcids addicionals a la regió N-terminal (veure figura D.1). Aquest error va ser degut a una assignació incorrecte de les fronteres entre dominis del monòmer de DnaB descrita a la literatura en el moment de dissenyar aquest constructe, la qual s'havia basat en l'anàlisi estructural dels fragments proteolítics de la DnaB d'*Escherichia coli* (Fass *et al.* 1999; Weigelt *et al.* 1999). A la figura I.12 de l'apartat II de la introducció es mostren les fronteres anteriors i actuals per al domini N-terminal, en base als anàlisis estructurals recents del monòmer sencer de la DnaB de *Thermus aquaticus* (Bailey *et al.* 2007a).

La presència d'aquesta regió addicional podria dificultar l'empaquetament de les molècules d'Aq_DnaB-CT per formar cristalls. De fet, a l'estructura del monòmer de *Thermus aquaticus* s'observa que aquesta regió és altament flexible i presenta múltiples orientacions respecte el domini compacte C-terminal quan no té les restriccions imposades per la formació d'una estructura hexamèrica (Bailey *et al.* 2007a). Per contra, aquesta regió es troba ordenada a les estructures dels hexàmers de DnaB de *B. Stearothermophilus* i SPP1 degut a que està involucrada en les interaccions que mantenen units els sis dominis C-terminals que conformen l'anell (Bailey *et al.* 2007b;

Wang *et al.* 2008). La proteïna Aq_DnaB-CT es purificà tant en forma hexamèrica, purificació a pH 7.6, com dimèrica, purificació a pH 6.5, encara que en el primer cas s'observava certa tendència de l'hexàmer a dissociar-se en dímers si la purificació es realitzava amb tampons de força iònica moderada (200 mM NaCl). Segurament a les proves de cristal·lització, on les condicions de pH i força iònica són més extremes, aquesta forma proteica es trobaria en forma dimèrica o, fins i tot monomèrica, i, per tant, la regió corresponent a la zona d'enllaç entre dominis de DnaB no estaria implicada en el procés d'hexamerització i prendria múltiples conformacions que dificultarien la cristal·lització d'aquesta proteïna.

Finalment, la redefinició del constructe inicial de la DnaB d'*A. aeolicus*, que eliminava 29 residus a la regió C-terminal de la proteïna, va resultar clau per millorar la qualitat dels cristalls obtinguts per aquesta proteïna, emfatitzant la rellevància de realitzar anàlisis bioinformàtiques prèvies de les proteïnes que es volen caracteritzar estructuralment per tal de treballar ja des d'un bon principi amb diferents constructes de la proteïna sencera, que difereixin en la presència o absència de les zones no conservades i/o predites com a intrínsecament desordenades.

Els primers cristalls obtinguts per a Aq_DnaB[1-439] ja presentaven un poder de difracció major que els de la proteïna sencera. La posterior optimització d'aquests cristalls va permetre obtindre cocristalls amb ADP que difractaven fins a una resolució de 3.4 Å, tot i que era una difracció feble, fet que va complicar l'obtenció de conjunts de dades complets de derivats amb àtoms pesants a partir d'aquests cristalls.

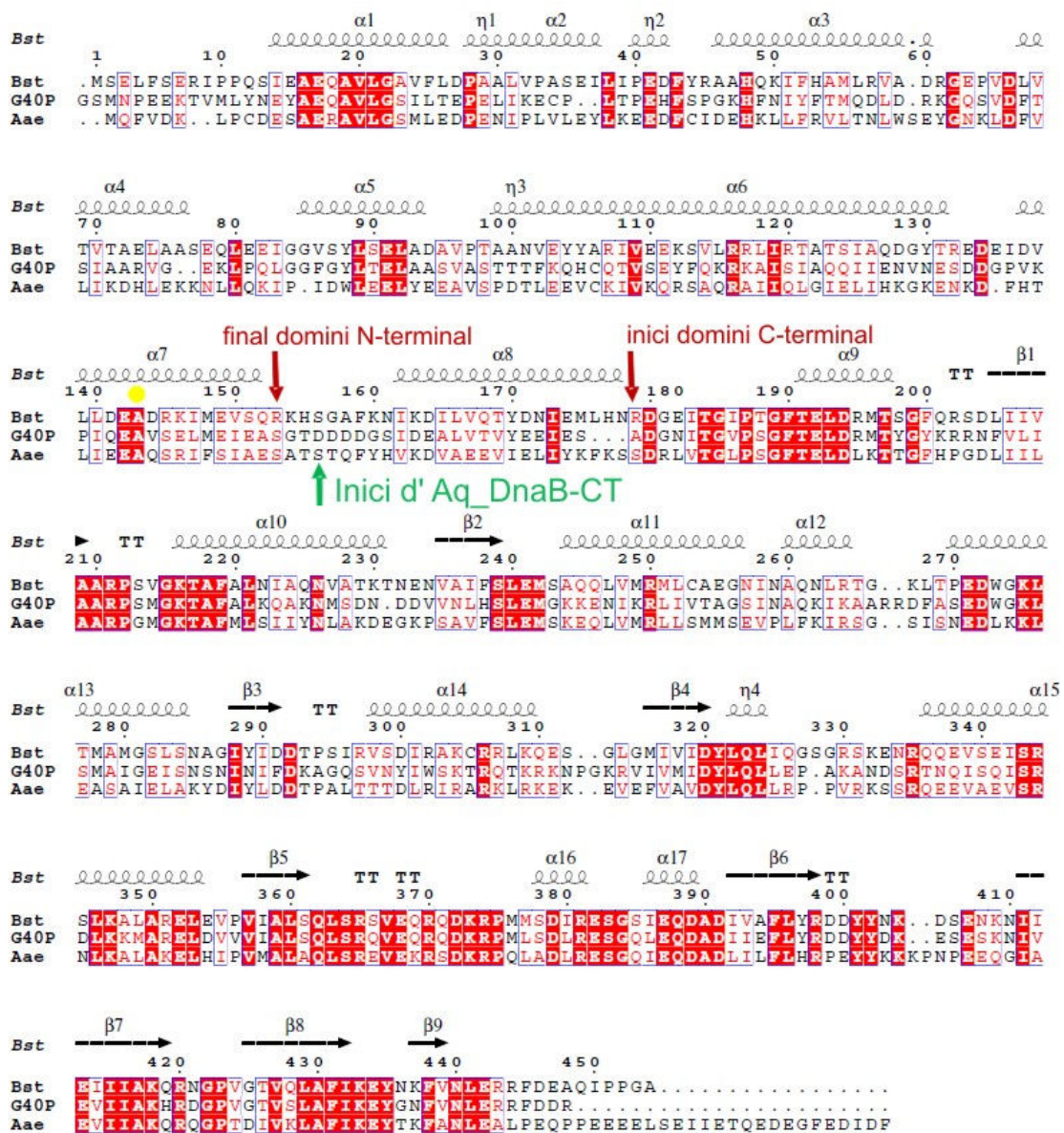


Figura D.1: Comparació de les fronteres entre els dominis de DnaB recentment descrites i les emprades en el disseny del constructe Aq_DnaB-CT. Es mostra una alineament de seqüències entre la DnaB d'*A. aeolicus* (Aae), la de *B. stearothermophilus* (Bst) i de G40P de SPP1 (G40P). A la part superior es representa l'estructura secundària determinada per cristal·lografia per als hexàmers de Bst i G40P. Les fletxes vermelles indiquen el final del domini N-terminal i l'inici del domini C-terminal. La fletxa verda indica el principi del constructe Aq_DnaB-CT amb que s'ha treballat en aquesta tesi. El punt groc representa el punt de finalització del domini N-terminal que havia estat descrit anteriorment mitjançant l'anàlisi dels fragments proteolítics de la DnaB d'*E. coli* per cristal·lografia, i que s'havia usat com a referència per a dissenyar el constructe Aq_DnaB-CT.

2. Resolució de l'estructura i construcció d'un model per a l'hexàmer d'Aq_DnaB

L'estructura de la forma curta de la DnaB d'*A. aeolicus* es va resoldre pel mètode del reemplaçament isomòrfic senzill amb dades anòmales (SIRAS) emprant com a conjunt de dades de derivat les imatges de difracció generades per un cocristall amb ADP submergit amb solució de bromur de tàntal a baixa concentració durant varis dies. Com a conjunt de dades natiu s'emprà el cocristall amb ADP que difractava a major resolució.

El model per a l'hexàmer d'Aq_DnaB present a la unitat asimètrica es construí encaixant manualment els models atòmics disponibles per als dominis C-terminals i N-terminals de G40P sobre el mapa de densitat electrònica experimental a una resolució de 4.5 Å. Les dimensions globals de l'hexàmer de la DnaB d'*A. aeolicus* resolt en aquest treball són similars a les prèviament determinades per als hexàmers de la DnaB de *B. steorothermophilus* i de G40P, tal i com es mostra a la figura D.2. Igualment, l'hexàmer d'Aq_DnaB també està organitzat en dues capes de simetria diferent, amb la regió formada pels dominis N-terminals en simetria 3 i la regió C-terminal en una disposició més irregular.

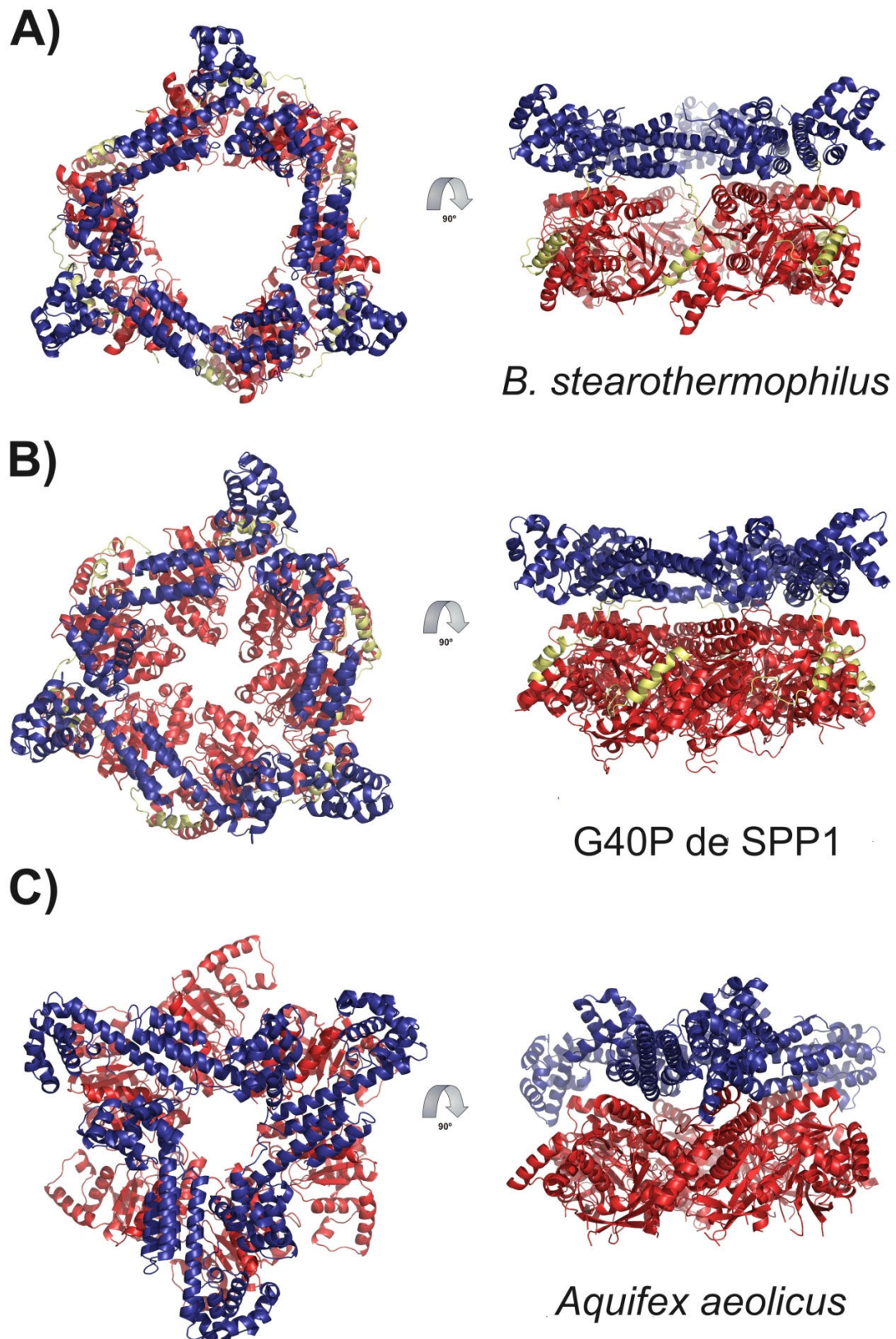


Figura D.2: Comparació de l'estructura global dels hexàmers de DnaB. A) DnaB de *B. stearothermophilus*, forma BH1. Té una amplada total de 115 Å i una alçada de 75 Å. **B)** G40P de SPP1, diàmetre de 120 Å i alçada de 75 Å. **C)** DnaB d'*Aquifex aeolicus*, amplada de 120 Å de i alçada de 80 Å.

3. Estructura tridimensional i moviment

3.1 Estructura dels glòbuls de l'hexàmer

La principal diferència amb les estructures prèvies es troba a la capa formada pels dominis N-terminalis de DnaB. A les dues estructures d'homòlegs de DnaB publicades, aquesta capa presenta un diàmetre per al seu canal central del voltant dels 45 Å, mentre que a la DnaB d'*A. aeolicus* aquest canal és molt més estret, d'uns 25 Å. Donat que l'amplada del DNA de doble cadena en la forma B és d'uns 20 Å, les estructures dels hexàmers de *B. stearothermophilus* i G40P poden acomodar tant dsDNA com ssDNA al seu interior, mentre que l'estructura de la DnaB d'*A. aeolicus* resulta en aquest treball possiblement només podria acomodar DNA de cadena senzilla. Conseqüentment, es trobaria més propera a la conformació que pren aquesta helicasa durant la seva activitat de separació de cadena, mentre que les estructures prèvies correspondrien al la forma de la proteïna inactiva com a helicasa o a la que funciona en la translocació sobre el dsDNA sense desenrotllar-lo (Kaplan *et al.* 2002).

3.1.1 Capa N-terminal

Tant a la DnaB de *B. stearothermophilus* com a G40P la capa formada pels dominis N-terminalis es troba estabilitzada en una simetria tres rígida gràcies a les interaccions que s'estableixen entre els monòmers N-terminalis, les quals són dobles i impliquen dues interfícies de contacte diferents. La interacció més forta i extensa s'estableix entre els subdominis α -hairpin, formats per dues hèlixs α en conformació estesa, de dos monòmers N-terminalis veïns. La segona interfície de dimerització implica el subdomini globular també de dos monòmers N-terminalis veïns. Així doncs, cada domini N-terminal dins l'hexàmer estableix dos tipus d'interaccions amb els monòmers adjacents, cadascuna d'elles mitjançada per un subdomini N-terminal diferent. A la figura D.3 es mostren aquestes interaccions per a la proteïna G40P.

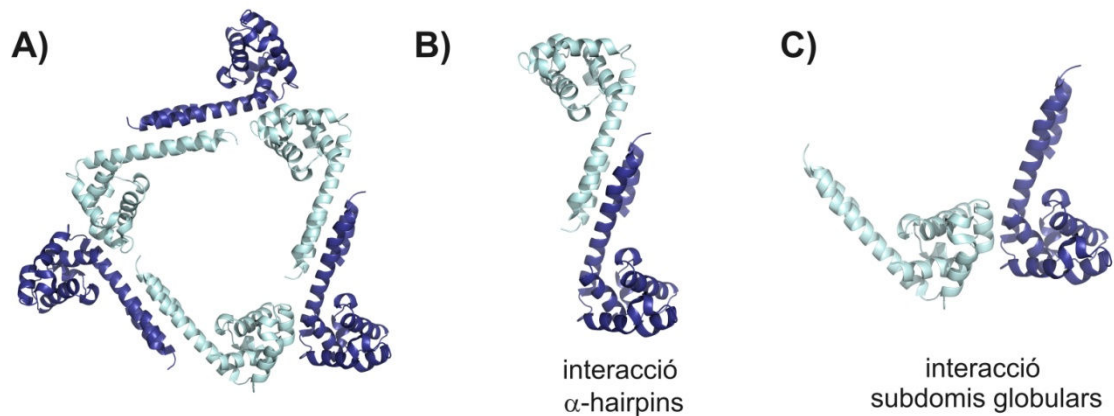


Figura D.3: Capa N-terminal de G40P. **A)** Vista frontal. **B)** Detall dels dímers formats mitjançant interaccions entre els subdominis α -hairpin. **C)** Detall dels dímers entre subdominis globulars.

El tancament d'aquest canal central al llarg de la capa N-terminal a la DnaB d'*A. aeolicus* es produeix degut a la desestabilització de les interaccions formades pels subdominis globulars, que en aquesta estructura es troben massa separats per a poder establir enllaços entre ells. En conseqüència, els subdominis globulars N-terminalns passen a estar més exposats al solvent i en general presenten una disposició més flexible a l'estructura resolta en aquest treball, doncs és la zona de l'hexàmer més poc definida dins el mapa de densitat electrònica i semblen prendre orientacions lleugerament diferents als diversos conjunts de dades de difracció dels quals disposem.

En canvi, la interacció entre els subdominis α -hairpin es manté a l'estructura de l'hexàmer de la DnaB d'*A. aeolicus*, fet que comporta que la regió N-terminal del hexàmer continuï presentant una organització en trímer de dímers.

A la figura D.4 es mostra la capa N-terminal d'Aq_DnaB[1-439] i les interaccions que s'estableixen entre monòmers.

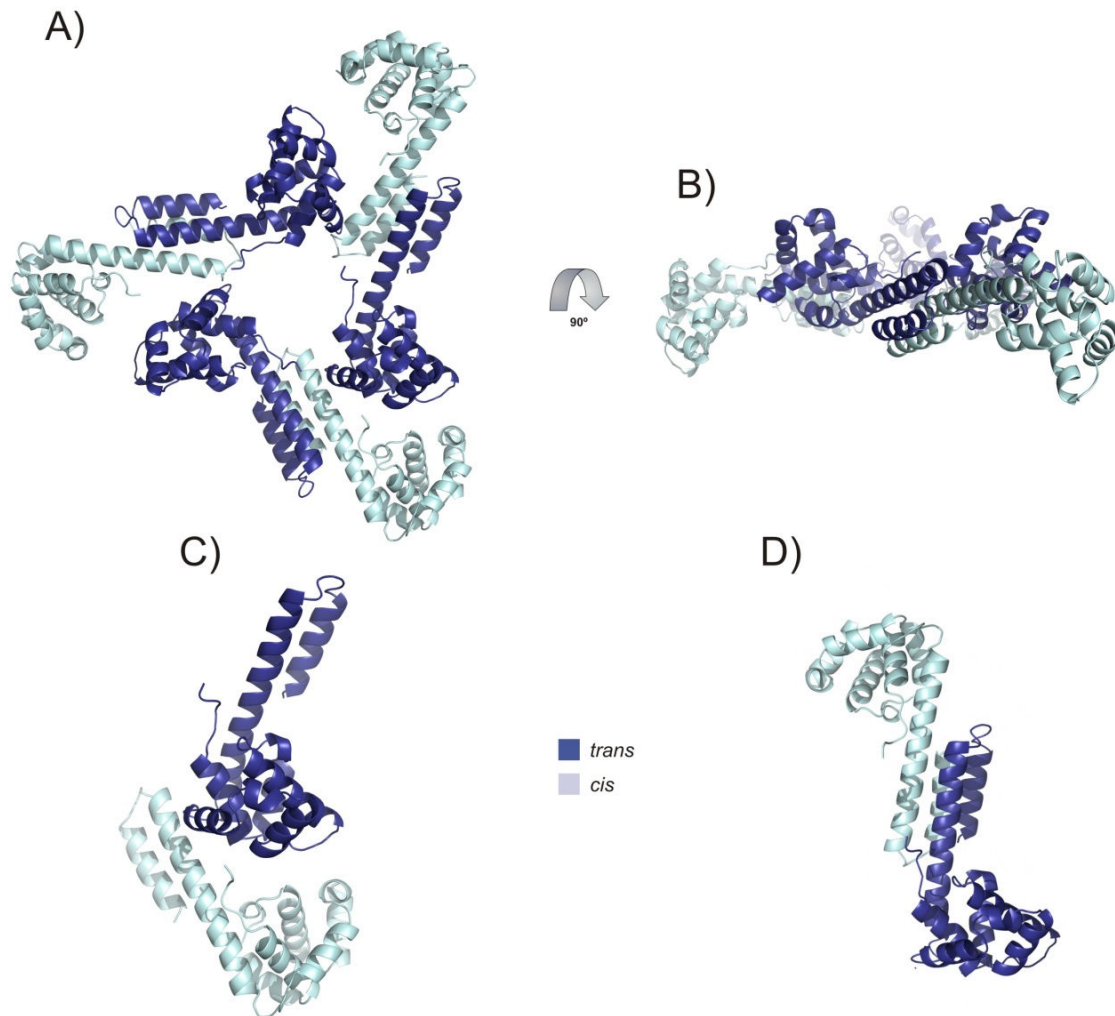


Figura D.4: Capa N-terminal de l'hexàmer de la DnaB d'*A. aeolicus*. **A)** Vista frontal del trímer de dímers. Els monòmers alternats es representen en blau clar i blau fosc. **B)** Vista lateral. **C)** Els subdominis globulars es troben massa separats per a interaccionar entre ells. **D)** El dímers formats mitjançant els subdominis α -hairpin es mantenen.

Per tant, la interacció entre els subdominis α -hairpin seria la responsable del manteniment de la simetria trimèrica de la capa N-terminal. Al seu torn, la interacció entre els subdominis globulars ajudaria a establir una conformació oberta de l'hexàmer, capaç d'acomodar DNA de cadena doble, però és prou feble per a trencar-se i permetre que el canal central de la regió N-terminal adopti una conformació més tancada, com l'observada a l'estructura de l'hexàmer d'*A. aeolicus* resolt en aquest treball.

3.1.2 Capa C-terminal

Pel que fa a la capa formada pels dominis C-terminal de DnaB, l'orientació relativa dels monòmers que la conformen, així com el diàmetre del canal central al llarg d'aquesta regió, és altament variable a les diferents estructures hexamèriques resoltes, tal i com es mostra a la figura D.5. Aquesta flexibilitat conformacional s'ha atribuït a l'absència d'interaccions estabilitzadores entre els diferents monòmers C-terminal, que bàsicament es mantenen propers gràcies a trobar-se enllaçats amb la capa N-terminal rígida.

Un element important en l'estabilització de la capa C-terminal és l'hèlix $\alpha 8$ que constitueix la regió d'enllaç entre dominis. En les estructures hexamèriques de DnaB prèvies, l'hèlix d'un monòmer es projectava cap al monòmer C-terminal veí, amb el qual establia interaccions. S'havia proposat que aquesta hèlix també estaria involucrada en la transmissió i la coordinació dels canvis conformacionals als sis monòmers. A l'estructura d'Aq_DnaB[1-439] que es descriu en aquest treball, encara no s'han localitzat inequívocament les posicions de aquestes hèlixs d'enllaç. Tot i que semblen trobar-se en una posició semblant a la descrita prèviament als altres hexàmers de DnaB, aquesta regió és més mòvil a la present estructura. Aquest augment de la mobilitat d'aquesta zona de transmissió de canvis podria estar relacionada amb els canvis a la capa N-terminal que s'han descrit a l'apartat anterior, i que situarien a l'estructura d'Aq_DnaB[1-439] més propera a la forma funcionalment activa de l'helicasa. Durant l'activitat de separació de cadena de l'helicasa, el ssDNA es translocat a través del seu canal central i això provocaria un augment de la flexibilitat de la capa C-terminal per possibilitar els moviments dels llaços d'unió a ssDNA que actuen durant la translocació de l'àcid nucleic.

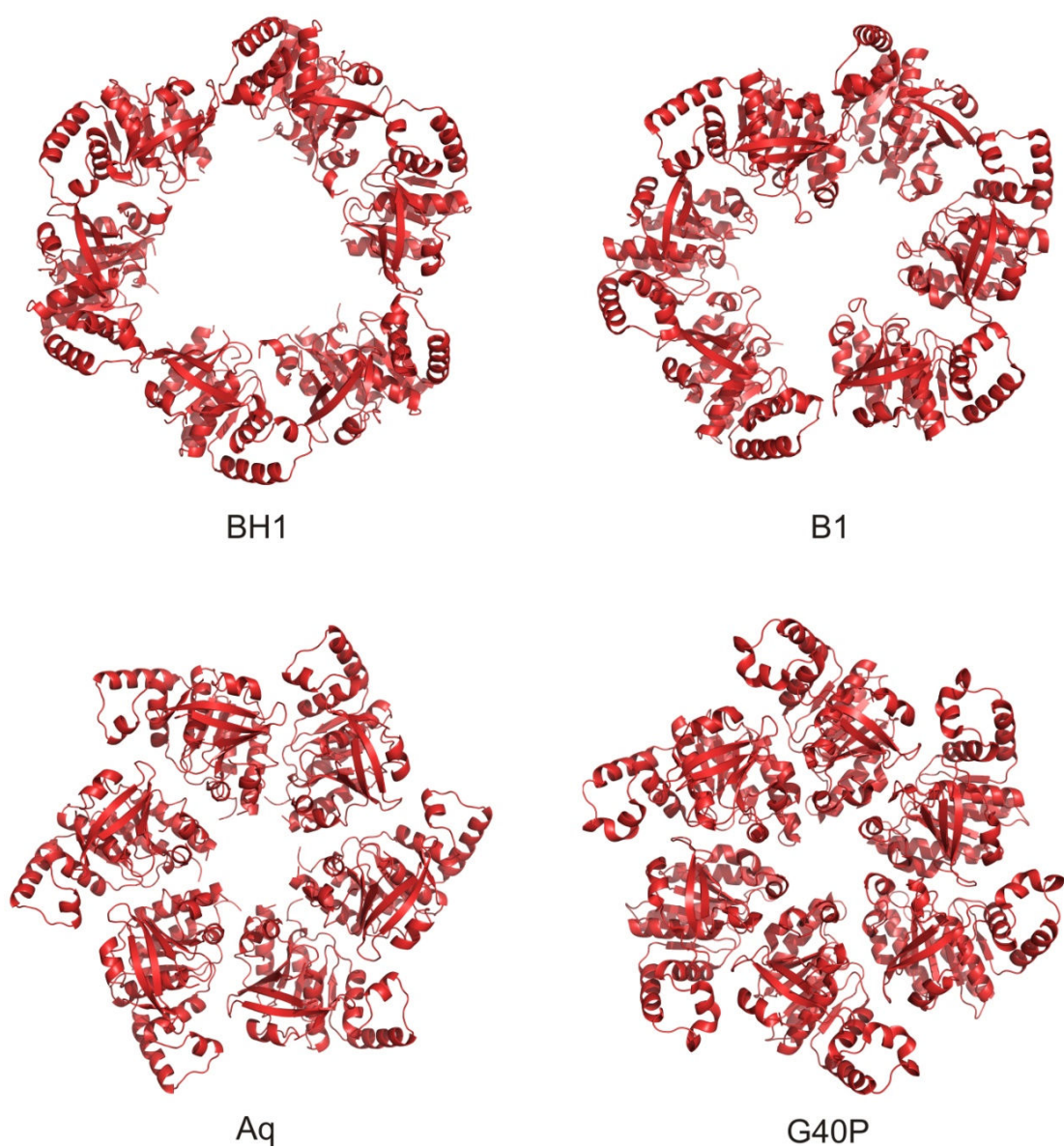


Figura D.5: Comparativa de les diferents orientacions que prenen els dominis C-terminals a les estructures hexamèriques de DnaB. BH1: forma cristal·lina BH1 de la DnaB de *Bacillus stearothermophilus* en complex amb el domini HBD de la primasa. Codi PDB: 2R6A. **B1:** forma cristal·lina B1 de la DnaB de *Bacillus stearothermophilus*. Codi PDB: 2R6D. **Aq:** Aq_DnaB[1-439]. Aquest treball. **G40P:** homòleg de DnaB del bacteriòfag SPP1. Codi PDB: 3B6W.

3.2 Centre actiu

Tal i com és característic de les helicases hexamèriques, les butxaques d'unió a nucleòtid de l'hexàmer de DnaB es troben a les interfícies entre dominis C-terminals adjacents, doncs necessiten la contribució de residus procedents de dos dominis C-terminals veïns. Conseqüentment, els canvis d'orientació relativa i diàmetre del canal central de la regió C-terminal estan probablement associats als diferents estadis dintre del cicle d'hidròlisi d'ATP de l'hexàmer de DnaB. És a dir, el moviment relatiu entre dominis C-terminals conduïts per l'estat d'unió a nucleòtid de la butxaca nucleotídica bipartida possibilitaria la translocació de l'hexàmer al llarg del DNA de cadena senzilla, així com la separació del DNA dúplex a través d'un mecanisme d'exclusió estèrica.

3.3 Unió a ssDNA

Per a poder descriure el mecanisme exacte d'aquesta translocació, caldria conèixer l'estructura que prenen els dominis C-terminals de l'hexàmer de DnaB en complex amb nucleòtid o ssDNA. Encara que els cristalls d'Aq_DnaB emprats en aquest treball contenen ADP, els mapes de densitat electrònica actuals no permeten localitzar-lo. Caldrà esperar a l'acabament de l'etapa d'afinat de l'estructura per saber quina ocupància presenten les sis butxaques nucleotídiques presents a l'hexàmer, així com si aquestes prenen conformacions diferents.

Igualment, encara que es disposa de dades de difracció d'un cocristall amb ssDNA, tampoc s'ha pogut localitzar densitat per a l'àcid nucleic als mapes actuals. Fins i tot si finalment s'aconseguís estendre les fases actuals fins a la resolució màxima de les dades natives (al voltant dels 3.5 Å), es tractaria d'una resolució moderada que permetria només una interpretació limitada. Per tant, caldrien nous estudis per a poder formular un model estructural acurat de funcionament de l'helicasa DnaB.

CONCLUSIONS

1. La proteïna DnaB d'*Aquifex aeolicus* sencera, així com la forma Aq_DnaB[1-439] i Aq_DnaB-CT, corresponents respectivament a una proteïna 29 residus més curta i al domini C-terminal, s'expressen de forma soluble a *E. coli*. També s'expressa soluble a *E. coli* la proteïna DnaB de *Thermotoga maritima* sencera.

2. S'ha establert un protocol de purificació per cadascuna de les quatre formes proteïques estudiades, aprofitant l'estabilitat d'aquestes proteïnes a les altes temperatures.

3. S'han obtingut cristalls per a la proteïna DnaB d'*A. aeolicus* sencera i la seva forma curta. Els cristalls de la proteïna sencera no generaven dades de difracció de qualitat i no es van poder millorar amb tècniques de millora pre o postcristal·lització. Els cristalls de la forma curta, Aq_DnaB[1-439], cocrystal·litzats amb ADP generaven dades de difracció al voltant dels 3.5 Å, pertanyien al grup espacial C2 amb uns paràmetres de cel·la de $a = 282.51 \text{ \AA}$, $b = 103.31 \text{ \AA}$, $c = 125.25 \text{ \AA}$ i $\beta = 102.43^\circ$ i contenien un hexàmer a la unitat asimètrica. Es van obtenir també dades de difracció d'un cocrystal amb ssDNA de 21 bases, isomòrfic amb el descrit anteriorment.

4. La utilització de reconstruccions de criomicroscòpia d'Aq_DnaB[1-439] com a models de cerca en el mètode de reemplaçament molecular enfront de dades de difracció recollides a molt baixa resolució (uns 200 Å) no va permetre resoldre l'estructura de la proteïna, encara que sí trobar l'orientació de la partícula hexamèrica.

5. La utilització dels models atòmics per al monòmer de la DnaB de *Thermus aquaticus*, l'hexàmer de la DnaB de *Bacillus stearothermophilus* i la G40P de SPP1 com a models de cerca per a reemplaçament molecular tampoc va permetre resoldre l'estructura d'Aq_DnaB[1-439].

6. L'estructura tridimensional d'Aq_DnaB es va resoldre pel mètode de SIRAS emprant com a conjunt de dades derivats un cristall submergit en una solució de bromur de tàntal recollit al pic d'absorció d'aquest àtom pesat i com a conjunt de dades natiu un cocristall amb ADP.

7. El model de l'hexàmer d'Aq_DnaB[1-439] construït a una resolució de 4.5 Å mostra una estructura en dos capes de simetria diferenciada, amb la regió formada pels dominis N-terminals en simetria tres i la regió formada pels dominis C-terminals amb simetria propera a sis. Les dimensions globals de l'hexàmer són de 80 Å d'alçada i 100 Å d'amplada, amb un diàmetre del canal interior central d'uns 25 Å tant al llarg de la capa N-terminal com C-terminal.

8. La principal diferència amb les estructures hexamèriques de DnaB resoltes anteriorment es troba a la capa N-terminal, que presenta un diàmetre entre 20 i 25 Å més estret. Aquest estretament bé marcat pel trencament d'una de les dues superfícies d'interacció entre monòmers N-terminals observades a les estructures prèvies. Concretament, la interfície entre els subdominis globular N-terminals es trenca, doncs aquest es troben massa separats per formar enllaços. Això provoca un augment de la flexibilitat d'aquests subdominis globulars, que es troben en posicions diferents als diferents cristalls analitzats en aquest treball.

BIBLIOGRAFIA

- Abdel-Monem, M. i Hoffmann-Berling, H. (1976). "Enzymic unwinding of DNA. 1. Purification and characterization of a DNA-dependent ATPase from *Escherichia coli*." *Eur J Biochem* **65**(2): 431-40.
- Abe, Y., Jo, T., Matsuda, Y., Matsunaga, C., Katayama, T. i Ueda, T. (2007). "Structure and function of DnaA N-terminal domains: specific sites and mechanisms in inter-DnaA interaction and in DnaB helicase loading on *oriC*." *J Biol Chem* **282**(24): 17816-27.
- Abrahams, J. P. i Leslie, A. G. W. (1996). "Methods used in the structure determination of bovine mitochondrial F1 ATPase." *Acta Crystallogr D Biol Crystallogr* **52**: 30-42.
- Adelman, J. L., Jeong, Y. J., Liao, J. C., Patel, G., Kim, D. E., Oster, G. i Patel, S. S. (2006). "Mechanochemistry of transcription termination factor Rho." *Mol Cell* **22**(5): 611-21.
- Allen, G. C., Jr. i Kornberg, A. (1991). "Fine balance in the regulation of DnaB helicase by DnaC protein in replication in *Escherichia coli*." *J Biol Chem* **266**(33): 22096-101.
- Arai, K. i Kornberg, A. (1979). "A general priming system employing only *dnaB* protein and primase for DNA replication." *Proc Natl Acad Sci U S A* **76**(9): 4308-12.
- Arai, K. i Kornberg, A. (1981a). "Mechanism of *dnaB* protein action. II. ATP hydrolysis by *dnaB* protein dependent on single- or double-stranded DNA." *J Biol Chem* **256**(10): 5253-9.
- Arai, K. i Kornberg, A. (1981b). "Mechanism of *dnaB* protein action. III. Allosteric role of ATP in the alteration of DNA structure by *dnaB* protein in priming replication." *J Biol Chem* **256**(10): 5260-6.
- Bailey, S., Eliason, W. K. i Steitz, T. A. (2007a). "The crystal structure of the *Thermus aquaticus* DnaB helicase monomer." *Nucleic Acids Res* **35**(14): 4728-36.
- Bailey, S., Eliason, W. K. i Steitz, T. A. (2007b). "Structure of hexameric DnaB helicase and its complex with a domain of DnaG primase." *Science* **318**(5849): 459-63.
- Bailey, S., Wing, R. A. i Steitz, T. A. (2006). "The structure of *T. aquaticus* DNA polymerase III is distinct from eukaryotic replicative DNA polymerases." *Cell* **126**(5): 893-904.

- Baker, T. A., Funnell, B. E. i Kornberg, A. (1987). "Helicase action of dnaB protein during replication from the Escherichia coli chromosomal origin in vitro." J Biol Chem **262**(14): 6877-85.
- Barcena, M., Ruiz, T., Donate, L. E., Brown, S. E., Dixon, N. E., Radermacher, M. i Carazo, J. M. (2001). "The DnaB.DnaC complex: a structure based on dimers assembled around an occluded channel." Embo J **20**(6): 1462-8.
- Berger, J. M. (2008). "SnapShot: nucleic acid helicases and translocases." Cell **134**(5): 888-888 e1.
- Bergfors, T. (2003). "Seeds to crystals." J Struct Biol **142**(1): 66-76.
- Betterton, M. D. i Julicher, F. (2005). "Opening of nucleic-acid double strands by helicases: active versus passive opening." Phys Rev E Stat Nonlin Soft Matter Phys **71**(1 Pt 1): 011904.
- Bhattacharyya, S. i Griep, M. A. (2000). "DnaB helicase affects the initiation specificity of Escherichia coli primase on single-stranded DNA templates." Biochemistry **39**(4): 745-52.
- Bird, L. E., Pan, H., Soultanas, P. i Wigley, D. B. (2000). "Mapping protein-protein interactions within a stable complex of DNA primase and DnaB helicase from Bacillus stearothermophilus." Biochemistry **39**(1): 171-82.
- Biswas, E. E. i Biswas, S. B. (1999). "Mechanism of DnaB helicase of Escherichia coli: structural domains involved in ATP hydrolysis, DNA binding, and oligomerization." Biochemistry **38**(34): 10919-28.
- Biswas, E. E., Biswas, S. B. i Bishop, J. E. (1986). "The dnaB protein of Escherichia coli: mechanism of nucleotide binding, hydrolysis, and modulation by dnaC protein." Biochemistry **25**(23): 7368-74.
- Biswas, S. B., Chen, P. H. i Biswas, E. E. (1994). "Structure and function of Escherichia coli DnaB protein: role of the N-terminal domain in helicase activity." Biochemistry **33**(37): 11307-14.
- Bloom, L. B. (2006). "Dynamics of loading the Escherichia coli DNA polymerase processivity clamp." Crit Rev Biochem Mol Biol **41**(3): 179-208.
- Bochman, M. L. i Schwacha, A. (2008). "The Mcm2-7 complex has in vitro helicase activity." Mol Cell **31**(2): 287-93.
- Bradford, M. M. (1976). "A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding." Anal Biochem **72**: 248-54.
- Braithwaite, D. K. i Ito, J. (1993). "Compilation, alignment, and phylogenetic relationships of DNA polymerases." Nucleic Acids Res **21**(4): 787-802.
- Bujalowski, W. (2003). "Expanding the physiological role of the hexameric DnaB helicase." Trends Biochem Sci **28**(3): 116-8.

- Bujalowski, W. i Jezewska, M. J. (1995). "Interactions of Escherichia coli primary replicative helicase DnaB protein with single-stranded DNA. The nucleic acid does not wrap around the protein hexamer." Biochemistry **34**(27): 8513-9.
- Bujalowski, W. i Klonowska, M. M. (1993). "Negative cooperativity in the binding of nucleotides to Escherichia coli replicative helicase DnaB protein. Interactions with fluorescent nucleotide analogs." Biochemistry **32**(22): 5888-900.
- Bujalowski, W., Klonowska, M. M. i Jezewska, M. J. (1994). "Oligomeric structure of Escherichia coli primary replicative helicase DnaB protein." J Biol Chem **269**(50): 31350-8.
- Buttner, K., Nehring, S. i Hopfner, K. P. (2007). "Structural basis for DNA duplex separation by a superfamily-2 helicase." Nat Struct Mol Biol **14**(7): 647-52.
- Byrd, A. K. i Raney, K. D. (2005). "Increasing the length of the single-stranded overhang enhances unwinding of duplex DNA by bacteriophage T4 Dda helicase." Biochemistry **44**(39): 12990-7.
- Caruthers, J. M. i McKay, D. B. (2002). "Helicase structure and mechanism." Curr Opin Struct Biol **12**(1): 123-33.
- Corn, J. E. i Berger, J. M. (2006). "Regulation of bacterial priming and daughter strand synthesis through helicase-primase interactions." Nucleic Acids Res **34**(15): 4082-8.
- Corn, J. E., Pease, P. J., Hura, G. L. i Berger, J. M. (2005). "Crosstalk between primase subunits can act to regulate primer synthesis in trans." Mol Cell **20**(3): 391-401.
- Corn, J. E., Pelton, J. G. i Berger, J. M. (2008). "Identification of a DNA primase template tracking site redefines the geometry of primer synthesis." Nat Struct Mol Biol **15**(2): 163-9.
- Costa, A. i Onesti, S. (2008). "The MCM complex: (just) a replicative helicase?" Biochem Soc Trans **36**(Pt 1): 136-40.
- Cowtan, K. (1994). "DMMULTI: a multi-crystal density modification package." Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography (31): 34-38.
- Crampton, D. J., Mukherjee, S. i Richardson, C. C. (2006). "DNA-induced switch from independent to sequential dTTP hydrolysis in the bacteriophage T7 DNA helicase." Mol Cell **21**(2): 165-74.
- Chang, P. i Marians, K. J. (2000). "Identification of a region of Escherichia coli DnaB required for functional interaction with DnaG at the replication fork." J Biol Chem **275**(34): 26187-95.
- Chayen, N. E. i Saridakis, E. (2008). "Protein crystallization: from purified protein to diffraction-quality crystal." Nat Methods **5**(2): 147-53.

- Chayen, N. E., Saridakis, E., El-Bahar, R. i Nemirovsky, Y. (2001). "Porous silicon: an effective nucleation-inducing material for protein crystallization." J Mol Biol **312**(4): 591-5.
- Cheng, W., Dumont, S., Tinoco, I., Jr. i Bustamante, C. (2007). "NS3 helicase actively separates RNA strands and senses sequence barriers ahead of the opening fork." Proc Natl Acad Sci U S A **104**(35): 13954-9.
- Daniel, R. M. i Cowan, D. A. (2000). "Biomolecular stability and life at high temperatures." Cell Mol Life Sci **57**(2): 250-64.
- Davey, M. J., Fang, L., McInerney, P., Georgescu, R. E. i O'Donnell, M. (2002). "The DnaC helicase loader is a dual ATP/ADP switch protein." Embo J **21**(12): 3148-59.
- Davey, M. J. i O'Donnell, M. (2003). "Replicative helicase loaders: ring breakers and ring makers." Curr Biol **13**(15): R594-6.
- Delagoutte, E. i von Hippel, P. H. (2002). "Helicase mechanisms and the coupling of helicases within macromolecular machines. Part I: Structures and properties of isolated helicases." Q Rev Biophys **35**(4): 431-78.
- Donate, L. E., Llorca, O., Barcena, M., Brown, S. E., Dixon, N. E. i Carazo, J. M. (2000). "pH-controlled quaternary states of hexameric DnaB helicase." J Mol Biol **303**(3): 383-93.
- Donmez, I. i Patel, S. S. (2008). "Coupling of DNA unwinding to nucleotide hydrolysis in a ring-shaped helicase." Embo J **27**(12): 1718-26.
- Donmez, I., Rajagopal, V., Jeong, Y.-J. i Patel, S. S. (2007). "Nucleic Acid Unwinding by Hepatitis C Virus and Bacteriophage T7 Helicases Is Sensitive to Base Pair Stability
10.1074/jbc.M702136200." J. Biol. Chem. **282**(29): 21116-21123.
- Doublet, S. (1997). "Preparation of selenomethionyl proteins for phase determination." Methods Enzymol **276**: 523-30.
- Dubendorff, J. W. i Studier, F. W. (1991). "Controlling basal expression in an inducible T7 expression system by blocking the target T7 promoter with lac repressor." J Mol Biol **219**(1): 45-59.
- Duderstadt, K. E. i Berger, J. M. (2008). "AAA+ ATPases in the initiation of DNA replication." Crit Rev Biochem Mol Biol **43**(3): 163-87.
- Egelman, E. H., Yu, X., Wild, R., Hingorani, M. M. i Patel, S. S. (1995). "Bacteriophage T7 helicase/primase proteins form rings around single-stranded DNA that suggest a general structure for hexameric helicases." Proc Natl Acad Sci U S A **92**(9): 3869-73.
- Enemark, E. J. i Joshua-Tor, L. (2006). "Mechanism of DNA translocation in a replicative hexameric helicase." Nature **442**(7100): 270-5.

- Enemark, E. J. i Joshua-Tor, L. (2008). "On helicases and other motor proteins." Curr Opin Struct Biol **18**(2): 243-57.
- Erzberger, J. P. i Berger, J. M. (2006a). "Evolutionary relationships and structural mechanisms of AAA+ proteins." Annu Rev Biophys Biomol Struct **35**: 93-114.
- Erzberger, J. P., Mott, M. L. i Berger, J. M. (2006b). "Structural basis for ATP-dependent DnaA assembly and replication-origin remodeling." Nat Struct Mol Biol **13**(8): 676-83.
- Erzberger, J. P., Pirruccello, M. M. i Berger, J. M. (2002). "The structure of bacterial DnaA: implications for general mechanisms underlying DNA replication initiation." Embo J **21**(18): 4763-73.
- Evans, P. R. (1993). "Data reduction." Proceedings of CCP4 Study Weekend On Data Collection & Processing(1993): 114-122.
- Fang, L., Davey, M. J. i O'Donnell, M. (1999). "Replisome assembly at oriC, the replication origin of E. coli, reveals an explanation for initiation sites outside an origin." Mol Cell **4**(4): 541-53.
- Fass, D., Bogden, C. E. i Berger, J. M. (1999). "Crystal structure of the N-terminal domain of the DnaB hexameric helicase." Structure **7**(6): 691-8.
- Fischer, C. J., Maluf, N. K. i Lohman, T. M. (2004). "Mechanism of ATP-dependent translocation of E.coli UvrD monomers along single-stranded DNA." J Mol Biol **344**(5): 1287-309.
- Fouts, E. T., Yu, X., Egelman, E. H. i Botchan, M. R. (1999). "Biochemical and electron microscopic image analysis of the hexameric E1 helicase." J Biol Chem **274**(7): 4447-58.
- Fujikawa, N., Kurumizaka, H., Nureki, O., Terada, T., Shirouzu, M., Katayama, T. i Yokoyama, S. (2003). "Structural basis of replication origin recognition by the DnaA protein." Nucleic Acids Res **31**(8): 2077-86.
- Funnell, B. E., Baker, T. A. i Kornberg, A. (1987). "In vitro assembly of a prepriming complex at the origin of the Escherichia coli chromosome." J Biol Chem **262**(21): 10327-34.
- Gai, D., Zhao, R., Li, D., Finkielstein, C. V. i Chen, X. S. (2004). "Mechanisms of conformational change for a replicative hexameric helicase of SV40 large tumor antigen." Cell **119**(1): 47-60.
- Gallant, P. (2007). "Control of transcription by Pontin and Reptin." Trends Cell Biol **17**(4): 187-92.
- Galletto, R., Jezewska, M. J. i Bujalowski, W. (2004a). "Unzipping Mechanism of the Double-stranded DNA Unwinding by a Hexameric Helicase: Quantitative Analysis of the Rate of the dsDNA Unwinding, Processivity and Kinetic Step-size of the Escherichia coli DnaB Helicase Using Rapid Quench-flow Method." Journal of Molecular Biology **343**(1): 83-99.

- Galletto, R., Jezewska, M. J. i Bujalowski, W. (2004b). "Unzipping mechanism of the double-stranded DNA unwinding by a hexameric helicase: the effect of the 3' arm and the stability of the dsDNA on the unwinding activity of the Escherichia coli DnaB helicase." J Mol Biol **343**(1): 101-14.
- Garman, E. i Murray, J. W. (2003). "Heavy-atom derivatization." Acta Crystallogr D Biol Crystallogr **59**(Pt 11): 1903-13.
- Garman, E. F. i Owen, R. L. (2006). "Cryocooling and radiation damage in macromolecular crystallography." Acta Crystallogr D Biol Crystallogr **62**(Pt 1): 32-47.
- Georgescu, R. E., Kim, S. S., Yurieva, O., Kuriyan, J., Kong, X. P. i O'Donnell, M. (2008). "Structure of a sliding clamp on DNA." Cell **132**(1): 43-54.
- Glover, B. P. i McHenry, C. S. (1998). "The chi psi subunits of DNA polymerase III holoenzyme bind to single-stranded DNA-binding protein (SSB) and facilitate replication of an SSB-coated template." J Biol Chem **273**(36): 23476-84.
- Gomez-Llorente, Y., Fletcher, R. J., Chen, X. S., Carazo, J. M. i San Martin, C. (2005). "Polymorphism and double hexamer structure in the archaeal minichromosome maintenance (MCM) helicase from Methanobacterium thermoautotrophicum." J Biol Chem **280**(49): 40909-15.
- Gorbalenya, A. E. i Koonin, E. V. (1993). "Helicases: amino acid sequence comparisons and structure-function relationships." Current Opinion in Structural Biology **3**(3): 419-429.
- Ha, T., Rasnik, I., Cheng, W., Babcock, H. P., Gauss, G. H., Lohman, T. M. i Chu, S. (2002). "Initiation and re-initiation of DNA unwinding by the Escherichia coli Rep helicase." Nature **419**(6907): 638-41.
- Hacker, K. J. i Johnson, K. A. (1997). "A hexameric helicase encircles one DNA strand and excludes the other during DNA unwinding." Biochemistry **36**(46): 14080-7.
- Hall, M. C. i Matson, S. W. (1999). "Helicase motifs: the engine that powers DNA unwinding." Mol Microbiol **34**(5): 867-77.
- Hamdan, S. M., Johnson, D. E., Tanner, N. A., Lee, J.-B., Qimron, U., Tabor, S., van Oijen, A. M. i Richardson, C. C. (2007). "Dynamic DNA Helicase-DNA Polymerase Interactions Assure Processive Replication Fork Movement." Molecular Cell **27**(4): 539-549.
- Heras, B. i Martin, J. L. (2005). "Post-crystallization treatments for improving diffraction quality of protein crystals." Acta Crystallogr D Biol Crystallogr **61**(Pt 9): 1173-80.
- Hickman, A. B. i Dyda, F. (2005). "Binding and unwinding: SF3 viral helicases." Curr Opin Struct Biol **15**(1): 77-85.
- Hickson, I. D. (2003). "RecQ helicases: caretakers of the genome." Nat Rev Cancer **3**(3): 169-178.

- Hingorani, M. M. i O'Donnell, M. (1998). "Toroidal proteins: running rings around DNA." Curr Biol **8**(3): R83-6.
- Hingorani, M. M., Washington, M. T., Moore, K. C. i Patel, S. S. (1997). "The dTTPase mechanism of T7 DNA helicase resembles the binding change mechanism of the F1-ATPase." Proc Natl Acad Sci U S A **94**(10): 5012-7.
- Hopfner, K. P. i Michaelis, J. (2007). "Mechanisms of nucleic acid translocases: lessons from structural biology and single-molecule biophysics." Curr Opin Struct Biol **17**(1): 87-95.
- Huang, S. G., Weisshart, K. i Fanning, E. (1998). "Characterization of the nucleotide binding properties of SV40 T antigen using fluorescent 3'(2')-O-(2,4,6-trinitrophenyl)adenine nucleotide analogues." Biochemistry **37**(44): 15336-44.
- Isupov, M. N., Brindley, A. A., Hollingsworth, E. J., Murshudov, G. N., Vagin, A. A. i Littlechild, J. A. (2004). "Crystallization and preliminary X-ray diffraction studies of a fungal hydrolase from *Ophiostoma novo-ulmi*." Acta Crystallogr D Biol Crystallogr **60**(Pt 10): 1879-82.
- Jacob, F. i Brenner, S. (1963). "[On the regulation of DNA synthesis in bacteria: the hypothesis of the replicon.]." C R Hebd Seances Acad Sci **256**: 298-300.
- Jankowsky, E. i Fairman, M. E. (2007). "RNA helicases--one fold for many functions." Curr Opin Struct Biol **17**(3): 316-24.
- Jeong, Y. J., Levin, M. K. i Patel, S. S. (2004). "The DNA-unwinding mechanism of the ring helicase of bacteriophage T7." Proc Natl Acad Sci U S A **101**(19): 7264-9.
- Jeruzalmi, D., O'Donnell, M. i Kuriyan, J. (2001a). "Crystal structure of the processivity clamp loader gamma (gamma) complex of *E. coli* DNA polymerase III." Cell **106**(4): 429-41.
- Jeruzalmi, D., Yurieva, O., Zhao, Y., Young, M., Stewart, J., Hingorani, M., O'Donnell, M. i Kuriyan, J. (2001b). "Mechanism of processivity clamp opening by the delta subunit wrench of the clamp loader complex of *E. coli* DNA polymerase III." Cell **106**(4): 417-28.
- Jezewska, M. J., Kim, U. S. i Bujalowski, W. (1996). "Binding of *Escherichia coli* primary replicative helicase DnaB protein to single-stranded DNA. Long-range allosteric conformational changes within the protein hexamer." Biochemistry **35**(7): 2129-45.
- Jezewska, M. J., Rajendran, S., Bujalowska, D. i Bujalowski, W. (1998a). "Does single-stranded DNA pass through the inner channel of the protein hexamer in the complex with the *Escherichia coli* DnaB Helicase? Fluorescence energy transfer studies." J Biol Chem **273**(17): 10515-29.
- Jezewska, M. J., Rajendran, S. i Bujalowski, W. (1997). "Strand specificity in the interactions of *Escherichia coli* primary replicative helicase DnaB protein with a replication fork." Biochemistry **36**(33): 10320-6.

- Jezewska, M. J., Rajendran, S. i Bujalowski, W. (1998b). "Complex of Escherichia coli primary replicative helicase DnaB protein with a replication fork: recognition and structure." Biochemistry **37**(9): 3116-36.
- Johnson, A. i O'Donnell, M. (2005). "Cellular DNA replicases: components and dynamics at the replication fork." Annu Rev Biochem **74**: 283-315.
- Johnson, D. S., Bai, L., Smith, B. Y., Patel, S. S. i Wang, M. D. (2007). "Single-molecule studies reveal dynamics of DNA unwinding by the ring-shaped T7 helicase." Cell **129**(7): 1299-309.
- Johnson, S. K., Bhattacharyya, S. i Griep, M. A. (2000). "DnaB helicase stimulates primer synthesis activity on short oligonucleotide templates." Biochemistry **39**(4): 736-44.
- Kabsch, W. (1993). "Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants." J. Appl. Cryst. **26**: 795-800.
- Kadare, G. i Haenni, A. L. (1997). "Virus-encoded RNA helicases." J Virol **71**(4): 2583-90.
- Kaguni, J. M. (2006). "DnaA: controlling the initiation of bacterial DNA replication and more." Annu Rev Microbiol **60**: 351-75.
- Kainov, D. E., Tuma, R. i Mancini, E. J. (2006). "Hexameric molecular motors: P4 packaging ATPase unravels the mechanism." Cell Mol Life Sci **63**(10): 1095-105.
- Kaplan, D. L. (2000). "The 3'-tail of a forked-duplex sterically determines whether one or two DNA strands pass through the central channel of a replication-fork helicase." J Mol Biol **301**(2): 285-99.
- Kaplan, D. L., Davey, M. J. i O'Donnell, M. (2003). "Mcm4,6,7 uses a "pump in ring" mechanism to unwind DNA by steric exclusion and actively translocate along a duplex." J Biol Chem **278**(49): 49171-82.
- Kaplan, D. L. i O'Donnell, M. (2002). "DnaB drives DNA branch migration and dislodges proteins while encircling two DNA strands." Mol Cell **10**(3): 647-57.
- Kaplan, D. L. i O'Donnell, M. (2004). "Twin DNA pumps of a hexameric helicase provide power to simultaneously melt two duplexes." Mol Cell **15**(3): 453-65.
- Kaplan, D. L. i Steitz, T. A. (1999). "DnaB from Thermus aquaticus unwinds forked duplex DNA with an asymmetric tail length dependence." J Biol Chem **274**(11): 6889-97.
- Kim, J. L., Morgenstern, K. A., Griffith, J. P., Dwyer, M. D., Thomson, J. A., Murcko, M. A., Lin, C. i Caron, P. R. (1998). "Hepatitis C virus NS3 RNA helicase domain with a bound oligonucleotide: the crystal structure provides insights into the mode of unwinding." Structure **6**(1): 89-100.

- Kim, S., Dallmann, H. G., McHenry, C. S. i Marians, K. J. (1996). "Coupling of a replicative polymerase and helicase: a tau-DnaB interaction mediates rapid replication fork movement." Cell **84**(4): 643-50.
- Kobori, J. A. i Kornberg, A. (1982). "The Escherichia coli dnaC gene product. III. Properties of the dnaB-dnaC protein complex." J Biol Chem **257**(22): 13770-5.
- Kong, X. P., Onrust, R., O'Donnell, M. i Kuriyan, J. (1992). "Three-dimensional structure of the beta subunit of E. coli DNA polymerase III holoenzyme: a sliding DNA clamp." Cell **69**(3): 425-37.
- Konieczny, I. (2003). "Strategies for helicase recruitment and loading in bacteria." EMBO Rep **4**(1): 37-41.
- Koonin, E. V. (1992). "DnaC protein contains a modified ATP-binding motif and belongs to a novel family of ATPases including also DnaA." Nucleic Acids Res **20**(8): 1997.
- Korhonen, J. A., Gaspari, M. i Falkenberg, M. (2003). "TWINKLE Has 5' -> 3' DNA helicase activity and is specifically stimulated by mitochondrial single-stranded DNA-binding protein." J Biol Chem **278**(49): 48627-32.
- Lamers, M. H., Georgescu, R. E., Lee, S. G., O'Donnell, M. i Kuriyan, J. (2006). "Crystal structure of the catalytic alpha subunit of E. coli replicative DNA polymerase III." Cell **126**(5): 881-92.
- Learn, B. A., Um, S. J., Huang, L. i McMacken, R. (1997). "Cryptic single-stranded-DNA binding activities of the phage lambda P and Escherichia coli DnaC replication initiation proteins facilitate the transfer of E. coli DnaB helicase onto DNA." Proc Natl Acad Sci U S A **94**(4): 1154-9.
- LeBowitz, J. H. i McMacken, R. (1986). "The Escherichia coli dnaB replication protein is a DNA helicase." J Biol Chem **261**(10): 4738-48.
- Lee, J. Y. i Yang, W. (2006). "UvrD helicase unwinds DNA one base pair at a time by a two-part power stroke." Cell **127**(7): 1349-60.
- Levin, M. K., Gurjar, M. i Patel, S. S. (2005). "A Brownian motor mechanism of translocation and strand separation by hepatitis C virus helicase." Nat Struct Mol Biol **12**(5): 429-35.
- Li, D., Zhao, R., Lilyestrom, W., Gai, D., Zhang, R., DeCaprio, J. A., Fanning, E., Jochimiak, A., Szakonyi, G. i Chen, X. S. (2003). "Structure of the replicative helicase of the oncoprotein SV40 large tumour antigen." Nature **423**(6939): 512-8.
- Liao, J. C., Jeong, Y. J., Kim, D. E., Patel, S. S. i Oster, G. (2005). "Mechanochemistry of t7 DNA helicase." J Mol Biol **350**(3): 452-75.
- Lionnet, T., Spiering, M. M., Benkovic, S. J., Bensimon, D. i Croquette, V. (2007). "Real-time observation of bacteriophage T4 gp41 helicase reveals an unwinding mechanism." Proc Natl Acad Sci U S A **104**(50): 19790-5.

- Lohman, T. M. (1993). "Helicase-catalyzed DNA unwinding." J Biol Chem **268**(4): 2269-72.
- Lohman, T. M. i Bjornson, K. P. (1996). "Mechanisms of helicase-catalyzed DNA unwinding." Annu Rev Biochem **65**: 169-214.
- Lohman, T. M., Tomko, E. J. i Wu, C. G. (2008). "Non-hexameric DNA helicases and translocases: mechanisms and regulation." Nat Rev Mol Cell Biol **9**(5): 391-401.
- Lu, Y. B., Ratnakar, P. V., Mohanty, B. K. i Bastia, D. (1996). "Direct physical interaction between DnaG primase and DnaB helicase of Escherichia coli is necessary for optimal synthesis of primer RNA." Proc Natl Acad Sci U S A **93**(23): 12902-7.
- Ludlam, A. V., McNatt, M. W., Carr, K. M. i Kaguni, J. M. (2001). "Essential amino acids of Escherichia coli DnaC protein in an N-terminal domain interact with DnaB helicase." J Biol Chem **276**(29): 27345-53.
- Lusty, C. (1999). "A gentle vapor-diffusion technique for cross-linking of protein crystals for cryocrystallography." Journal of Applied Crystallography **32**(1): 106-112.
- Mackintosh, S. G. i Raney, K. D. (2006). "DNA unwinding and protein displacement by superfamily 1 and superfamily 2 helicases." Nucleic Acids Res **34**(15): 4154-9.
- Maluf, N. K., Ali, J. A. i Lohman, T. M. (2003). "Kinetic mechanism for formation of the active, dimeric UvrD helicase-DNA complex." J Biol Chem **278**(34): 31930-40.
- Marians, K. J. (2008). "Understanding how the replisome works." Nat Struct Mol Biol **15**(2): 125-7.
- Marszalek, J. i Kaguni, J. M. (1994). "DnaA protein directs the binding of DnaB protein in initiation of DNA replication in Escherichia coli." J Biol Chem **269**(7): 4883-90.
- Marszalek, J., Zhang, W., Hupp, T. R., Margulies, C., Carr, K. M., Cherry, S. i Kaguni, J. M. (1996). "Domains of DnaA protein involved in interaction with DnaB protein, and in unwinding the Escherichia coli chromosomal origin." J Biol Chem **271**(31): 18535-42.
- Matias, P. M., Gorynia, S., Donner, P. i Carrondo, M. A. (2006). "Crystal structure of the human AAA+ protein RuvBL1." J Biol Chem **281**(50): 38918-29.
- Matson, S. W., Bean, D. W. i George, J. W. (1994). "DNA helicases: enzymes with essential roles in all aspects of DNA metabolism." BioEssays **16**(1): 13-22.
- Matthews, B. W. (1968). "Solvent content of protein crystals." J Mol Biol **33**(2): 491-7.
- Mayans, O., van der Ven, P. F., Wilm, M., Mues, A., Young, P., Furst, D. O., Wilmanns, M. i Gautel, M. (1998). "Structural basis for activation of the titin kinase domain during myofibrillogenesis." Nature **395**(6705): 863-9.

- McCoy, A., Grosse-Kunstleve, R., Adams, P., Winn, M., Storoni, L. i Read, R. (2007). "Phaser crystallographic software." J. Appl. Cryst. **40**: 658-674
- McHenry, C. S. (2003). "Chromosomal replicases as asymmetric dimers: studies of subunit arrangement and functional consequences." Mol Microbiol **49**(5): 1157-65.
- Messer, W., Blaesing, F., Jakimowicz, D., Krause, M., Majka, J., Nardmann, J., Schaper, S., Seitz, H., Speck, C., Weigel, C., Wegrzyn, G., Welzeck, M. i Zakrzewska-Czerwinska, J. (2001). "Bacterial replication initiator DnaA. Rules for DnaA binding and roles of DnaA in origin unwinding and helicase loading." Biochimie **83**(1): 5-12.
- Miles, C. S., Weigelt, J., Stamford, N. P., Dammerova, N., Otting, G. i Dixon, N. E. (1997). "Precise limits of the N-terminal domain of DnaB helicase determined by NMR spectroscopy." Biochem Biophys Res Commun **231**(1): 126-30.
- Miller, R., Gallo, S. M., Khalak, H. G. i Weeks, C. M. (1994). "SnB: crystal structure determination via shake-and-bake." J. Appl. Cryst. **27**: 613-621.
- Mitkova, A. V., Khopde, S. M. i Biswas, S. B. (2003). "Mechanism and stoichiometry of interaction of DnaG primase with DnaB helicase of Escherichia coli in RNA primer synthesis." J Biol Chem **278**(52): 52253-61.
- Mohaghegh, P. i Hickson, I. D. (2001). "DNA helicase deficiencies associated with cancer predisposition and premature ageing disorders." Hum Mol Genet **10**(7): 741-6.
- Mott, M. L. i Berger, J. M. (2007). "DNA replication initiation: mechanisms and regulation in bacteria." Nat Rev Microbiol **5**(5): 343-54.
- Mott, M. L., Erzberger, J. P., Coons, M. M. i Berger, J. M. (2008). "Structural synergy and molecular crosstalk between bacterial helicase loaders and replication initiators." Cell **135**(4): 623-34.
- Nakayama, N., Arai, N., Kaziro, Y. i Arai, K. (1984). "Structural and functional studies of the dnaB protein using limited proteolysis. Characterization of domains for DNA-dependent ATP hydrolysis and for protein association in the primosome." J Biol Chem **259**(1): 88-96.
- Nakayama, N., Bond, M. W., Miyajima, A., Kobori, J. i Arai, K. (1987). "Structure of Escherichia coli dnaC. Identification of a cysteine residue possibly involved in association with dnaB protein." J Biol Chem **262**(22): 10475-80.
- Navaza, J. (1994). "AMoRe: a molecular replacement package." Acta Crystallogr A **50**: 157-163.
- Nonin, S., Leroy, J. L. i Gueron, M. (1995). "Terminal base pairs of oligodeoxynucleotides: imino proton exchange and fraying." Biochemistry **34**(33): 10652-9.

- Nunez-Ramirez, R., Robledo, Y., Mesa, P., Ayora, S., Alonso, J. C., Carazo, J. M. i Donate, L. E. (2006). "Quaternary polymorphism of replicative helicase G40P: structural mapping and domain rearrangement." J Mol Biol **357**(4): 1063-76.
- Nunez-Ramirez, R., Velten, M., Rivas, G., Polard, P., Carazo, J. M. i Donate, L. E. (2007). "Loading a ring: structure of the Bacillus subtilis DnaB protein, a co-loader of the replicative helicase." J Mol Biol **367**(3): 764-9.
- Ozaki, S., Kawakami, H., Nakamura, K., Fujikawa, N., Kagawa, W., Park, S. Y., Yokoyama, S., Kurumizaka, H. i Katayama, T. (2008). "A common mechanism for the ATP-DnaA-dependent formation of open complexes at the replication origin." J Biol Chem **283**(13): 8351-62.
- Patel, S. S. i Donmez, I. (2006). "Mechanisms of helicases." J Biol Chem **281**(27): 18265-8.
- Patel, S. S. i Picha, K. M. (2000). "Structure and function of hexameric helicases." Annu Rev Biochem **69**: 651-97.
- Pomerantz, R. T. i O'Donnell, M. (2007). "Replisome mechanics: insights into a twin DNA polymerase machine." Trends Microbiol **15**(4): 156-64.
- Pyle, A. M. (2008). "Translocation and unwinding mechanisms of RNA and DNA helicases." Annu Rev Biophys **37**: 317-36.
- Rasnik, I., Myong, S. i Ha, T. (2006). "Unraveling helicase mechanisms one molecule at a time." Nucleic Acids Res **34**(15): 4225-31.
- Rowen, L. i Kornberg, A. (1978). "Primase, the dnaG protein of Escherichia coli. An enzyme which starts DNA chains." J Biol Chem **253**(3): 758-64.
- Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. i Erlich, H. A. (1988). "Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase." Science **239**(4839): 487-91.
- San Martin, C., Radermacher, M., Wolpensinger, B., Engel, A., Miles, C. S., Dixon, N. E. i Carazo, J. M. (1998). "Three-dimensional reconstructions from cryoelectron microscopy images reveal an intimate complex between helicase DnaB and its loading partner DnaC." Structure **6**(4): 501-9.
- San Martin, M. C., Stamford, N. P., Dammerova, N., Dixon, N. E. i Carazo, J. M. (1995). "A structural model for the Escherichia coli DnaB helicase based on electron microscopy data." J Struct Biol **114**(3): 167-76.
- Saridakis, E. i Chayen, N. E. (2000). "Improving protein crystal quality by decoupling nucleation and growth in vapor diffusion." Protein Sci **9**(4): 755-7.
- Saridakis, E. i Chayen, N. E. (2003). "Systematic improvement of protein crystals by determining the supersolubility curves of phase diagrams." Biophys J **84**(2 Pt 1): 1218-22.

- Schneider, T. R. i Sheldrick, G. M. (2002). "Substructure solution with SHELXD." Acta Crystallogr D Biol Crystallogr **58**.
- Schubot, F. D., Cherry, S., Austin, B. P., Tropea, J. E. i Waugh, D. S. (2005). "Crystal structure of the protease-resistant core domain of Yersinia pestis virulence factor YopR." Protein Sci **14**(6): 1679-83.
- Schuck, S. i Stenlund, A. (2005). "Assembly of a double hexameric helicase." Mol Cell **20**(3): 377-89.
- Seitz, H., Weigel, C. i Messer, W. (2000). "The interaction domains of the DnaA and DnaB replication proteins of Escherichia coli." Mol Microbiol **37**(5): 1270-9.
- Sheldrick, G. M. (2008). "A short history of SHELX." Acta Crystallogr A **64**(Pt 1): 112-22.
- Shereda, R. D., Kozlov, A. G., Lohman, T. M., Cox, M. M. i Keck, J. L. (2008). "SSB as an organizer/mobilizer of genome maintenance complexes." Crit Rev Biochem Mol Biol **43**(5): 289-318.
- Singleton, M. R., Dillingham, M. S. i Wigley, D. B. (2007). "Structure and mechanism of helicases and nucleic acid translocases." Annu Rev Biochem **76**: 23-50.
- Singleton, M. R., Sawaya, M. R., Ellenberger, T. i Wigley, D. B. (2000). "Crystal Structure of T7 Gene 4 Ring Helicase Indicates a Mechanism for Sequential Hydrolysis of Nucleotides." Cell **101**(6): 589-600.
- Skordalakes, E. i Berger, J. M. (2003). "Structure of the Rho transcription terminator: mechanism of mRNA recognition and helicase loading." Cell **114**(1): 135-46.
- Skordalakes, E. i Berger, J. M. (2006). "Structural insights into RNA-dependent ring closure and ATPase activation by the Rho termination factor." Cell **127**(3): 553-64.
- Soultanas, P., Dillingham, M. S., Wiley, P., Webb, M. R. i Wigley, D. B. (2000a). "Uncoupling DNA translocation and helicase activity in PcrA: direct evidence for an active mechanism." Embo J **19**(14): 3799-810.
- Soultanas, P. i Wigley, D. B. (2000b). "DNA helicases: 'inching forward'." Curr Opin Struct Biol **10**(1): 124-8.
- Speck, C. i Messer, W. (2001). "Mechanism of origin unwinding: sequential binding of DnaA to double- and single-stranded DNA." Embo J **20**(6): 1469-76.
- Spelbrink, J. N., Li, F. Y., Tiranti, V., Nikali, K., Yuan, Q. P., Tariq, M., Wanrooij, S., Garrido, N., Comi, G., Morandi, L., Santoro, L., Toscano, A., Fabrizi, G. M., Somer, H., Croxen, R., Beeson, D., Poulton, J., Suomalainen, A., Jacobs, H. T., Zeviani, M. i Larsson, C. (2001). "Human mitochondrial DNA deletions associated with mutations in the gene encoding Twinkle, a phage T7 gene 4-like protein localized in mitochondria." Nat Genet **28**(3): 223-31.

- Stitt, B. L. i Xu, Y. (1998). "Sequential hydrolysis of ATP molecules bound in interacting catalytic sites of Escherichia coli transcription termination protein Rho." J Biol Chem **273**(41): 26477-86.
- Studier, F. W. (1991). "Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system." J Mol Biol **219**(1): 37-44.
- Stukenberg, P. T., Studwell-Vaughan, P. S. i O'Donnell, M. (1991). "Mechanism of the sliding beta-clamp of DNA polymerase III holoenzyme." J Biol Chem **266**(17): 11328-34.
- Su'etsugu, M., Shimuta, T. R., Ishida, T., Kawakami, H. i Katayama, T. (2005). "Protein associations in DnaA-ATP hydrolysis mediated by the Hda-replicase clamp complex." J Biol Chem **280**(8): 6528-36.
- Su'etsugu, M., Takata, M., Kubota, T., Matsuda, Y. i Katayama, T. (2004). "Molecular mechanism of DNA replication-coupled inactivation of the initiator protein in Escherichia coli: interaction of DnaA with the sliding clamp-loaded DNA and the sliding clamp-Hda complex." Genes Cells **9**(6): 509-22.
- Sun, P. D., Radaev, S. i Kattah, M. (2002). "Generating isomorphous heavy-atom derivatives by a quick-soak method. Part I: test cases." Acta Crystallogr D Biol Crystallogr **58**(Pt 7): 1092-8.
- Sutton, M. D., Carr, K. M., Vicente, M. i Kaguni, J. M. (1998). "Escherichia coli DnaA protein. The N-terminal domain and loading of DnaB helicase at the E. coli chromosomal origin." J Biol Chem **273**(51): 34255-62.
- Tackett, A. J., Chen, Y., Cameron, C. E. i Raney, K. D. (2005). "Multiple full-length NS3 molecules are required for optimal unwinding of oligonucleotide DNA in vitro." J Biol Chem **280**(11): 10797-806.
- Takahashi, T. S., Wigley, D. B. i Walter, J. C. (2005). "Pumps, paradoxes and ploughshares: mechanism of the MCM2-7 DNA helicase." Trends Biochem Sci **30**(8): 437-44.
- Tanner, N. A., Hamdan, S. M., Jergic, S., Schaeffer, P. M., Dixon, N. E. i van Oijen, A. M. (2008). "Single-molecule studies of fork dynamics in Escherichia coli DNA replication." Nat Struct Mol Biol **15**(2): 170-6.
- Terwilliger, T. C. (2000). "Maximum likelihood density modification." Acta Crystallogr D **56**: 965-972.
- Toth, E. A., Li, Y., Sawaya, M. R., Cheng, Y. i Ellenberger, T. (2003). "The crystal structure of the bifunctional primase-helicase of bacteriophage T7." Mol Cell **12**(5): 1113-23.
- Tougu, K. i Marians, K. J. (1996a). "The extreme C terminus of primase is required for interaction with DnaB at the replication fork." J Biol Chem **271**(35): 21391-7.
- Tougu, K. i Marians, K. J. (1996b). "The interaction between helicase and primase sets the replication fork clock." J Biol Chem **271**(35): 21398-405.

- Tougu, K., Peng, H. i Marians, K. J. (1994). "Identification of a domain of Escherichia coli primase required for functional interaction with the DnaB helicase at the replication fork." J Biol Chem **269**(6): 4675-82.
- Ueda, K., McMacken, R. i Kornberg, A. (1978). "dnaB protein of Escherichia coli. Purification and role in the replication of phiX174 DNA." J Biol Chem **253**(1): 261-9.
- Uson, I. i Sheldrick, G. M. (1999). "Advances in direct methods for protein crystallography." 9(643-648).
- Vagin, A. i Teplyakov, A. (1997). "MOLREP: an Automated Program for Molecular Replacement." J. Appl. Cryst. **30**: 1022-1025.
- Valle, M., Gruss, C., Halmer, L., Carazo, J. M. i Donate, L. E. (2000). "Large T-antigen double hexamers imaged at the simian virus 40 origin of replication." Mol Cell Biol **20**(1): 34-41.
- van Brabant, A. J., Stan, R. i Ellis, N. A. (2000). "DNA helicases, genomic instability, and human genetic disease." Annu Rev Genomics Hum Genet **1**: 409-59.
- Velankar, S. S., Soultanas, P., Dillingham, M. S., Subramanya, H. S. i Wigley, D. B. (1999). "Crystal structures of complexes of PcrA DNA helicase with a DNA substrate indicate an inchworm mechanism." Cell **97**(1): 75-84.
- Venteicher, A. S., Meng, Z., Mason, P. J., Veenstra, T. D. i Artandi, S. E. (2008). "Identification of ATPases pontin and reptin as telomerase components essential for holoenzyme assembly." Cell **132**(6): 945-57.
- von Hippel, P. H. i Delagoutte, E. (2001). "A general model for nucleic acid helicases and their "coupling" within macromolecular machines." Cell **104**(2): 177-90.
- Wahle, E., Lasken, R. S. i Kornberg, A. (1989a). "The dnaB-dnaC replication protein complex of Escherichia coli. I. Formation and properties." J Biol Chem **264**(5): 2463-8.
- Wahle, E., Lasken, R. S. i Kornberg, A. (1989b). "The dnaB-dnaC replication protein complex of Escherichia coli. II. Role of the complex in mobilizing dnaB functions." J Biol Chem **264**(5): 2469-75.
- Walter, T. S., Meier, C., Assenberg, R., Au, K. F., Ren, J., Verma, A., Nettleship, J. E., Owens, R. J., Stuart, D. I. i Grimes, J. M. (2006). "Lysine methylation as a routine rescue strategy for protein crystallization." Structure **14**(11): 1617-22.
- Wang, G., Klein, M. G., Tokonzaba, E., Zhang, Y., Holden, L. G. i Chen, X. S. (2008). "The structure of a DnaB-family replicative helicase and its interactions with primase." Nat Struct Mol Biol **15**(1): 94-100.
- Watt, P. M. i Hickson, I. D. (1996). "Failure to unwind causes cancer. Genome stability." Curr Biol **6**(3): 265-7.

- Wechsler, J. A. i Gross, J. D. (1971). "Escherichia coli mutants temperature-sensitive for DNA synthesis." Mol Gen Genet **113**(3): 273-84.
- Weigel, C. i Seitz, H. (2002). "Strand-specific loading of DnaB helicase by DnaA to a substrate mimicking unwound oriC." Mol Microbiol **46**(4): 1149-56.
- Weigelt, J., Brown, S. E., Miles, C. S., Dixon, N. E. i Otting, G. (1999). "NMR structure of the N-terminal domain of E. coli DnaB helicase: implications for structure rearrangements in the helicase hexamer." Structure **7**(6): 681-90.
- Wickner, S., Wright, M. i Hurwitz, J. (1974). "Association of DNA-dependent and -independent ribonucleoside triphosphatase activities with dnaB gene product of Escherichia coli." Proc Natl Acad Sci U S A **71**(3): 783-7.
- Wing, R. A., Bailey, S. i Steitz, T. A. (2008). "Insights into the replisome from the structure of a ternary complex of the DNA polymerase III alpha-subunit." J Mol Biol **382**(4): 859-69.
- Wong, I. i Lohman, T. M. (1992). "Allosteric effects of nucleotide cofactors on Escherichia coli Rep helicase-DNA binding." Science **256**(5055): 350-5.
- Wu, C. A., Zechner, E. L. i Mariani, K. J. (1992). "Coordinated leading- and lagging-strand synthesis at the Escherichia coli DNA replication fork. I. Multiple effectors act to modulate Okazaki fragment size." J Biol Chem **267**(6): 4030-44.
- Xiao, H., Dong, Z. i O'Donnell, M. (1993). "DNA polymerase III accessory proteins. IV. Characterization of chi and psi." J Biol Chem **268**(16): 11779-84.
- Yang, S., Yu, X., VanLoock, M. S., Jezewska, M. J., Bujalowski, W. i Egelman, E. H. (2002). "Flexibility of the rings: structural asymmetry in the DnaB hexameric helicase." J Mol Biol **321**(5): 839-49.
- Yao, J. X. (2002). "ACORN in CCP4 and its applications." Acta Crystallogr D **58**: 1941-1947.
- Yarranton, G. T. i Gefter, M. L. (1979). "Enzyme-catalyzed DNA unwinding: studies on Escherichia coli rep protein." Proc Natl Acad Sci U S A **76**(4): 1658-62.
- Yu, X., Hingorani, M. M., Patel, S. S. i Egelman, E. H. (1996a). "DNA is bound within the central hole to one or two of the six subunits of the T7 DNA helicase." Nat Struct Biol **3**(9): 740-3.
- Yu, X., Jezewska, M. J., Bujalowski, W. i Egelman, E. H. (1996b). "The hexameric E. coli DnaB helicase can exist in different Quaternary states." J Mol Biol **259**(1): 7-14.
- Yuzhakov, A., Kelman, Z. i O'Donnell, M. (1999). "Trading places on DNA--a three-point switch underlies primer handoff from primase to the replicative DNA polymerase." Cell **96**(1): 153-63.