

ESTUDIS CINÈTICS I ESTRUCTURALS DE LA TRANS INTEÏNA *Npu DnaE* DE “*Nostoc punctiforme*”

David Soler Campmajó

Dipòsit legal: Gi. 287-2015

<http://hdl.handle.net/10803/285784>

ADVERTIMENT. L'accés als continguts d'aquesta tesi doctoral i la seva utilització ha de respectar els drets de la persona autora. Pot ser utilitzada per a consulta o estudi personal, així com en activitats o materials d'investigació i docència en els termes establerts a l'art. 32 del Text Refós de la Llei de Propietat Intel·lectual (RDL 1/1996). Per altres utilitzacions es requereix l'autorització prèvia i expressa de la persona autora. En qualsevol cas, en la utilització dels seus continguts caldrà indicar de forma clara el nom i cognoms de la persona autora i el títol de la tesi doctoral. No s'autoritza la seva reproducció o altres formes d'explotació efectuades amb finalitats de lucre ni la seva comunicació pública des d'un lloc aliè al servei TDX. Tampoc s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX (framing). Aquesta reserva de drets afecta tant als continguts de la tesi com als seus resums i índexs.

ADVERTENCIA. El acceso a los contenidos de esta tesis doctoral y su utilización debe respetar los derechos de la persona autora. Puede ser utilizada para consulta o estudio personal, así como en actividades o materiales de investigación y docencia en los términos establecidos en el art. 32 del Texto Refundido de la Ley de Propiedad Intelectual (RDL 1/1996). Para otros usos se requiere la autorización previa y expresa de la persona autora. En cualquier caso, en la utilización de sus contenidos se deberá indicar de forma clara el nombre y apellidos de la persona autora y el título de la tesis doctoral. No se autoriza su reproducción u otras formas de explotación efectuadas con fines lucrativos ni su comunicación pública desde un sitio ajeno al servicio TDR. Tampoco se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR (framing). Esta reserva de derechos afecta tanto al contenido de la tesis como a sus resúmenes e índices.

WARNING. Access to the contents of this doctoral thesis and its use must respect the rights of the author. It can be used for reference or private study, as well as research and learning activities or materials in the terms established by the 32nd article of the Spanish Consolidated Copyright Act (RDL 1/1996). Express and previous authorization of the author is required for any other uses. In any case, when using its content, full name of the author and title of the thesis must be clearly indicated. Reproduction or other forms of for profit use or public communication from outside TDX service is not allowed. Presentation of its content in a window or frame external to TDX (framing) is not authorized either. These rights affect both the content of the thesis and its abstracts and indexes.



TESI DOCTORAL

ESTUDIS CINÈTICS I ESTRUCTURALS DE LA TRANS INTEÏNA *Npu DnaE* DE *Nostoc punctiforme*

Memòria presentada per optar al títol de doctor per la Universitat de
Girona

David Soler Campmajó

2014

Programa de Doctorat en Ciències Experimentals i Sostenibilitat

Dirigida per

Dirigida per

Marc Ribó Panosa

Professor titular

de Bioquímica i Biologia Molecular

Maria Vilanova Brugués

Catedràtica

de Bioquímica i Biologia Molecular

Als meus pares

A la meva germana

A la Gemma

“Hi havia tot un món esperant a ser descobert. O redescobert. Les llegendes no naixien del no res. Sempre hi havia la llavor d’una gran veritat en tota llegenda, que floriria amb la comprensió i la voluntat de buscar el que realment estava al darrera de cada història.”

“El despertar de les ovelles”

AGRAÏMENTS

L'inici de tot aquest treball va començar tot just acabar la carrera, és a dir, a l'estiu del 2008 i culmina 6 anys després amb l'escriptura d'aquesta tesi. Durant tots aquests anys puc recordar un munt de vivències que es tradueixen bàsicament en molts bons records. En tot aquest temps hi ha hagut també mals moments, però que s'obliden fàcilment quan penso amb tants i tants bons moments que he viscut. Per aquest motiu necessito donar les gràcies a un munt de gent, això va per tots vosaltres!

Primer de tot vull donar les gràcies a la Maria, en Marc i en Toni per haver-me deixat formar part del seu grup de recerca i ensenyar-me en què consisteix tot aquest món. Vull agrair especialment a en Marc per haver-me guiat durant tots aquests anys donant-me bons consells i compartint la seva experiència en el camp de la recerca, com també, la confiança que ha dipositat en mi per dur a terme aquest treball.

Vull agrair també a tots els becaris que he conegut del 107, tant els que m'han ensenyat molt per la seva experiència que ja tenien quan jo vaig començar com són en Gerard, en Pere, la Jess, la Mariona, en Roger i l'Anna com pels que han vingut darrere meu, la Vero, la Glòria i en Santi i que ens hem pogut ajudar mutuament en tots els entrebancs i dubtes que sorgeixen en el nostre camp de treball. Vull aprofitar aquest paràgraf per agrair també els bons moments que he passat amb les "veïnes" la Montse, la Cristina, la Clara, la Dolors, la Sònia, la Txell i la Marta. Per altra banda, aprofito també unes línies per agrair a tots els becaris del departament de Biologia que he anat coneixent durant aquests anys i que he anat compartint molt bones experiències. Agraeixo també a l'Imma per estar sempre pendent de tot i procurar que en el laboratori no ens falti mai de res.

Sortint de la UdG, vull donar les gràcies a la Marta i en Douglas de l'institut de química-física Rocasolano del CSIC de Madrid. Vosaltres em vau ajudar durant 4 mesos a entendre una mica el complicadíssim món de la ressonància magnètica nuclear, com també a fer que l'estada fos més agradable per la vostra amabilitat, simpatia i per les bones recomenacions de llocs a visitar i restaurants on menjar durant els caps de setmana.

Fora de l'àmbit de la recerca, necessito agrair també a en Vador, l'Edu, l'Anna i la Mireia que com a bons amics que sou, sempre us heu preocupat i interessat per mi tant quan les coses van bé com no tant bé. Espero poder continuar mantenint una bona relació amb vosaltres per molt que passin els anys i per molt que es vagin separant els nostres camins.

I la família? Crec que són les persones a qui més he d'agrair ja que tot això ha sigut possible gràcies als meus pares i la meva germana. Moltíssimes gràcies per la confiança que sempre teniu amb mi i per valorar-me sempre molt per damunt del que jo em valoro. Gràcies de tot cor! Vull acabar aquest paràgraf amb un record als avis, que allà on sigueu, estic ben segur que us sentiu molt orgullosos de veure'm culminar una etapa molt important en la meua vida.

Finalment, els meus agraïments més sincers a la Gemma. Gràcies per compartir amb mi totes les vivències bones i dolentes. Per comprendre'm, valorar-me i dipositar una enorme confiança amb mi. Has sigut capaç d'aguantar tots els meus moments de mal humor i de baixa moral amb un somriure i uns "*mimos*" que tant m'han ajudat a tirar endavant. Moltíssimes gràcies per tot!

Aquesta tesi ha estat realitzada amb el suport del Ministerio de Educación y Ciencia (beca AP2009-2817 i projecte BFU2006-15543-CO2-02/BMC) i del Ministerio de Ciencia e Innovación (projecte BFU2009-06935/BMC).

ÍNDIX FIGURES

Fig. 1.1. Comparació entre el mecanisme de tall i unió proteic i el de RNA	28
Fig. 1.2. Tipus d'inteiines trobades a la natura	29
Fig. 1.3. Estructura esquemàtica de les inteïnes	30
Fig. 1.4. Sistema de numeració dels aminoàcids	31
Fig. 1.5. Mecanisme de tall i unió	32
Fig. 1.6. Tall i unió de proteïnes i de les possibles reaccions secundàries	35
Fig. 1.7. Mecanisme de tall i unió de les tres classes d'inteiines	36
Fig. 1.8. Mecanisme de tall i unió en trans	38
Fig. 1.9. Generació de proteïnes cícliques	40
Fig. 1.10. Expressió de pèptids mitjançant inteïnes	41
Fig. 1.11. Procés d'obtenció de proteïnes fusionades a partir del mètode EPL	42
Fig. 1.12. Marcatge isotòpic de segments de proteïnes multi domini <i>in vivo</i>	44
Fig. 1.13. Procés de regulació de l'activitat proteica a partir del mecanisme de tall i unió condicionat	44
Fig. 1.14. Utilització d'una inteïna com a biosensor	45
Fig. 1.15. Seqüències aminoacídiques de les subunitats I _N i I _C de la <i>Npu DnaE</i>	47
Fig. 1.16. Estructura tridimensional de la <i>Npu DnaE</i>	49
Fig. 1.17. Especificitat en l'intercanvi de càrregues en l'activitat de <i>Npu DnaE</i>	52
Fig. 1.18. Mecanisme de captura i col·lapse proposat per Muir i col·laboradors	53
Fig. 4.1. Estratègia d'obtenció del vector pRSET-GFP _N - <i>Npu DnaE</i> -GFP _C	93
Fig. 4.2. SDS-PAGE per tal de determinar si <i>Npu DnaE</i> es localitzava a la fracció soluble o a la insoluble	94
Fig. 4.3. Perfil cromatogràfic de l'elució de <i>Npu DnaE</i>	95
Fig. 4.4. Mecanisme de tall i unió per l'obtenció de cHPR i de I _N i I _C associades	96
Fig. 4.5. Perfil cromatogràfic de l'elució de I _N i I _C associades	96
Fig. 4.6. Estratègia d'obtenció del vector pBADm41-I _N	98
Fig. 4.7. SDS-PAGE de les proves de tall i unió entre la construcció MBP-GB1-I _N	98
Fig. 4.8. Perfil cromatogràfic de l'elució de la subunitat I _N (cromatografia bescanvi aniònic)	100
Fig. 4.9. Perfil cromatogràfic de l'elució de la subunitat I _N (cromatografia gel filtració)	100
Fig. 4.10. Corbes normalitzades de desplegament i replegament per temperatura de <i>Npu DnaE</i>	101
Fig. 4.11. SDS-PAGE de les proves de tall i unió <i>in vitro</i> a 40, 50 i 60°C a diferents temps de GFP _N - <i>Npu DnaE</i> -GFP _C amb la mutació C+1S	105
Fig. 4.12. Western blots de les proves de tall i unió <i>in vitro</i> a 40, 50 i 60°C a 1h i 24h de GFP _N - <i>Npu DnaE</i> -GFP _C amb la mutació C+1S	106
Fig. 4.13. Formació del producte de tall i unió <i>in vitro</i> al llarg del temps a diferents temperatures amb la mutació C+1S	106
Fig. 4.14. Gràfics de les proves de tall i unió <i>in vitro</i> a 20, 30, 40, 50 i 60°C a diferents temps	107
Fig. 4.15. Gràfic on es mostra per cada temperatura el percentatge a les 24h del producte de tall i unió, C-escissió i N-escissió	108
Fig. 4.16. SDS-PAGE de les proves de tall i unió <i>in vitro</i> a 50 i 60°C a diferents temps de GFP _N - <i>Npu DnaE</i> -GFP _C amb la mutació C+1S amb i sense DTT	109
Fig. 4.17. Gràfic on es mostra per cada temperatura el percentatge a les 24h del producte de tall i unió de les proves realitzades a 0 i 1mM de DTT	110
Fig. 4.18. Esquema i gels SDS-PAGE de les proves de tall i unió <i>in vitro</i> a 20, 30, 40, 50 i 60°C a diferents temps de GB1-I _N I _C -GB1 i GB1-I _N I _C -GB1 (C+1S)	112
Fig. 4.19. Western blots de les proves de tall i unió <i>in vitro</i> a 30, 40, 50 i 60°C a 30min, 2h, 6h i 24h de GB1-I _N I _C -GB1 (C+1S)	113
Fig. 4.20. Formació del producte de tall i unió <i>in vitro</i> al llarg del temps a diferents temperatures de GB1-I _N I _C -GB1 C+1S	113
Fig. 4.21. Western blots de les proves de tall i unió <i>in vitro</i> a 20, 30, 40, 50 i 60°C a, 6h i 24h	116
Fig. 4.22. Elements d'estructura secundària de <i>Npu DnaE</i>	118
Fig. 4.23. Proteïnes de fusió I _N I _C _sc construïdes per a la realització d'aquest treball	119
Fig. 4.24. Construcció d el vector pBADm41	119
Fig. 4.25. Perfils cromatogràfics de l'elució de les diferents variants de I _N I _C _sc	120

Fig. 4.26. Gràfic de càrrega-hidrofobicitat comparatiu de <i>Npu DnaE</i> , les diferents variants de $I_{N C_sc}$, la subunitat I_N i la subunitat I_C	121
Fig. 4.27. Anàlisi CDF de les tres variants $I_{N C_sc}$	122
Fig. 4.28. Anàlisi on es compara la composició relativa d'aminoàcids de les diferents variants $I_{N C_sc}$	123
Fig. 4.29. Espectres de dicroisme circular normalitzats a l'ultraviolat llunyà	124
Fig. 4.30. Espectres monodimensionals de 1H de la subunitat I_N i de la inteïna <i>Npu DnaE</i>	127
Fig. 4.31. Espectres monodimensionals de 1H de les variants $I_{N C_sc114}$, $I_{N C_sc123}$ i $I_{N C_sc131}$	128
Fig. 4.32. Diferències entre els valors de desplaçament conformacional dels protons α de les variants $I_{N C_sc131}$, $I_{N C_sc123}$ i la inteïna <i>Npu DnaE</i>	130
Fig. 4.33. Espectres bidimensionals de 1H de $I_{N C_sc123}$	131
Fig. 4.34. Mapejat de les variants $I_{N C_sc123}$ i $I_{N C_sc131}$	133
Fig. 4.35. NOE's a llarga distància identificats en les diferents cadenes β de les variants $I_{N C_sc123}$ i $I_{N C_sc131}$	135
Fig. 5.1. Ponts d'hidrogen que s'estableixen entre el residu 1 i els residus Thr 69 i Asp 118 en la inteïna <i>Npu DnaE</i>	144
Fig. 5.2. Grau d'identitat que presenten les seqüències de la inteïna <i>Ssp DnaE</i> i <i>Npu DnaE</i>	145
Fig. 5.3. Esquema del procés de tall i unió de la variant C+1S on s'observa que la Ser +1 es trobaria a l'exteïna al final del procés	147
Fig. 5.4. Estructura tridimensional de <i>Npu DnaE</i> on s'indiquen els parells iònics identificats per Shah i col·laboradors.	149
Fig. 5.5. Representació esquemàtica de l'estructura tridimensional de <i>Npu DnaE</i> i <i>Ssp DnaE</i>	151

ÍNDEX TAULES

Taula 1.1. Seqüències aminoacídiques consens de les diferents regions de les inteïnes	31
Taula 1.2. Inteïnes partides més utilitzades en el camp de la biologia molecular	47
Taula 4.1. Efecte del pH en el repliegament de les diferents variants de <i>Npu DnaE</i>	101
Taula 4.2. Efecte del pH en el valor de T_m en les diferents variants de <i>Npu DnaE</i>	102
Taula 4.3. Valors de la K_{obs} observada en les diferents reaccions de tall i unió <i>in vitro</i> de <i>Npu DnaE C+1S</i> en cis	108
Taula 4.4. Valors de $t_{1/2}$ en les diferents reaccions de tall i unió <i>in vitro</i> de <i>Npu DnaE C+1S</i> en cis	108
Taula 4.5. Taula on es mostren els valors-p obtinguts en el test estadístic t-student per veure si hi havia diferències entre la quantitat total de producte de tall i unió format amb presència o sense DTT	110
Taula 4.6. Valors de K_{obs} (min^{-1}) i valors de vida mitja $t_{1/2}$ (min) observats en les diferents reaccions de tall i unió <i>in vitro</i> de <i>Npu DnaE C+1S</i> en trans	114
Taula 4.7. Taula on es mostren els valors-p obtinguts en el test estadístic t-student per veure si hi havia diferències entre la quantitat total de producte de tall i unió format entre GB1-I _N I _C -GB1 C+1S i GB1-I _N I _C -GB1 (wt)	114
Taula 4.8. Rendiments de purificació obtinguts per cada una de les variants de I _N I _C -sc	121
Taula 4.9. Contingut en hèlix α i fulls β de les diferents proteïnes estudiades	125
Taula 4.10. NOE's de metils, HA-NH i NH-NH de mitja i llarga distància assignats	132
Taula 4.11. Relació entre les diferents cadenes β de la inteïna <i>Npu DnaE</i>	134
Taula 5.1. Inteïnes de bacteris termòfils on es mostra el residu que tenen en la posició 1 i +1	142
Taula 5.2. Trans inteïnes recollides a InBase	143

ABREVIATURES

A	adenina
Amp	ampicil·lina
C	citosa
CaCl ₂	clorur de calci
CCD	dispositiu de càrrega acoblada
CD	dicroisme circular
ChBD	domini d'unió a quitina
C-terminal	carboxi terminal
D ₂ O	aigua deuterada
Da, KDa	dalton, kilodalton
ddNTP	dideoxiribonucleòtid
Deg, mdeg	(mil·li)graus el·lipticitat
DNA	àcid desoxiribonucleic
DO	densitat òptica
DTT	1,4-ditio-DL-treitol
e	exponencial
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	àcid etilendiaminotetracètic
G	guanina
G, mg, µg	gram, mil·ligram, microgram
GB1	domini B1 de la proteïna G
GFP	<i>green fluorescent protein</i>
gpD	<i>head protein D</i>
H	hidrogen
h	hora
HEPES	àcid 2-(4-hidroxietil)-1-piperazinyloetanesulfonic
HPLC	cromatografia líquida d'alta resolució
I _C	inteïna C
I _N	inteïna N
IPTG	isopropil-β-D-tiogalactopiranòsid
Kan	kanamicina
Kd	constant de dissociació
l, ml, µl	litre, mil·lilitre, microlitre
LB	medi Lùria-Bertrani
M, mM, µM	molar, mil·limolar, micromolar
MALDI-TOF	matrix-assisted laser desorption/ionization time-of-flight mass spectrometry

MES	àcid 2-(N-morpholino)ethanesulfonic)
min	minut
Mw	massa molecular
N	nitrogen
NOE	nuclear Overhauser effect
nm	nanòmetre
N-terminal, N-t	amino terminal
o/n	<i>overnight</i> , tota la nit
°C	graus centígrads
p/p	relació pes pes
p/v	relació pes volum
PAGE	electroforesi en gel de poliacrilamida
pb	parells de bases
PCR	reacció en cadena de la polimerasa
PDB	banc de dades de proteïnes
pI	punt isoelèctric
pN	piconewton
ppm	parts per milió
Psi	lliura per polzada quadrada
RMN	ressonància magnètica nuclear
RNA	àcid ribonucleic
rpm	revolucions per minut
s	segon
SDS	dodecil fostaf sòdic
T	timina
t _{1/2}	temps de vida mitja
TAE	amortidor Tris/àcid acètic/EDTA
TCEP-HCl	tris (2-carboxietil) fosfina clorhidrat
TEMED	N,N,N',N'-tetrametiletildiamina
TEV	<i>tobacco etch virus</i>
Tm	temperatura de fusió
Tris	tris-hidroximetil-aminometà
U	unitat
UV	ultraviolat
v	volts
V	volum
v/v	relació volum volum
wt	<i>wild type</i> salvatge

CODI AMINOÀCIDS

Codi de tres lletres	Codi d'una lletra	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
Trp	W	Triptòfan
Tyr	Y	Tirosina
Val	V	Valina

ÍNDEX GENERAL

Resum / Resumen / Summary	19
1 INTRODUCCIÓ	25
1.1 Les inteïnes	27
1.1.1 Característiques generals	27
1.1.2 Descobriment	28
1.1.3 Organització	29
1.1.4 Nomenclatura	31
1.2 El mecanisme de tall i unió	32
1.2.1 El procés	32
1.2.2 Classificació	35
1.2.3 Alteracions en residus clau de la seqüència de la inteïna	37
1.2.4 El procés de tall i unió en trans	38
1.3 Aplicacions	40
1.3.1 Expressió i modificació de proteïnes	40
1.3.2 Inteïnes com a mediadores de l'etiquetatge de proteïnes	42
1.3.3 Regulació de l'activitat proteica a partir del mecanisme de tall i unió condicionat	44
1.3.4 Inteïnes com a biosensors	45
1.3.5 Requisits de les inteïnes per ser utilitzades com a eines biotecnològiques	46
1.4 La inteïna <i>Npu DnaE</i>	47
1.4.1 L'estructura tridimensional	48
1.4.2 Cinètica i afinitat d'unio de la inteïna	49
1.4.3 Robustesa de la inteïna	50
1.4.4 Promiscuïtat de la inteïna	51
1.4.5 El procés associatiu dels fragments I _N i I _C	52
2 OBJECTIUS	55
3 MATERIAL I MÈTODES	59
3.1 Material biològic	61
3.1.1 Soques bacterianes	61
3.1.2 Vectors procariotes	62
3.2 Mètodes microbiològics	63
3.2.1 Medis de cultiu	63
3.2.2 Antibiótics	64
3.2.3 Manteniment de les soques	65
3.2.4 Creixement dels cultius	65
3.3 Tècniques de DNA recombinant	66
3.3.1 Manipulació del DNA	66
3.3.2 Obtenció de DNA plasmídic	66
3.3.3 Electroforesi de DNA en gel d'agarosa	67
3.3.4 Purificació de fragments de DNA a partir de gels d'agarosa	67
3.3.5 Estimació espectrofotomètrica de concentracions de DNA	67
3.3.6 Utilització d'enzims que modifiquen el DNA	68
3.3.7 Mutagènesi dirigida per oligonucleòtid: Mètode "QuikChange™"	69
3.3.8 Obtenció de cèl·lules competents d' <i>E. coli</i>	70
3.3.9 Transformació de cèl·lules competents d' <i>E. coli</i>	71
3.3.10 Seqüenciació de DNA	71

3.4 Mètodes d'expressió i purificació de proteïnes recombinants	74
3.4.1 Expressió de la inteïna <i>Npu DnaE</i> , I _N – I _C , I _N -GB1 i I _C -GB1	74
3.4.2 Expressió de I _N , les variants I _N I _C _sc i de la proteasa TEV	75
3.4.3 Purificació de la <i>NpuDnaE</i> i de I _N -I _C	75
3.4.4 Purificació de I _N i de les variants I _N I _C _sc	76
3.4.5 Purificació de I _N -GB1 i I _C -GB1	78
3.4.6 Purificació de la proteasa TEV	79
3.5 Anàlisi i caracterització proteïca	79
3.5.1 Determinació de la massa molecular	79
3.5.2 Determinació de la quantitat de proteïna	79
3.5.3 Electroforesi discontinua en gel de poliacrilamida	80
3.5.4 Western blot	81
3.5.5 Determinació de l'estabilitat tèrmica	82
3.5.6 Estudi del procés de tall i unió	83
3.5.7 Càlcul de paràmetres cinètics	84
3.5.8 Anàlisi de l'estructura secundària per dicroisme circular	84
3.6 Espectroscòpia de ressonància magnètica nuclear (NMR)	86
3.6.1 Assignació de les variants I _N I _C _sc	86
3.7 Anàlisi bioinformàtica	86
3.8 Anàlisi estadística	87
4 RESULTATS I DISCUSSIÓ	89
4.1 Expressió, purificació i caracterització de l'estabilitat tèrmica de la inteïna <i>Npu DnaE</i>, de les subunitats I_N – I_C associades i de la I_N aïllada	91
4.1.1 Construcció, expressió i purificació de <i>Npu DnaE</i> formant una única cadena polipeptídica	91
4.1.2 Expressió i purificació de les subunitats I _N -I _C de <i>Npu DnaE</i> associades	
4.1.3 Construcció, expressió i purificació de la subunitat I _N de <i>Npu DnaE</i>	97
4.1.4 Caracterització de la inteïna <i>Npu DnaE</i> , de les subunitats I _N – I _C associades i de la I _N aïllada	100
4.2 Estudi del procés de tall i unió	103
4.2.1 Estudi del procés de tall i unió en cis de la <i>Npu DnaE</i> C+1S utilitzant la GFP com a exteïna	104
4.2.2 Estudi del procés de tall i unió en trans de la <i>Npu DnaE</i> C+1S utilitzant GB1 com a exteïna	110
4.2.3 Estudi del procés de tall i unió en trans de la <i>Npu DnaE</i> utilitzant GB1 com a exteïna amb la doble mutació C1S i C+1S	115
4.3 Estudi estructural mitjançant RMN del procés associatiu de la inteïna <i>Npu DnaE</i>	117
4.3.1 Construcció, expressió i purificació de les variants I _N I _C _sc	118
4.3.2 Anàlisi bioinformàtica i predicció del grau de plegament	121
4.3.3 Anàlisi de l'estructura secundària per dicroisme circular	124
4.3.4 Anàlisi de l'estructura per ressonància magnètica nuclear	125
5 DISCUSSIÓ GENERAL	137
6 CONCLUSIONS	153
7 BIBLIOGRAFIA	159

RESUM

Les inteïnes són dominis proteics interns que són capaços d'autoescindir-se de la cadena polipeptídica precursora i catalitzar la unió de les seqüències flanquejants (exteïnes) mitjançant un enllaç peptídic. La capacitat d'autoescindir-se és un procés post traduccional anomenat tall i unió. El procés és intramolecular, autocatalític i no requereix d'enzims auxiliars, cofactors ni energia metabòlica. Aquest mecanisme de tall i unió pot dur-se a terme també en trans. En aquest cas, la inteïna es troba partida en els fragments N i C terminals on cadascun d'ells porta unida la seva exteïna. Aquests dos fragments de la inteïna són capaços d'associar i lligar les exteïnes que porten unides donant lloc a una proteïna funcional. Entendre aquest mecanisme de tall i unió i com els dominis de les inteïnes en trans són capaces d'associar, ha permès el desenvolupament d'aplicacions basades en inteïnes com són, per exemple, la purificació de proteïnes, la ciclació o el marcatge isotòpic de dominis proteics entre moltes altres. Per tal que una inteïna pugui ser utilitzada en aquestes tècniques cal que aquesta presenti una bona termoestabilitat, catalitzi el mecanisme de tall i unió ràpidament, sigui soluble i la seva eficiència sigui independent de l'exteïna que porta unida. Fins a l'actualitat, la trans inteïna *Npu DnaE* és una de les inteïnes que compleix millor aquests requisits.

En aquest treball s'ha estudiat l'estabilitat tèrmica que presenta aquesta inteïna en un rang de diferents pHs tant quan aquesta es troba formant una única cadena polipeptídica, com partida en forma de fragments I_N i I_C i s'ha estudiat també l'estabilitat del fragment I_N aïllat. Els resultats han mostrat que en tots els casos l'estabilitat és més elevada a pHs lleugerament bàsics i, quan la inteïna es troba formant una única cadena polipeptídica, presenta una major estabilitat respecte quan es troba en forma dels dos fragments I_N i I_C . L'estabilitat és inferior encara quan el fragment I_N es troba aïllat. El replegament d'aquestes tres variants també segueix aquest patró on s'observa un major grau de plegament quan la inteïna forma una única cadena polipeptídica i és completament inapreciable en el cas del fragment I_N . Aquests resultats suggereixen que hi hauria una sinèrgia mútua entre els fragments I_N i I_C per tal que la inteïna adoptés la seva estructura tridimensional nativa.

Per altra banda, s'ha estudiat com es veu afectada l'activitat de la inteïna quan es substitueixen les Cys+1 i Cys1 per Ser, tant quan la inteïna forma una única cadena polipeptídica com quan es troba partida en els fragments I_N i I_C . En ambdós casos, la mutació C+1S bloqueja l'activitat d'aquesta a 20°C, però es reconstitueix l'activitat a mesura que s'incrementa la temperatura fins als 60°C. A partir de 50°C, les inteïnes mutants presenten una activitat igual a la inteïna salvatge. En el cas de la mutació C1S s'abolia completament l'activitat d'aquesta inteïna en tot el rang de temperatures. Igual que en el cas de la doble mutació C+1S i C1S. Per tant, la variant C+1S podria ser utilitzada com a eina biotecnològica ja que presenta un mecanisme de tall i unió condicionat per temperatura.

Finalment, en aquest treball s'ha volgut estudiar i complementar el que es coneix sobre el procés associatiu dels fragments I_N i I_C. Per dur a terme això, s'ha purificat i estudiat el plegament de tres variants de *Npu DnaE* que presenten diferents llargades del fragment I_C. El plegament s'ha analitzat mitjançant una anàlisi bioinformàtica prèvia, dicromisme circular i RMN. Els resultats obtinguts, mostren que malgrat el fragment I_N no presenta una estructura pròpia, afegint ja els primers 12 residus del fragment I_C, aquests són capaços de començar a reconstituir l'estructura global de la *Npu DnaE*. Aquesta reconstitució de l'estructura global és veu ja de manera significativa amb els 21 primers residus del fragment I_C. Confirmant d'aquesta manera que el plegament global d'aquesta inteïna ve condicionat per una sinèrgia de plegament entre els fragments I_N i I_C.

RESUMEN

Las inteínas son dominios proteicos que pueden autoescindirse de la cadena polipeptídica precursora y catalizar el empalme de las secuencias flanqueantes (exteínas) mediante un enlace peptídico. La capacidad de autoescindirse es un proceso post traduccional llamado corte y empalme. El proceso es intramolecular, autocatalítico y no requiere de enzimas auxiliares, cofactores ni de energía metabólica. El mecanismo de corte y empalme puede realizarse en trans. En este caso, la inteína se encuentra partida en los fragmentos N y C terminales donde cada uno de ellos lleva unida su exteína. Estos dos fragmentos de la inteína son capaces de asociar e empalmar las exteínas que llevan unidas obteniéndose así una proteína funcional. Entender este mecanismo y como los dominios de las inteínas en trans son capaces de asociar, ha permitido el desarrollo de aplicaciones basadas en inteínas como son por ejemplo, la purificación de proteínas, la ciclación o el marcaje isotópico de dominios proteicos entre muchas otras. Para que una inteína sea utilizada en esas técnicas debe de presentar una buena termoestabilidad, catalizar el corte y empalme de manera rápida, ser soluble y su eficiencia en el corte y empalme debe de ser independiente de la exteína que lleva unida. Hasta la actualidad, la trans inteína *Npu DnaE* es una de las que cumple mejor estos requisitos.

En este trabajo se ha estudiado la estabilidad térmica que presenta esta inteína en un rango de distintos pHs tanto cuando esta se encuentra formando una única cadena polipeptídica, como partida en forma de las subunidad I_N e I_C y se ha estudiado también la estabilidad de la subunidad I_N aislada. Los resultados han mostrado como en todos los casos la estabilidad es más elevada en valores de pHs ligeramente básicos y, la estabilidad es mayor cuando la inteína se encuentra formando una única cadena polipeptídica respecto cuando se encuentra en forma de las subunidades I_N e I_C . Esta estabilidad es aún inferior cuando el fragmento I_N se encuentra aislado. El replegamiento de estas tres variantes también sigue este patrón donde se observa un mejor grado de replegamiento cuando la inteína forma una única cadena polipeptídica y es completamente inapreciable en el caso del fragmento I_N aislado. Estos resultados sugieren que habría una sinergia mutua entre los fragmentos de I_N e I_C para que la inteína adoptara su estructura tridimensional nativa.

Además, se ha estudiado como se ve afectada la actividad de la inteína cuando se sustituyen las Cys+1 y Cys1 por una Ser, tanto cuando la inteína forma una única cadena polipeptídica como cuando se encuentra partida en las dos subunidad I_N e I_C . En ambos casos, la mutación C+1S bloquea la actividad de esta a 20°C, pero se reconstituye la actividad a medida que se incrementa la temperatura hasta los 60°C. A partir de los 50°C, las inteínas mutantes presentaban una actividad igual a la inteína salvaje. En el caso de la mutación C1S se inhibía totalmente la actividad de esta inteína en todos los rangos de temperatura. Igual que en el caso de la doble mutación C+1S – C1S.

Así pues, la variante C+1S podría ser utilizada como herramienta biotecnológica ya que presenta un mecanismo de corte y empalme condicionado por temperatura.

Finalmente, en este trabajo se ha estudiado y complementado lo que se conoce sobre el proceso asociativo de los fragmentos I_N e I_C . Para hacer esto, se ha purificado y estudiado el plegamiento de tres variantes de *Npu DnaE* que presentan distintas longitudes del fragmento I_C . El plegamiento se ha analizado mediante un análisis informático previo, dicroísmo circular i RMN. Los resultados obtenidos muestran que a pesar que el fragmento I_N no presenta estructura propia, añadiendo los 12 primeros residuos del fragmento I_C , estos ya son capaces de empezar a reconstituir la estructura global de la *Npu DnaE*. Esta reconstitución de la estructura global se ve de forma significativa con los 21 primeros residuos del fragmento I_C . Se confirma de esta manera que el plegamiento global de esta inteína viene condicionado por una sinergia en el plegamiento de los fragmentos I_N e I_C .

SUMMARY

Inteins are protein domains that self-excise from a polypeptide chain precursor and catalyze ligation of the flanking sequences (exteins) with a peptide bond. The capacity to self-excise is a posttranslational process called protein splicing. This process is intramolecular, autocatalytic and does not require auxiliary enzymes, cofactors and metabolic energy. Protein splicing can also occur in *trans*. In this case, the intein is separated into N and C terminal domains, each joined to an extein. The intein domains reassemble and link the joined exteins into a single functional protein. Understanding protein splicing mechanism and how the N and C domains reassemble has led to the development of intein-mediated protein-engineering applications such as protein purification, cyclization or segmental isotopic labeling of multi-domain proteins among many other methods. In order to use inteins in these applications, they must possess a good thermoestability, a fast protein splicing rate, high solubility and their efficiency must be independent of the leading extein. To date, *Npu DnaE* is one of the best split inteins that satisfy these requirements.

In this PhD thesis we have studied the thermal stability of this intein in a range of different pHs when the N and C domains are forming a single polypeptide chain, when they are forming an I_N-I_C complex and in addition, we have studied the stability of the isolated I_N fragment. Results have shown that, in all cases the stability is higher when the pH is slightly basic and when the intein is forming a single polypeptide chain whereas the stability is lower when the intein is in the form of two fragments I_N and I_C forming a complex. This stability is even lower when the I_N fragment is isolated. Refolding of these three variants also follows this pattern where a greater degree of refolding appears with the intein as a single polypeptide chain and is completely negligible in the case of the isolated I_N fragment. These results suggest that there would be a mutual synergy between I_N and I_C fragments so that intein adopts its native three-dimensional structure and becomes stabilized only after association.

Furthermore, we studied how the activity of the intein is altered when substituting Cys+1 and Cys1 by Ser, both when intein is a single polypeptide chain and when it is split into fragments I_N and I_C. In both cases, C+1S mutation blocks the activity at 20 °C, but the activity was reconstituted when the temperature increased up to 60 °C. From 50 °C and on, mutant inteins showed the same activity that the wild-type inteins. The C1S mutation completely abolished the activity of this intein in the entire range of temperatures, result that was also observed in the case of the double mutation C+1S-C1S. Therefore, the variant C+1S could be used as a biotechnological tool because it has a conditional protein splicing mediated by temperature.

Finally, in this work we wanted to study and complement what is known about the associative process between I_N and I_C fragments. To accomplish this, we purified and studied the folding of three variants that have different lengths of the I_C fragment by performing bioinformatics analyses, circular

dichroism and NMR spectroscopy. The attained results show that despite the I_N fragment does not have a high degree of structure, the variant comprising the first 12 residues of the fragment I_C, shows a significant number of structural characteristics that are present in the global fold of the Npu *DnaE*. This reconstitution of the global structure is also considerably observed in the variant with the first 21 residues of fragment I_C. This fact confirms that the complete folding of this intein depends on the synergy of folding between fragments I_N and I_C that occurs upon association.

1. Introducció

1.1 LES INTEÏNES

1.1.1 Característiques generals

Les inteïnes formen part d'una família de proteïnes que es troba englobada dins la súper família *Hint*. Aquesta superfamília es caracteritza per presentar el domini *Hint* el qual té una llargada de 130 a 160 aminoàcids i la conservació de 4 o 5 regions (Amitai et al., 2003; Dassa and Pietrokovski, 2005). Presenten un plegament compacte amb una estructura simètrica composta majoritàriament per fulls β i els extrems N i C terminals tancats mútuament (Duan et al., 1997). Tots els membres d'aquesta superfamília es caracteritzen per participar en processos postraduccionalis com és el cas del trencament d'extrems N o C terminals en determinades proteïnes. Concretament les inteïnes són les responsables de dur a terme el mecanisme de tall i unió de proteïnes (Hall et al., 1997).

Les inteïnes són presents en els tres regnes de vida i també s'han descobert en genomes de virus i fags. No obstant, es troben confinades només en els organismes unicel·lulars (Evans and Xu, 2002). Aquest tipus de distribució ha generat diferents teories sobre la seva evolució, incloent el paper que tenen en la mobilitat del genoma. Actualment però, encara no hi ha una explicació del perquè existeixen les inteïnes i quina és la seva funció en els éssers vius (Gogarten et al., 2002; Liu, 2000; Perler, 1998; Pietrokovski, 2001).

Les inteïnes poden ser vistes com l'equivalent proteic als introns (figura 1.1). Els gens eucariotes es troben fragmentats, els quals les regions codificants o exons estan separades per regions no codificants o introns. Aquesta estructura es manté en el procés de transcripció, generant un pre-mRNA el qual per un procés de tall i unió donarà lloc al mRNA madur. Si el procés es dona per tall i unió alternatiu es pot aconseguir més d'una proteïna (mRNA madur) a partir d'un sol gen. El mecanisme de tall i unió proteic és un procés post-traduccional on les inteïnes promouen la seva pròpia escissió de la proteïna hoste i la lligació de les regions proteiques flanquejants, les extreïnes N i C. Aquest procés és intramolecular en cis però no en trans (apartat 1.2 de la introducció), autocatalític i no requereix ni de coenzims ni de fonts d'energia metabòlica per a dur a terme un seguit de reordenaments d'enllaços, que donen lloc a la formació de noves molècules (Kenneth and Paulus, 2005; Paulus, 1998). El mecanisme de tall i unió de RNA i proteic podrien ser vistos com a mètodes d'expressió alternatiu de gens interromputs que difereixen només en el nivell on es du a terme l'escissió de les seqüències interventores i on aquestes podrien tenir funcions reguladores. No obstant, aquestes dos mecanismes presenten una finalitat diferent: el mecanisme de tall i unió de RNA, en el cas que sigui alternatiu, confereix la capacitat d'acoblar diferents exons i formar diferents mRNAs a partir d'un sol gen, obtenint així una major variabilitat (Hirata et al., 1990) i el mecanisme de tall i unió proteic es troba sempre en dominis molt ben conservats i la seva funció és bàsicament l'activació d'una proteïna precursora inactiva (Paulus, 1998). Fins a l'actualitat només s'ha descrit, de

1. Introducció

manera no natural, un estudi en el qual s'ha observat un procés de tall i unió alternatiu a partir de l'intercanvi de dominis *in vitro* de diferents inteïnes (Aranko et al., 2013).

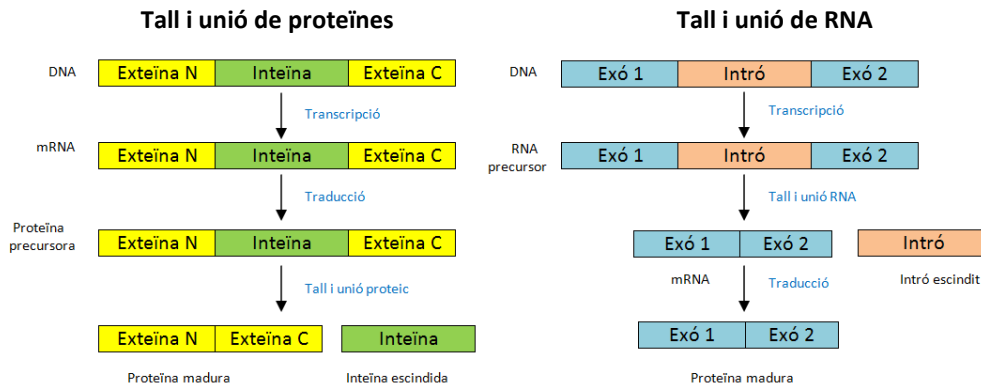


Fig. 1.1. Esquema comparatiu entre el mecanisme de tall i unió proteic i el de RNA.

1.1.2 Descobriment

El procés de tall i unió en proteïnes es va descobrir per primera vegada l'any 1990 en la subunitat catalítica de la ATPasa translocadora de H^+ de la membrana vacuolar (TFP1/VMA1) de *Saccharomyces cerevisiae* (Hirata et al., 1990; Kane et al., 1990).

Els autors van veure que, un cop purificada la subunitat catalítica de la ATPasa translocadora de H^+ de la membrana vacuolar, en el gel SDS-PAGE la proteïna purificada tenia una massa molecular de 70 kDa enlloc dels 119 kDa esperats. En estudiar la seqüència aminoacídica van veure que només les seqüències terminals tenien una homologia del 70% respecte a altres ATPases translocadores de H^+ . A més, la seqüència central presentava una homologia bastant elevada amb l'endonucleasa del llevat.

En un primer moment es va pensar que en un mateix gen hi havia dues proteïnes codificades i que el mecanisme de tall i unió alternatiu d'RNA explicaria l'eliminació de la seqüència central, per tal d'obtenir així, la seqüència de 70 kDa observada i no la de 119 kDa esperada. Es van fer estudis posteriors amb el mRNA de la ATPasa translocadora de H^+ per tal de verificar que el mecanisme de tall i unió de RNA explicava l'eliminació de la seqüència central. Aquests estudis van permetre observar que una vegada eliminats els introns, que es caracteritzen per tenir unes seqüències altament conservades en llevats, el mRNA madur seguia codificant per l'endonucleasa. Aquesta evidència experimental indicava que el mecanisme de tall i unió no era a nivell de RNA sinó de proteïna (Evans and Xu, 2002). Posteriorment es va demostrar que el mRNA de VMA1 és traduït sencer i que la part central és eliminada de forma postraduccional de la seqüència polipeptídica precursora mitjançant el mecanisme de tall i unió proteic (Perler et al., 1994).

Fins a l'actualitat s'han descobert noves inteïnes que es caracteritzen per presentar una baixa homologia entre elles, malgrat que presenten regions amb seqüències força conservades. Actualment, ja hi ha aproximadament 600 inteïnes catalogades a InBase, la base de dades d'inteïnes (<http://tools.neb.com/inbase/>), seguint una nomenclatura estandarditzada (Perler, 2002).

1.1.3 Organització

La longitud de les inteïnes varia entre els 134 i 650 residus (Perler, 2002) tot i que la inteïna més gran descrita és la *Ter DnaBI* (Yang et al., 2004) que presenta 1986 residus i la més curta descrita fins a l'actualitat és la *AceL-TerL* de 129 residus (Thiel et al., 2014). De manera convencional, les inteïnes s'organitzen en tres grans grups (figura. 1.2) (Volkman and Mootz, 2013):

- **Inteïnes clàssiques:** Constituïdes per tres dominis, dos dominis catalitzadors del procés de tall i unió que flanquegen un domini central amb activitat endonucleasa.
- **Mini-inteïnes:** Iguales que les anteriors però sense el domini endonucleasa.
- **Inteïnes partides:** Iguales que les mini-inteïnes però en aquest cas els dominis catalitzadors del procés de tall i unió es troben discontinus i els fragments I_N (corresponent a la regió N-terminal de la inteïna) i I_C (corresponent a la regió C-terminal de la inteïna) es troben fusionats a les respectives exteïnes i codificats en dos gens diferents.

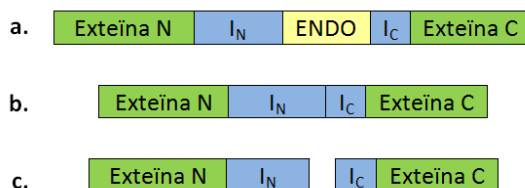


Fig. 1.2. Esquemes dels tipus de inteïnes trobades a la natura. (a) representa les inteïnes clàssiques amb el domini endonucleasa (ENDO), (b) representa les mini-inteïnes i (c) les inteïnes partides.

La anàlisi de diferents inteïnes ha permès identificar 6 regions de seqüències aminoacídiques conservades entre elles: A, N2, B, N4, F i G que participen en la seva activitat. Les inteïnes clàssiques presenten a més les regions centrals C, D, E i H (Pietrokovski, 1998) que constitueixen el domini endonucleasa tal i com s'observa en la figura 1.3. A continuació es detallen les característiques principals de cada regió (Pietrokovski, 1998; Starokadomskyy, 2007) i a la taula 1.1 es mostren les seqüències aminoacídiques consens de cada regió:

Les regions A, N2, B i N4 formen part de la regió N-terminal composta per uns 150-200 residus. La seva funció principal és facilitar la ruptura de l'enllaç peptídic a l'extrem N-terminal de la inteïna:

Regió A: Seqüència de 13 residus a l'extrem N-terminal. El primer residu és una Cys extremadament conservada i en alguns casos rars s'hi troba una Ala, Gln o Ser. El residu de la posició 13 és ocupat per una Gly i en alguns casos s'hi troba Ala, Lys, Thr, Arg, Tyr o Asn. Es suggereix que aquesta regió és extremadament important per l'inici i execució del procés de tall i unió.

1. Introducció

Regió N2: Composada per 7 residus on l'Asp Glu de la posició 5 es troben molt conservats i generalment precedits per una Gly.

Regió B: Constituïda per 14 residus on en la posició 10 hi ha una His i en la majoria de casos una Thr en la posició 7 involucrats en l'inici del procés de tall i unió.

Regió N4: formada per 16 residus on la posició 11 és ocupada per un Asp o Glu precedits per una Gly. En algunes inteïnes aquesta regió és absent.

A la regió C-terminal s'han identificat les regions F i G i generalment presenta una longitud de 25 a 40 residus:

Regió F: Constituïda per 16 residus, la meitat dels quals es troben ben conservats.

Regió G: Situada a l'extrem C-terminal de la inteïna i formada per 8 residus, set pertanyen a la inteïna i un a l'exteïna C. El penúltim aminoàcid de la inteïna és una His i l'últim una Asn que en algunes ocasions pot ser una Gln. Aquests 2 aminoàcids permeten la hidròlisi de l'enllaç peptídic de l'extrem C-terminal. El residu +1 és crític alhora de realitzar la lligació de les exteïnes i aquesta posició la sol ocupar una Cys, Ser o Thr.

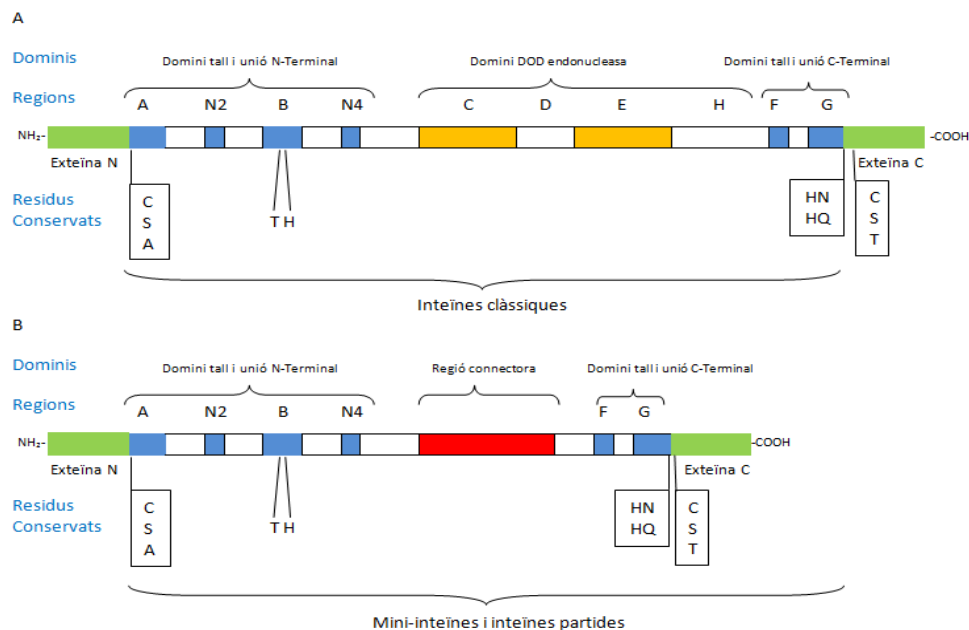


Fig. 1.3. Estructura esquemàtica de les inteïnes: Inteïnes clàssiques (A) i mini-inteïnes i partides (B) amb indicació d'alguns dels residus més conservats dels extrems N i C-terminals. La resta de residus amb un grau de conservació inferior es troba detallat a la taula 1.1.

Les inteïnes clàssiques tenen també les regions C, D, E i H centrals que formen el domini DOD endonucleasa (Matsumura et al., 2006). Aquest domini no és necessari per dur a terme el procés de tall i unió. El domini DOD endonucleasa presenta la seqüència aminoacídica LAGLIDADG en el seu centre actiu (Guhan and Muniyappa, 2003) que és capaç de reconèixer una seqüència de DNA específica no palindròmica de 14 a 40 pb. El domini endonucleasa assegura la transferència de la seqüència codificadora de la inteïna des del gen hoste (al·lel (+)) al gen de l'exteïna que no presenta

1. Introducció

la seqüència codificadora (al·lel (-)) de la inteïna. A la vegada, el domini DOD involucra una part de l'extrem N-terminal de la inteïna (DRR) que reconeix una seqüència de DNA en l'al·lel (-). Per tant, el reconeixement del DNA de l'al·lel (-) es fa a partir de dos centres actius, un és el domini DOD que reconeix el lloc de restricció i l'altre és la regió DRR que reconeix una regió distal de l'al·lel (-) (Bakhrat et al., 2004). A partir d'aquests reconeixements es fa la transferència de la seqüència codificadora de la inteïna de l'al·lel (+) a l'al·lel (-).

Taula 1.1. Seqüències aminoacídiques consens de les diferents regions de les inteïnes

Domini	Regió	Seqüència aminoacídica	Designacions
N-terminal (I _N)	A	Ch..Dp.hhh..G	h: residus hidrofòbics (G, V, L, I, A, M) a: residus àcids (D, E) r: residus aromàtics (F, Y, W) p: residus polars (S, T, C) .: residus no conservats p ⁺¹ : residu +1 de la exteïna C (C, S, T) ↓: lloc trencament en l'extrem C-terminal
	N2	...GD..	
	B	G..h.hT..H.hhh	
	N4GD.....	
Endonucleasa	C	LhG..hhaG	
	D	.K.IP..h	
	E	.L.GhFahDG	
	H	p.S..hh..h..LL..hGI	
C-terminal (I _C)	F	rVYDLpV..a..HNFh	
	G	NGhhhHN↓p ⁺¹	

1.1.4 Nomenclatura

La primera inteïna descrita es va anomenar *spacer* però a mesura que se'n descobrien de noves, se les va anomenar seqüències proteiques interventores i finalment se les va anomenar inteïnes (*internal protein sequence*). Seguint aquest conveni, els pèptids o proteïnes que flanquegen les inteïnes es van anomenar exteïnes (*external protein sequence*) (Perler et al., 1994).

El nom específic de cada inteïna consisteix en una abreviació del nom de l'espècie que la conté i el nom del gen que codifica l'exteïna que porta acoblada. Per exemple, la inteïna *Vma1* (vacuolar membrane H⁺-ATPase) de *Saccharomyces cerevisiae* s'anomena *SceVMA*. En el cas que una mateixa exteïna tingui diferents inteïnes, aquestes es diferencien a partir de xifres romanes: *TerRIRI*, *TerRIRII*... i en el cas que presenti activitat endonucleasa s'afegeix PI a davant: *PI-Pful* (Anraku et al., 2005).

Hi ha també una nomenclatura basada en la posició de l'aminoàcid utilitzada normalment quan les inteïnes s'enumeren fora del context de la proteïna hoste. El residu -1 fa referència a l'aminoàcid de l'exteïna adjacent a l'extrem N-terminal de la inteïna i vindrà seguit del -2, -3... El primer residu de la C-exteïna s'anomena +1 seguit del +2, +3... En el cas de la inteïna el primer residu de l'extrem N-terminal és el residu 1 seguit del 2, 3..., tal com s'esquematitza a la figura 1.4.

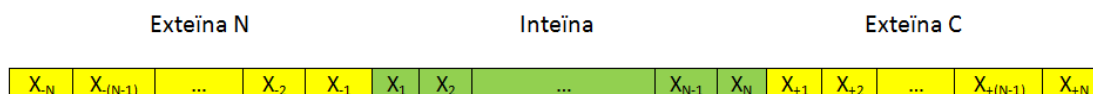


Fig. 1.4. Sistema de numeració dels aminoàcids per tal de facilitar la comparació entre inteïnes.

1.2 EL MECANISME DE TALL I UNIÓ

1.2.1 El procés

El mecanisme de tall i unió va romandre desconegut durant 6 anys a causa de la rapidesa del procés. L'any 1996 es va definir el procés tal i com és conegut a l'actualitat per Xu i Perler. Aquest procés té lloc a través d'un conjunt d'atacs nucleofílics dels residus dels extrems N i C terminals que provoquen atacs a enllaços amida i éster i fins i tot als seus grups amino. A la vegada es produeix la ciclació d'una Asn de l'extrem C-terminal de la inteïna i la formació d'un enllaç peptídic que és una característica exclusiva del procés de tall i unió en proteïnes (Xu and Perler, 1996).

Aquest mecanisme consisteix en una sèrie de reaccions separades en 4 passos (figura 1.5) detallats a continuació:

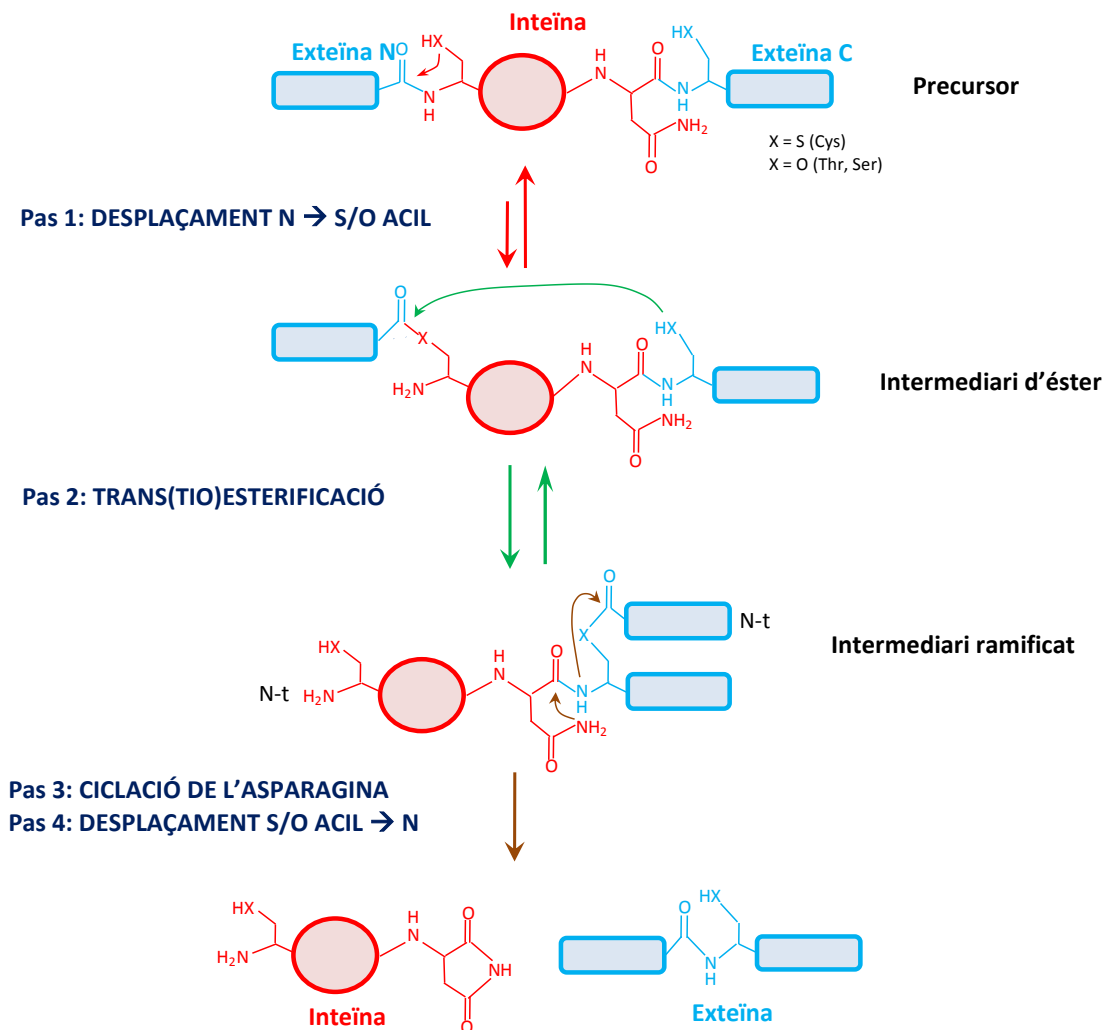


Fig. 1.5. Esquema del mecanisme de tall i unió.

1. Introducció

Pas 1 – Desplaçament N → S/O Acil: El procés s'inicia amb un atac nucleofílic del grup OH o SH de la cadena lateral dels residus de Ser, Cys o Thr de l'extrem N-terminal de la inteïna al carboni carbonílic de l'enllaç peptídic que hi ha entre els residus -1 i 1. Aquest atac genera un enllaç (tio)éster entre la inteïna i l'exteïna N formant el que s'anomena intermediari éster (Shao et al., 1996). El pas 1 és el més ràpid de tot el mecanisme de tall i unió i la velocitat de la reacció d'aquest pas pot ser fins a 100 vegades més ràpid que tot el procés global (Romanelli et al., 2004).

El trencament d'un enllaç peptídic per formar-se un enllaç d'elevada energia com és el (tio)éster és des d'un punt de vista termodinàmic desfavorit. Això fa que sigui un dels passos de tot el mecanisme que ha estat més estudiat per tal d'entendre el perquè es duu a terme i se sap que no es duu a terme quan les inteïnes es troben desnaturalitzades. Per tant, aquest pas es donaria gràcies a les relacions entre estructura i activitat de les inteïnes degut a la conservació de determinats residus en determinades regions molt concretes de la inteïna (Romanelli et al., 2004). S'ha observat que la His 10 de la regió B (B:10) conservada en la majoria de les inteïnes és essencial per dur a terme aquest pas (Binschik and Mootz, 2013; Ghosh et al., 2001; Kawasaki et al., 1997). La mutació d'aquesta His per una Ala bloqueja o inhibeix significativament el primer pas en la majoria d'inteïnes (Du et al., 2011a). Du i col·laboradors van determinar que la His B:10 actua primer com a base per desprotonar la cadena lateral de la Cys 1 (A:1) i llavors cedeix els protons al nitrogen amida de l'enllaç peptídic (Du et al., 2009) afavorint així la reacció. Un altre residu important en aquest pas és l'Asp present a la regió F (F:4). Diferents estudis han indicat que el residu d'Asp forma un pont d'hidrogen amb la Cys A:1 i el residu protonat d'Asp estabilitzaria el grup tiol de la cadena lateral de la Cys, facilitant així l'atac d'aquest grup tiol al carboni carbonílic situat a l'enllaç peptídic (Du et al., 2011b). Alguns estudis han mostrat que els residus hidrofòbics de la regió G es podrien trobar involucrats en aquest procés mitjançant forces de van der Waals, i a la vegada, la cadena lateral del residu +1 pot estar involucrada en la polarització i trencament de l'enllaç peptídic (Appleby et al., 2009; Volkmann and Liu, 2011).

Pas 2 – Trans(tio)esterificació: En aquest pas es produeix un altre atac nucleofílic del grup OH o SH de la cadena lateral del residu +1 de la exteïna C sobre el carboni carbonílic del (tio)éster format en l'etapa anterior. Aquest atac constitueix una reacció de transesterificació en el qual la exteïna N és transferida a la cadena lateral del residu +1 de la exteïna C. Aquesta reacció produeix un intermediari ramificat, és a dir, una proteïna amb dos extrems N-terminal. Durant aquest procés hi ha una reordenació de càrregues i a la vegada, s'indueix la ciclació de la cadena lateral de l'Asn (Xu et al., 1993).

El pas dos és probablement el més complex d'estudiar experimentalment. S'han postulat diferents requeriments per dur a terme aquesta catàlisi similars als del pas 1. Diferents estudis han conclòs que l'Asp de la regió F (F:4) és capaç de captar un protó del grup tiol de la Cys de la regió G (G:8).

1. Introducció

Aquest tiolat, aleshores, ataca el carboni carbonílic de l'intermediari d'éster formant-se l'intermediari ramificat (Johnson et al., 2007; Pereira et al., 2011).

Pas 3 – Ciclació de l'Asn: En aquest tercer pas, el més lent de tot el procés (Kurpiers and Mootz, 2008), es produeix un atac del nitrogen β -amida al carboni carbonílic de l'enllaç peptídic donant lloc al trencament d'aquest i conduint a la ciclació de la cadena lateral de l'Asn (Shao et al., 1995; Xu and Perler, 1996). La ciclació de la cadena lateral de l'Asn provoca la formació d'un anell de succinimida i el trencament de l'enllaç peptídic a l'extrem C-terminal de la inteïna (Xu et al., 1994).

La ciclació de l'Asn és la primera reacció irreversible del procés de tall i unió ja que es produeix el trencament entre la inteïna i la exteïna C. Diversos estudis han demostrat que aquesta reacció va estretament lligada a la primera i segona reacció del procés de tall i unió (Frutos et al., 2010; Martin et al., 2001; Zettler et al., 2009). En aquest pas, es creu que les cadenes laterals de la His F:13 i His G:6 juntament amb molècules d'aigua en la posició correcta són capaces de reduir les barreres energètiques crucials en aquest estat de transició (Mujika et al., 2009). L'estat de protonació d'aquestes dues His també és essencial per dur a terme aquest procés: només quan la penúltima His es troba completament protonada (càrregada positivament) i la His de la posició F es troba en la forma neutre es dona la ciclació de l'Asn (Mujika et al., 2012).

Pas 4 – Desplaçament S/O acil \rightarrow N: La inteïna és escindida amb la ciclació de l'Asn i el trencament de l'enllaç peptídic. No obstant, la inteïna conté l'anell de succinimida i les exteïnes es troben unides per un enllaç (tio)éster. L'enllaç (tio)éster és energèticament desavantatjós i es produeix un canvi espontani formant-se un enllaç peptídic entre les exteïnes. La molècula de succinimida és hidrolitzada per donar lloc de nou a una Asn (Shao and Paulus, 1997).

En aquest pas, la inteïna no contribueix en la formació de l'enllaç peptídic i, a la vegada, és completament independent a la temperatura amb energies d'activació molt baixes (4-5 kcal/mol) (Frutos et al., 2010; Shao and Paulus, 1997).

In vitro, s'ha observat en determinades ocasions que el mecanisme de tall i unió no és eficient del tot i es produeixen altres reaccions secundàries i no sempre s'arriba a obtenir les exteïnes unides per una banda i la inteïna escindida per l'altra. Aquestes reaccions secundàries depenen sobretot de l'exteïna que porta unida la inteïna, de les seqüències aminoacídiques de les regions flanquejants exteïna-inteïna i en determinats casos el pH, temperatura o estat redox poden facilitar aquestes reaccions (Amitai et al., 2009; Ellila et al., 2011; Nicastri et al., 2013; Nichols and Evans, 2004; Shah et al., 2013a; Shemella et al., 2011; Wasmuth et al., 2013). Les reaccions secundàries més comunes que es produeixen són la N escissió (*N-cleavage*) i la C escissió (*C-cleavage*). La N escissió consisteix en un trencament de l'enllaç entre l'exteïna N i l'extrem N-terminal de l'inteïna donant lloc a la inteïna

1. Introducció

sencera unida a l'exteïna C i l'exteïna N alliberada. Aquesta reacció secundària es produeix per un trencament per hidròlisi o tiòlisi de l'enllaç (tio)éster format en els passos 1 i 2. En el cas de la C escissió, consisteix en un trencament de l'enllaç entre l'exteïna C i l'extrem C-terminal de la inteïna donant lloc a la inteïna sencera unida a l'exteïna N i l'exteïna C alliberada. Aquesta reacció succeeix a causa de la ciclació de l'Asn desacoblada als passos 1 i 2 (Topilina and Mills, 2014). La figura 1.6 mostra un esquema de les possibles reaccions secundàries que poden tenir lloc.

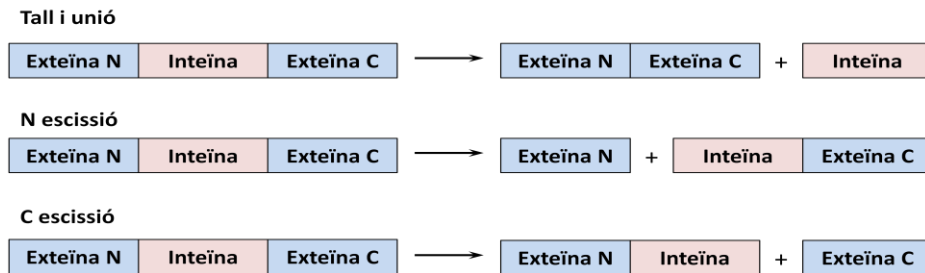


Fig. 1.6. Esquema del tall i unió de proteïnes i de les possibles reaccions secundàries que poden succeir com la N i C escissió.

El mecanisme de tall i unió es troba estandarditzat en la majoria de inteïnes gràcies a la conservació de determinades regions tant de la inteïna com de l'exteïna tal i com s'ha comentat en l'apartat 1.1.4. Algunes inteïnes però es caracteritzen per presentar alguna modificació en aquestes regions tant ben conservades i fins i tot algunes d'elles es troben codificades en dos gens diferents i són anomenades inteïnes partides o trans inteïnes. Totes aquestes excepcions es comentaran més endavant.

1.2.2 Classificació

Des del descobriment de les primeres inteïnes i el seu mecanisme de funcionament s'han descrit noves inteïnes que presenten mecanismes diferents als descrits originalment (figura 1.7). Això ha conduït a una nova classificació d'aquestes atenent al seu mecanisme de tall i unió (Tori et al., 2010). La classificació d'inteïnes proposada segons els mecanismes utilitzats per dur a terme el tall i unió és la següent:

Classe 1: La majoria d'inteïnes (>90%) pertanyen a aquest grup i presenten el mecanisme de tall i unió descrit a l'apartat 1.2.1. La característica principal d'aquestes inteïnes és que presenten un residu nucleofílic a l'extrem N-terminal (Kenneth and Paulus, 2005).

Classe 2: Aquesta classe d'inteïnes es caracteritza per no presentar un residu nucleofílic a la primera posició de la regió A, la qual és ocupada per una Ala. El mecanisme de tall i unió en aquestes inteïnes comença directament pel pas 2 del mecanisme de tall i unió descrit a l'apartat 1.2, ja que es produeix un atac nucleofílic directament de la cadena lateral del residu +1 de la exteïna C sobre el carboni carbonílic de l'enllaç peptídic present entre els residus -1 i 1. El trencament de l'enllaç peptídic de

1. Introducció

l'extrem N-terminal de la inteïna sembla ser que és facilitat per les estructures tridimensionals específiques que adopten aquest tipus d'inteïna. La inteïna *Mja* K1bA és un exemple d'aquesta classe d'inteïnes (Perler, 2005; Saleh et al., 2011; Southworth et al., 2000).

Classe 3: Aquest grup d'inteïnes tampoc presenten el residu nucleofílic a la primera posició de la regió A, però es caracteritzen per presentar el triplet Trp-Cys-Thr (WCT) a la regió F, del qual la Cys ataca nucleofílicament l'extrem N-terminal de la inteïna, entre el residu -1 i 1, per tal d'iniciar-se el mecanisme de tall i unió. D'aquest triplet, mutacions del Trp i la Thr no afecten l'eficiència de la inteïna, mentre que s'ha vist que la Cys és essencial en el procés. Dues inteïnes que pertanyen a aquest grup són *MP-Be DnaB* (Tori et al., 2010) i *Dra Snf2* (Brace et al., 2010).

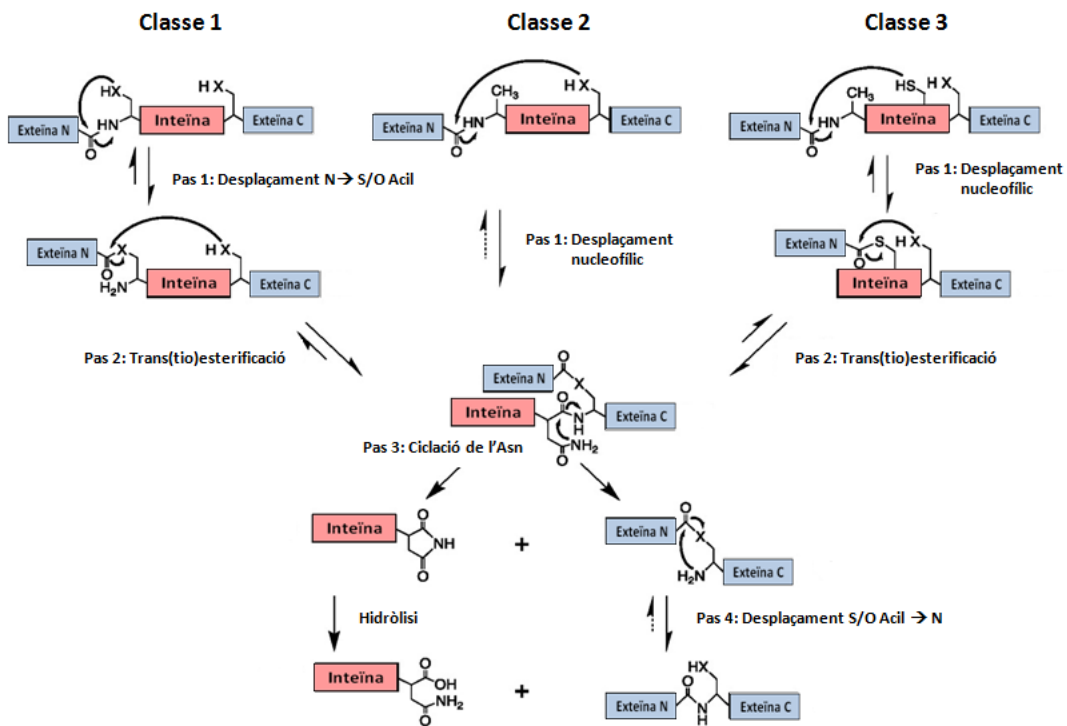


Fig. 1.7. Mecanismes de tall i unió de les tres classes d'inteïnes diferents. Esquema reproduït de (Tori and Perler, 2011).

A causa de l'elevada divergència que s'observa a mesura que es descobreixen noves inteïnes, s'estan trobant inteïnes que presenten característiques mixtes corresponents a les diferents classes. Per exemple, les inteïnes *Tfu2914* i *Nsp-1S614 TOPRIM* presenten una Cys en el primer residu de la regió A i el triplet WCT. Estudis mutagènics han mostrat que la Cys de la regió A és essencial pel procés de tall i unió mentre que la Cys del triplet no (Reitter and Mills, 2011). En canvi, en la inteïna *MP-Catera Gp206* que presenta les mateixes característiques que les anteriors, la Cys del triplet és essencial per dur a terme el procés de tall i unió, per tant, aquesta inteïna es troba catalogada com a classe 3 (Tori and Perler, 2011).

1.2.3 Alteracions en residus clau de la seqüència de la inteïna

1.2.3.1 Inteïnes sense la His 10 de la regió B

La His 10 de la regió B (figura 1.3) és essencial pel primer pas del mecanisme de tall i unió. Aquesta His és present en totes les inteïnes descrites fins a l'actualitat excepte en la Tko CDC21-1 de *Thermococcus kodakaraensis*. Aquesta inteïna és completament activa ja la funció de la His 10 la duu a terme una Lys en la posició 58, fora de tots els dominis altament conservats en inteïnes. Aquesta Lys catalitza el primer pas del mecanisme de tall i unió (figura 1.5), estabilitzant l'intermediari que es forma quan es trenca l'enllaç peptídic entre l'exteïna N i l'extrem N-terminal de la inteïna per formar l'enllaç tioéster (Tori et al., 2012).

1.2.3.2 Inteïnes sense la penúltima His

El penúltim residu de la seqüència d'una inteïna, en la majoria dels casos, és una His i es creu que té un paper molt important en la ciclació de l'Asn (Ding et al., 2003; Mills et al., 2005). No obstant, algunes inteïnes com la *Ssp DnaE* no presenten aquesta His la qual és substituïda per una Ala. En aquesta inteïna, l'Arg 73 de la regió B actua com a substituta de la penúltima His, atès que la flexibilitat que presenta la cadena lateral li permet formar ponts d'hidrogen amb l'àtom d'oxigen de la cadena lateral de l'Asn, la qual queda estabilitzada facilitant la seva ciclació. Un exemple d'aquestes inteïnes és la *Ssp DnaE*, la qual pertany a la família *DnaE* d'inteïnes, cap de les quals presenta la penúltima His. En totes elles però, en el bloc B hi ha una Arg, Lys o Gln que actua com a substitut (Sun et al., 2005).

Per mutagènesi s'ha reintroduït aquesta penúltima His en inteïnes que no la presentaven i totes elles han presentat el mecanisme de tall i unió actiu en *E. coli*, exceptuant la inteïna *ClpP*. Aquests resultats suggereixen que la penúltima His prové d'un ancestre comú i que per mutació, algunes l'han perdut i el seu paper l'han assumit altres residus (Chen et al., 2000; Wang and Liu, 1997).

1.2.3.3 Inteïnes amb Gln o Asp com a últim residu

Algunes inteïnes presenten una Gln o un Asp enlloc d'una Asn com a últim residu (figura 1.3). En el cas de la Gln, es produeix també una ciclació formant-se aminoglutarimida i en el cas de l'Asp es creu que també es produeix una ciclació formant-se anhídrid succínic. No obstant, aquest últim és molt inestable en medis aquosos i és molt difícil d'identificar. Dos exemples d'aquest tipus d'inteïnes són la *Chy* que presenta un Asp i la *CIV* que presenta una Gln. Una mutació de l'extrem C-terminal d'aquestes inteïnes per una Asn no altera el mecanisme de tall i unió a *CIV*, mentre que en el cas de la *Chy* el mecanisme de tall i unió queda aturat a l'etapa de formació de l'intermediari ramificat (Amitai et al., 2004).

1.2.4 El procés de tall i unió en trans

Tal i com s'ha mencionat anteriorment, algunes inteïnes es troben, de manera natural, partides en dos fragments codificats a partir de dos gens diferents que donen lloc a les regions N- i C-terminals de la inteïna. Aquest tipus de inteïnes representa el 5% del total de inteïnes conegudes fins a l'actualitat. De manera general, els dos fragments presenten una elevada afinitat l'un per l'altre i a la vegada cada un d'ells porta unit l'exteïna corresponent. L'afinitat que presenten els dos fragments es deu bàsicament als nuclis hidrofòbics que contenen (Contreras Martinez et al., 2008) i a les interaccions electrostàtiques que s'estableixen entre els dos fragments. Aquests fragments generalment presenten afinitats de l'ordre de nanomolar baixes i l'elevada capacitat d'associar-se és independent del mecanisme de tall i unió (Shi and Muir, 2005). L'associació entre els dos fragments no es pot considerar com una interacció proteïna-ligand, però sí que la seva unió s'ha vist que condueix al replegament dels dos fragments i a activar l'activitat de la inteïna (Shah et al., 2013b; Zheng et al., 2012). Es va calcular la força d'unió que presentaven els fragments de tres inteïnes partides dels cianobacteris *Oscillatoria limnetica*, *Thermosynechococcus vulcanus* i *Nostoc sp PCC7120* mitjançant microscòpia de forces atòmiques (AFM). Es va veure que la força d'unió entre els fragments era de 50-52 pN entre les tres espècies i també, quan aquestes inteïnes es creuaven en les 9 combinacions possibles. Per tant, es va veure que les diferents forces d'unió eren similars entre diferents inteïnes partides de cianobacteris i que aquesta afinitat d'unió depenia de les interaccions electrostàtiques i dels nuclis hidrofòbics que es trobaven concentrats en una interfase d'unió "en cremallera" dels fragments (Sorci et al., 2013).

L'elevada afinitat fa que els dos fragments siguin capaços d'associar-se i iniciar el mecanisme de tall i unió seguint un procés tal com el que s'ha descrit en l'apartat 1.2.1. En aquest cas el mecanisme s'anomena tall i unió en trans (figura 1.8) i les inteïnes que hi participen són anomenades trans inteïnes o inteïnes partides "split inteïns" (Kenneth and Paulus, 2005). La majoria d'elles s'han descrit en cianobacteris (Choi et al., 2006; Liu and Yang, 2003).

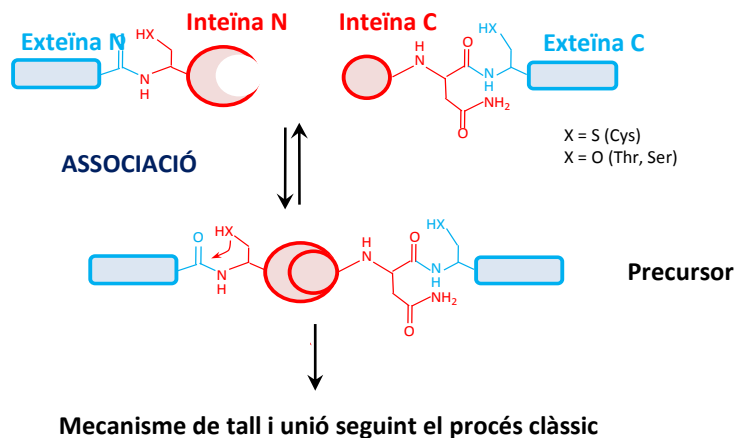


Fig. 1.8. Esquema del mecanisme de tall i unió en trans.

1. Introducció

El mecanisme de tall i unió en trans es va descobrir en la inteïna *DnaE* de *Synechocystis sp.* (*Ssp DnaE*) (Codi PDB: 1ZD7) (Sun et al., 2005). Aquesta inteïna es troba al mig de la seqüència de la subunitat α de la DNA polimerasa III, en forma d'exteïna, en aquests organismes (Dassa et al., 2007). Els seus gens precursors es troben en cadenes oposades del DNA separades per unes 80 kb (Starokadomsky, 2007). Aquesta disposició en el genoma suggereix que hi va haver una inversió del DNA just en el gen de la inteïna. Aquesta inversió però, no va ser letal i, a més, les noves inteïnes de la família *DnaE* identificades fins a l'actualitat comparteixen la mateixa exteïna i els mateixos punts de fragmentació (Iwai et al., 2006). Aquesta característica va permetre proposar la hipòtesi segons la qual totes elles provenen d'un ancestre comú (Saleh and Perler, 2006).

L'organització de les regions de les inteïnes partides és la mateixa que la de les convencionals i la majoria d'elles es troben fragmentades entre les regions B i F, coincidint amb el lloc d'inserció del domini endonucleasa de les inteïnes clàssiques. De manera general el fragment N-terminal (I_N) sol tenir una longitud superior a 100 aminoàcids, tot i que en la inteïna *Ssp DnaB* presenta només 11 aminoàcids (Ludwig et al., 2006) i el fragment C-terminal (I_C) varia entre 24 i 44 aminoàcids (Mootz, 2009; Saleh and Perler, 2006), amb excepcions com és el cas de la *Ssp GyrB* que només en presenta 6 (Appleby et al., 2009).

Malgrat que se'n coneixen alguns detalls aïllats, no disposem d'una visió completa de com es produeix l'associació dels dos fragments de la inteïna. En el cas de la *Ssp DnaE* utilitzant la GFP com a exteïna es va concloure que l'associació dels dos fragments es donava en 2 fases i consistia, bàsicament, en interaccions electrostàtiques. La primera fase era dependent de la concentració dels fragments i la segona es donava a causa d'una reorganització del complex format en la primera fase (Shi and Muir, 2005). Posteriorment Liu i col·laboradors van postular que els fragments de la inteïna *Ssp DnaE* es trobaven desplegats i que un cop aquests associaven, s'indueïen canvis conformacionals que dotaven d'activitat a la inteïna. El procés el van definir com a sinèrgia de plegament mútua entre els dos fragments de la inteïna (Zheng et al., 2012). Els estudis fets sobre el procés associatiu dels fragments de la inteïna *Npu DnaE*, inteïna estudiada en aquesta tesi doctoral, es detallaran a l'apartat 1.4.5 d'aquesta introducció.

La majoria d'inteïnes partides són ortogonals (no presenten el mateix punt d'inserció en el genoma de les diferents espècies) exceptuant el cas de la família *DnaE* que són al·lèliques (presenten el mateix punt d'inserció en el genoma de les diferents espècies). Barrejant fragments de diferents inteïnes ortogonals es va observar que el mecanisme de tall i unió es donava només entre els fragments complementaris de la mateixa inteïna i no es va detectar una interferència apreciable entre inteïnes (Brenzel et al., 2006; Shi and Muir, 2005). En canvi, quan es tracta d'inteïnes al·lèliques com és el cas de la *Ssp DnaE* i la *Npu DnaE*, sí que poden presentar activitat creuada i donar lloc a un

procés de tall i unió productiu, la qual cosa es deu bàsicament a l'elevada similitud entre les seqüències (Dassa et al., 2007; Iwai et al., 2006).

Les inteïnes partides han esdevingut una eina molt utilitzada en diferents aplicacions de treball amb proteïnes dins el camp de la biotecnologia i la biologia molecular. Algunes d'aquestes aplicacions es descriuen en l'apartat següent d'aquesta introducció.

1.3 APLICACIONS

L'interès en l'estudi de les inteïnes i el seu mecanisme no es troba només en el coneixement en sí, sinó també en les aplicacions que se'n deriven. Les investigacions realitzades fins a l'actualitat ha permès el desenvolupament de noves tècniques en el camp de la biotecnologia i la biologia molecular basades en les inteïnes i la seva activitat. Han aparegut tècniques de purificació, lligació i ciclació de proteïnes, i recentment, noves aplicacions *in vivo* tant en cèl·lules com en organismes (Topilina and Mills, 2014). Actualment, ja hi ha més de 20 tècniques basades en el mecanisme de tall i unió d'inteïnes. A continuació se'n descriuen algunes:

1.3.1 Expressió i modificació de proteïnes

1.3.1.1 Ciclació de proteïnes

Les proteïnes cíclics són proteïnes unides per enllaç peptídic entre els seus extrems N i C-terminals. Això fa que presentin una major estabilitat a nivell químic i tèrmic i una major bioactivitat i resistència a la degradació per exoproteases respecte a les seves homòlogues lineals (Jeffries et al., 2006; Kanno et al., 2007; Sakamoto et al., 2013; Williams et al., 2005).

Per posar en pràctica aquesta aplicació (figura 1.9) cal fusionar la proteïna de la qual se'n vol fer la ciclació a una inteïna partida en els seus extrems N i C terminals (I_C -proteïna- I_N). Mitjançant el mecanisme de tall i unió s'escindiran els fragments de la inteïna i la proteïna quedarà unida pels seus extrems N i C terminals. Emprant aquesta tècnica s'ha aconseguit de manera ràpida l'obtenció d'àmplies llibreries de pèptids cíclics (Abel-Santos et al., 2003; Kritzer et al., 2009; Tavassoli and Benkovic, 2007).

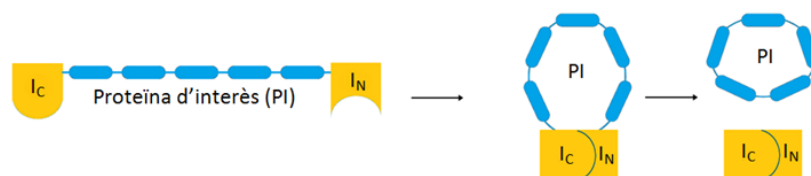


Fig. 1.9. Esquema de la generació de proteïnes cícliques.

La ciclació *in vivo* (interior de la cèl·lula) de proteïnes utilitzant inteïnes, permet ciclar pèptids i proteïnes sense limitacions de la seva mida i permet la incorporació de dianes per tal de facilitar-ne el procés de purificació (Tavassoli and Benkovic, 2007). És per tant, una tècnica sense restriccions de tipus i mida de la proteïna a ciclar i a més, a diferència de les tècniques de ciclació de pèptids convencionals, no cal dissenyar els extrems dels pèptids per tal que siguin resistents a la maquinària catabòlica de la cèl·lula hoste que pot fer variar l'activitat d'aquests pèptids (Gururaja et al., 2000; Jermutus et al., 2001).

1.3.1.2 Expressió de pèptids i proteïnes tòxiques

En els darrers anys, els pèptids estan esdevenint una classe important d'agents terapèutics (Bellmann-Sickert and Beck-Sickinger, 2010; Vlieghe et al., 2010) ja que permeten la regulació de funcions fisiològiques, interaccionant la majoria d'ells amb receptors cel·lulars i induint cascades de senyals cel·lulars. No obstant, es troben limitats per la seva estabilitat (Frokjaer and Otzen, 2005). Alguns d'ells requereixen, a més, d'algun grup funcional en el seu extrem N- o C- terminal com pot ser un α -amida per tal que sigui actiu. Aquests grups funcionals, generalment s'han d'introduir mitjançant síntesi química la qual dificulta i encareix el procés (Merkler, 1994; Vlieghe et al., 2010).

L'expressió de petits pèptids i proteïnes citotòxiques es pot aconseguir mitjançant el mecanisme de tall i unió. Els pèptids es poden expressar de manera conjunta amb la inteïna i un domini d'afinitat per facilitar-ne la purificació (figura 1.10) i les proteïnes tòxiques es poden expressar de manera inactiva fent que la inteïna parteixi la seva seqüència (Topilina and Mills, 2014). Aquesta tècnica ha permès, per exemple, l'estudi de pèptids del complex major d'histocompatibilitat (Farfan-Arribas et al., 2012) o la purificació a gran escala del pèptid YY involucrat en tractaments d'obesitat (Greenwood et al., 2011; Karra et al., 2009), amb el grup funcional α -amida en el seu extrem C-terminal (Albertsen et al., 2013). També, s'ha aconseguit purificar proteïnes tòxiques com les endonucleases (Miao et al., 2005; Wu et al., 2002) o proteïnes amb activitats antimicrobianes (Wang et al., 2012).



Fig. 1.10. Esquema de l'expressió de pèptids mitjançant inteïnes. D.A. fa referència a domini d'afinitat 1 i 2 respectivament. N- i C- term. fa referència als extrems N i C terminals de la proteïna o pèptid i int fa referència a la inteïna involucrada en el procés.

1.3.1.3 Lligació de proteïnes expressades (EPL)

L'EPL és una tècnica que permet obtenir proteïnes amb modificacions post-traduccionals i introduir aminoàcids no naturals (Muir et al., 1998). És una tècnica semisintètica on es necessita una proteïna, la qual, al seu extrem C-terminal hi hagi fusionada una inteïna mutada per tal que no es pugui formar o trencar l'intermediari ramificat (pas 2 del procés de tall i unió, figura 1.5). A més, es necessita un

pèptid, el qual, en el seu extrem N-terminal hi hagi una Cys. Mitjançant síntesi en fase sòlida de pèptids (SPPS) es pot construir de manera sintètica l'extrem N o C terminal de la proteïna el qual se li vol introduir les mutacions post-traduccionals i per tant, se li podran introduir una àmplia varietat de modificacions. L'altre extrem de la proteïna que no hagi estat sintetitzat mitjançant SPPS s'obté a partir de l'expressió recombinant de proteïnes.

La tècnica (figura 1.11) comença amb el mecanisme de tall i unió que quedarà aturat en el pas 2. Aleshores s'elimina la inteïna amb el tractament d'un agent reductor i s'obté la proteïna recombinant amb un grup tioéster en l'extrem C-terminal. El grup tioéster reacciona, amb l'ajuda d'una lligació química nativa (NCL), amb la Cys que conté el pèptid en l'extrem N-terminal formant-se un enllaç peptídic entre les dues molècules. D'aquesta manera s'obté la proteïna fusionada a un pèptid amb les modificacions desitjades (Berrade and Camarero, 2009; Flavell and Muir, 2009). Mitjançant aquesta tècnica, s'ha aconseguit recentment, per exemple introduir modificacions a receptors transmembrana (Singla et al., 2011) o realitzar modificacions específiques en anticossos monoclonals (Vila-Perello et al., 2013).

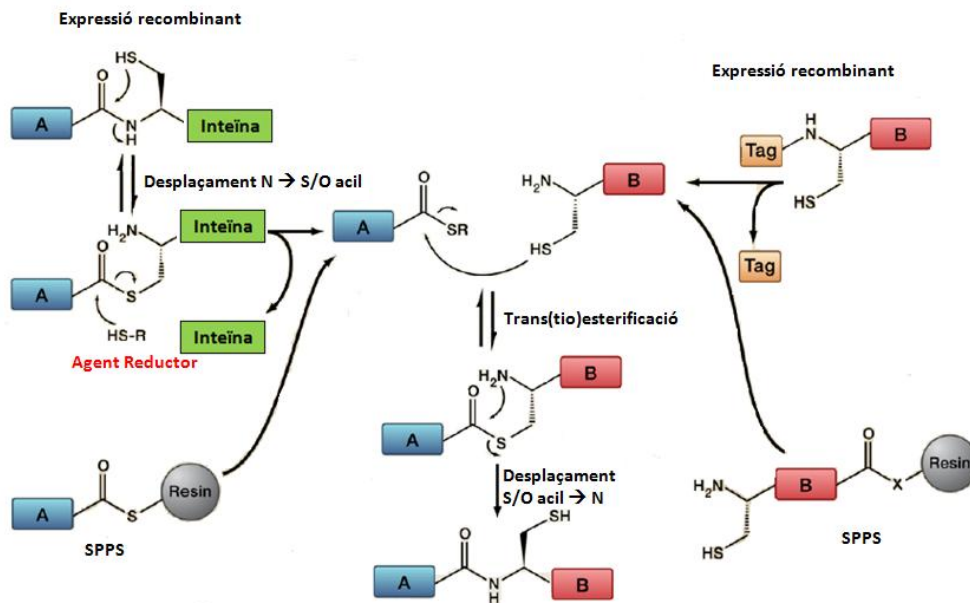


Fig. 1.11. Esquema del procés d'obtenció de proteïnes fusionades a partir del mètode EPL. Al centre es mostra la reacció de NCL en la qual la trans(tio)esterificació de la proteïna α-tioéster amb el pèptid amb la Cys en el seu extrem N-terminal és seguit per un desplaçament S/O acil → N per generar un nou enllaç peptídic per unir els dos polipèptids. Els extrems N i C-terminals poden ser obtinguts de manera recombinant o per síntesi química mitjançant síntesi en fase sòlida de pèptids (SPPS). Esquema reproduït de (Vila-Perello and Muir, 2010).

1.3.2 Inteïnes com a mediadores de l'etiquetatge de proteïnes

La tècnica descrita anteriorment (EPL) permet la modificació del l'extrem N i C-terminal d'una proteïna i també ha permès introduir fosforilacions, glicosilacions etc. Però hi ha un altre conjunt de tècniques que permeten fer altres modificacions d'aquestes proteïnes *in vivo* incloent la semi síntesi

de proteïnes a la superfície cel·lular (Dhar and Mootz, 2011; Volkmann and Liu, 2009) o el marcatge isotòpic de proteïnes multi domini (Zuger and Iwai, 2005).

1.3.2.1 Marcatge isotòpic de proteïnes multi domini

Aquesta tècnica (figura 1.12a) permet marcar isotòpicament només una regió o domini d'una proteïna gran de la qual es volen estudiar aspectes estructurals o dinàmics mitjançant RMN reduint significativament el soroll de fons generat en la RMN de proteïnes grans. La tècnica consisteix en partir la proteïna d'interès en dos dominis i fusionar cada domini amb la I_N o la I_C introduint les construccions en dos vectors diferents. Cada vector porta com a factor de selecció un gen de resistència a un antibiòtic diferent i presenten també un promotor diferent. Els dos vectors s'utilitzen per transformar la mateixa cèl·lula i se'n fa créixer un cultiu en medi marcat amb l'isòtop escollit. Quan el cultiu ha crescut, s'indueix l'expressió del domini que es vol estudiar per RMN, el qual s'expressarà juntament amb la I_N o la I_C . Una vegada produït el domini marcat amb l'isòtop, es recullen les cèl·lules i es resuspenen en un medi no marcat. Aleshores s'indueix l'expressió del segon domini que no estarà marcat isotòpicament i que s'ha fusionat a l'altre fragment de la proteïna. A mesura que es vagi produint el segon domini, es donarà el mecanisme de tall i unió i es formarà una proteïna amb un domini marcat i l'altre no (Jung et al., 2013; Muona et al., 2010; Volkmann and Iwai, 2010). Aquesta estratègia permet abordar l'estudi estructural i la dinàmica d'un domini no de forma aïllada, sinó en el context de la proteïna sencera.

Utilitzant trans inteïnes es pot també fer el marcatge d'un domini central d'una proteïna que presenti tres dominis (A-B-C). Aquesta tècnica (figura 1.12b) utilitza una trans inteïna a partir de dos punts de tall diferents. Per exemple *Npu DnaE* té la capacitat de ser activa tant quan la seva subunitat I_C presenta 36 (Npu_{C36}) que és com es troba en el seu estat natural com quan se li escurcen els primers 21 residus i forma una subunitat I_C de només 15 residus (Npu_{C15}). Les característiques d'aquesta inteïna com també els seus punts de tall seran explicats amb major detall a l'apartat 1.4 d'aquesta introducció. Per altra banda, les subunitats I_N presenten la resta de residus totals de la inteïna i són la Npu_{NAC36} i la Npu_{NAC15} . El primer pas a seguir és la construcció de dos vectors amb les construccions Npu_{C15} -B- Npu_{NAC36} i Npu_{C36} -C. Cada un d'aquest vector porta com a factor de selecció un gen de resistència a un antibiòtic diferent i presenten també un promotor diferent. Els dos vectors s'utilitzen per transformar la mateixa cèl·lula i seguint el mateix protocol que l'anterior (figura 1.11a) es marca només el domini B de la proteïna. El producte obtingut Npu_{C15} -B-C es purifica per una cromatografia d'afinitat. Per una altra banda s'expressa i es purifica mitjançant cromatografia d'afinitat el precursor N-terminal amb A- Npu_{NAC15} . Aleshores i de manera *in vitro* es barregen els 2 precursors A- Npu_{NAC15} i Npu_{C15} -B-C i es durà a terme el mecanisme de tall i unió obtenint-se d'aquesta manera la unió dels tres dominis A-B-C, on el domini B es trobarà marcat isotòpicament (Busche et al., 2009).

1. Introducció

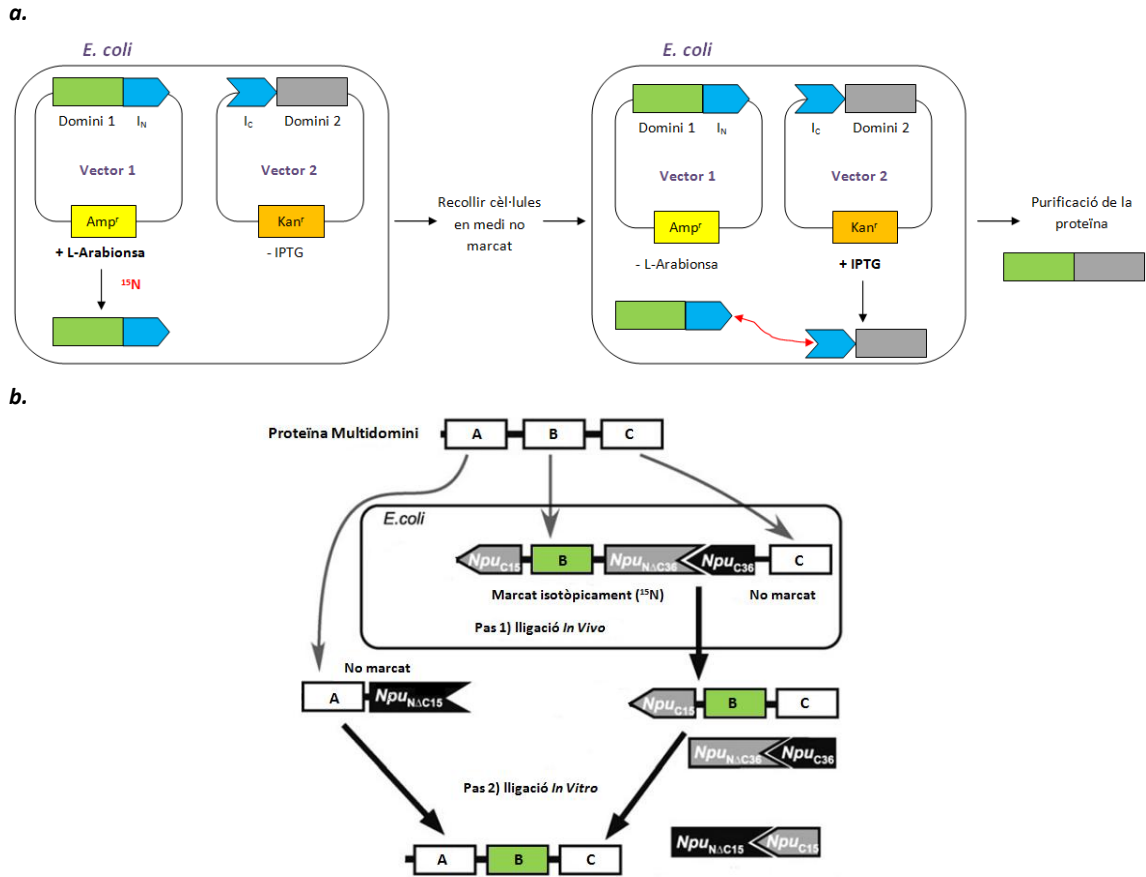


Fig. 1.12. (a) Esquema del marcatge isotòpic de segments de proteïnes multi domini *in vivo*. En taronja es mostren els segments no marcats isotòpicament i en verd els dominis marcats isotòpicament. (b) Marcatge isotòpic d'un domini intern d'una proteïna multidomini. En verd es mostra el domini marcat isotòpicament. Esquema extret de (Busche et al., 2009).

1.3.3 Regulació de l'activitat proteica a partir del mecanisme de tall i unió condicionat

La finalitat del mecanisme de tall i unió condicionat és regular l'activitat proteica *in vivo* d'una proteïna. Per això, és essencial que el mecanisme de tall i unió, tant si és en cis com en trans, pugui ser controlat per un estímul que provingui d'una altra molècula (figura 1.13) o d'un canvi de pH, temperatura, estat redox o de llum (Chen et al., 2012; Mootz, 2009; Topilina and Mills, 2014).

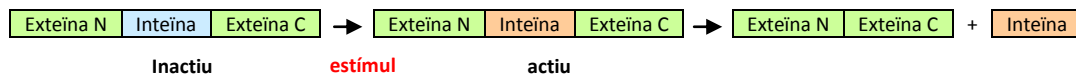


Fig. 1.13. Esquema del procés de regulació de l'activitat proteica a partir del mecanisme de tall i unió condicionat.

Per exemple, Perrimon i col·laboradors van desenvolupar una variant de la inteïna *Sce VMA* sensible a temperatura la qual va ser funcional tant en llevats com en *D. melanogaster* (Zeidler et al., 2004). De manera natural, hi ha inteïnes dependents de temperatura en les quals, el mecanisme de tall i unió es dona només a altes temperatures. Aquests tipus d'inteïnes són les que es troben en organismes termòfils (Cambon-Bonavita et al., 2000; Choi et al., 2006).

Altres investigadors han dissenyat inteïnes fotoactivables a partir de la fusió d'un grup protector que és fotosensible (Berrade et al., 2010; Vila-Perello et al., 2007). Mootz i col·laboradors van dissenyar una trans inteïna que indueix l'escissió C a partir d'irradiacions i canvis de pH (Binschik et al., 2011).

En altres investigacions s'ha aconseguit controlar el mecanisme de tall i unió controlant l'estat redox dels ponts disulfur. Sota condicions oxidants i amb l'adició d'agents reductors el mecanisme de tall i unió es veu facilitat per la reassociació dels fragments de la inteïna en trans (Shi and Muir, 2005). En el cas de la inteïna *Mma PolII* la seva activitat és depenent d'un pont disulfur intern i presenta diferents activitats en funció de l'estat redox de la cèl·lula o de si aquesta es troba en el periplasma o el citoplasma d'*E. coli* (Nicastri et al., 2013). Callahan i col·laboradors van dissenyar una variant de la inteïna *Ssp DnaE* la qual presentava una Cys -3 a l'exteïna N que funcionava com a trampa redox. En aquesta variant només es duia a terme l'escissió N sota condicions reductores en *E. coli*. Això va permetre purificar el precursor i provocar l'escissió N *in vivo* després de l'adició d'agents reductors (Callahan et al., 2011). Aquesta inteïna va ser utilitzada com a biosensor per determinar l'estat redox de la cèl·lula i com pot condicionar l'activitat de l'exteïna que porta unida la inteïna (Callahan et al., 2013).

1.3.4 Inteïnes com a biosensors

Les inteïnes poden ser utilitzades com a biosensors per tal de detectar (a) interaccions entre proteïnes (Kanno et al., 2009; Ozawa et al., 2000; Xu and Evans, 2005), (b) per a la localització i internalització de proteïnes (Kim et al., 2004; Zhang et al., 2011), (c) canvis en la metilació del DNA (Huang et al., 2012), (d) detecció de molècules petites (Gierach et al., 2012; Li et al., 2011), (e) l'activitat d'una proteasa (Kanno et al., 2007) o com ja s'ha avançat, (f) conèixer l'estat redox d'una cèl·lula (Callahan et al., 2013). Aquesta tècnica consisteix bàsicament en què un estímul provoqui el mecanisme de tall i unió i que aquest activi la proteïna *reporter* (figura 1.14). Algunes de les tècniques citades es descriuen a continuació:

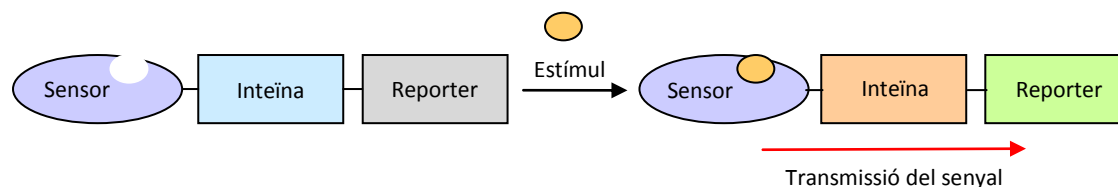


Fig. 1.14. Esquema general de l'utilització d'una inteïna com a biosensor.

1.3.4.1 Interacció de proteïnes

Per detectar interaccions proteïques emprant inteïnes es necessita una inteïna partida els fragments de la qual tinguin baixa afinitat d'associació. Per dissenyar aquests biosensors, es creen dues proteïnes de fusió, de les quals cada una portarà una proteïna de les quals se'n vol estudiar la interacció, un fragment de la inteïna partida i un fragment de la proteïna *reporter* (luciferasa o GFP).

1. Introducció

En el cas que, les proteïnes que s'estudien, interaccionen, s'induirà el procés de tall i unió ja que els dos fragments d'inteïna estaran molt propers i es reconstituirà la proteïna *reporter* i emetrà fluorescència o es detectarà l'activitat de la luciferasa. L'emissió de fluorescència o llum podrà ser monitoritzada i per tant es podran fer anàlisis quantitius de la interacció entre les proteïnes sotmeses a estudi (Ozawa et al., 2000).

1.3.4.2 Localització i internalització de proteïnes

En aquest cas caldrà unir la proteïna *reporter* (luciferasa o GFP) partida en dos fragments i unir cada un dels fragments a un fragment d'una inteïna partida amb elevada afinitat entre les dues subunitats. El fragment N-terminal és dissenyat per tal que vagi a nucli (o un altre compartiment desitjat) i el fragment C-terminal porta unit la proteïna d'interès per la qual es vol estudiar la localització, internalització o bé localitzar el seu receptor. Quan el fragment C-terminal es transloqui a nucli o un altre compartiment desitjat en funció de la proteïna que s'en vol estudiar la seva localització i/o internalització, es durà a terme el mecanisme de tall i unió i es reconstituirà la proteïna *reporter* i aquesta adquirirà fluorescència o llum. L'emissió de fluorescència o llum podrà ser monitoritzada i per tant es podran fer anàlisis quantitius de la internalització de la proteïna (Kim et al., 2004).

1.3.5 Requisits de les inteïnes per ser utilitzades com a eines biotecnològiques

Amb l'aparició de noves tècniques on s'utilitzen inteïnes i l'aplicació d'aquestes en diferents camps de la biologia molecular, fa que les inteïnes siguin cada vegada més estudiades i siguin una eina biotecnològica potencialment interessant. Des del seu descobriment fins ara, van apareixent noves inteïnes naturals en els tres regnes de la vida. A la vegada, algunes s'han dissenyat de manera artificial a partir de les inteïnes naturals. Això s'ha fet per tal de millorar les seves característiques natives i així intentar que siguin més eficients en les diferents tècniques en les quals s'han d'utilitzar. Algunes d'aquestes inteïnes artificials són la S11 de *Ssp GyrB* (Appleby et al., 2009) o una variant de la inteïna fúngica *PRP8* (Elleuche and Poggeler, 2007).

S'accepta que una inteïna en concret és potencialment interessant per ser utilitzada en qualsevol tècnica on s'utilitzen inteïnes quan compleix les següents característiques: és capaç de realitzar la reacció de manera ràpida a diferents rangs de temperatura, no apareixen de manera significativa reaccions secundàries en el mecanisme de tall i unió (l'escissió N i C), té bona solubilitat i estabilitat, és activa sota condicions desnaturalitzants i l'eficiència del mecanisme de tall i unió ha de ser independent de la naturalesa de l'exteïna (Volkman and Mootz, 2013). A la taula 1.2 es mostren les inteïnes partides més utilitzades en el camp de la biologia molecular:

1. Introducció

Fins a l'actualitat, una de les inteïnes que compleix millor aquests requisits és la trans inteïna *Npu DnaE*. Aquesta és la inteïna que s'ha utilitzat per a la realització d'aquesta tesi doctoral. Els estudis fets fins a l'actualitat es descriuen en el següent apartat d'aquesta introducció.

Taula 1.2. Inteïnes partides més utilitzades en el camp de la biologia molecular per al desenvolupament de diferents tècniques

	Mida		
	I _N	I _C	t _{1/2} (min)
Inteïnes naturals			
<i>Ssp DnaE</i>	123	36	35-175
<i>Npu DnaE</i>	102	36	1
Inteïnes artificials			
<i>Mtu RecA</i>	105	38	60-120
<i>Sce VMA</i>	184	55	6
<i>Ssp DnaB</i>	104	47	12

A la taula es mostra el nombre d'aminoàcids per cada fragment de la inteïna i el temps de vida mitja (t_{1/2}) calculat a partir de la constant de primer ordre (Vila-Perello and Muir, 2010). t_{1/2} és el temps el qual tarda en formar-se la meitat del producte de tall i unió en un assaig d'activitat de la inteïna.

1.4 LA INTEÏNA *Npu DnaE*

La inteïna partida *Npu DnaE* de *Nostoc punctiforme* PCC73102 forma part de la família *DnaE* d'inteïnes descrites en cianobacteris (Caspi et al., 2003) i pertany a la classe 1 de inteïnes tot i que no presenta la penúltima His en la regió G. Aquesta inteïna es troba formant part del precursor de la subunitat α de la DNA polimerasa III d'aquest organisme. La seva massa molecular és de 15832,97 Da i està composta per 138 residus, 102 dels quals pertanyen a la I_N i 36 a la I_C (Iwai et al., 2006). La seqüència aminoacídica de cada fragment es detalla a la figura 1.15:

Subunitat I_N:

CLSYETEILT	VEYGLLPICK	IVEKRIECTV	YSVDNNGNIY
10	20	30	40
TQPVAQWHDR	GEQEVFEYCL	EDGSLIRATK	DHKFMTVDGQ
50	60	70	80
MLPIDEIFER	ELDLMRVDNL	PN	
90	100		

Subunitat I_C:

MIKIATRKYL	GKQNVYDIGV	ERDHNFALKN	GFIASN
10	20	30	

Fig. 1.15. Seqüències aminoacídiques de les subunitats I_N i I_C de la *Npu DnaE*. En blau es mostren els residus pertanyents a la regió A, en taronja la regió N2, en vermell la regió B, en verd la regió F i en lila la regió G. No s'ha trobat cap regió que pugui correspondre a la regió N4 (Perler, 2002).

1. Introducció

Aquesta seqüència presenta una regió rica en residus hidrofòbics (residus 1 al 50 de la subunitat I_N), una regió rica en residus de càrrega negativa (residus 51 al 102 de la subunitat I_N) i per últim una regió corresponent a la subunitat I_C rica en residus de càrrega positiva. Aquesta disposició fa que la inteïna en global presenti un pI de 4.84 i, de manera separada, la subunitat I_N i la subunitat I_C presenten pIs de 4.39 i 9.70 respectivament (Artimo et al., 2012).

1.4.1 L'estructura tridimensional

L'estructura d'aquesta inteïna resolta per Oeemig i col·laboradors mitjançant RMN a partir de 20 conformacions d'energia mínima, es troba dipositada en el *protein data bank* (Codi PDB 2KEQ). La resolució de l'estructura d'aquesta inteïna, es va realitzar a partir d'una única cadena polipeptídica corresponent a la fusió de les subunitats I_N i I_C i amb la mutació C1A per tal d'evitar que la inteïna fos activa (figura 1.16). Els resultats obtinguts per RMN van mostrar que la inteïna es trobava composta bàsicament per catorze fulls β (46%) i dues hèlix α curtes (9%) que formen una proteïna en forma de disc amb un diàmetre d'entre 19 i 22 Å i de 14 Å de gruix amb un plegament tipus ferradura típic dels dominis HINT, dels quals forma part aquesta inteïna (figura 1.16) (Hall et al., 1997; Johnson et al., 2007). La regió connectora de les subunitats I_N i I_C forma un *loop* sense estructura secundària regular, la qual cosa fa pensar que aquests residus es podrien eliminar sense afectar la funció de la inteïna. En general, la inteïna presenta una estructura globular compacte on la regió més llarga desordenada identificada correspon al *loop* entre β12 i β13. (Oeemig et al., 2009). Cal remarcar també que malgrat presentar tres Cys en la seva seqüència, la inteïna no té cap pont disulfur.

Oeemig i col·laboradors van realitzar també experiments de relaxació nuclear de ¹⁵N. Aquests experiments permeten un bon coneixement de la mobilitat de l'esquelet de la proteïna (Kay et al., 1989; Luginbuhl et al., 1997). Amb aquests experiments es va poder veure quines eren les regions que presentaven una major mobilitat interna i van assumir que aquestes regions serien punts de tall potencials per tal de dissenyar nous punts de partició d'aquesta trans inteïna. Els residus que presentaven una major mobilitat van ser el 10, 40, 76, 78, 96-98, 102, 104, 120 i 132. Tots aquests residus es troben en loops excepte el residu 40. En el cas dels residus 96-98, 102 i 104 es troben localitzats en el loop que connecta les subunitats I_N i I_C, suggerint que en aquesta regió la mobilitat és alta (Oeemig et al., 2009). Amb l'identificació d'aquests residus, es van realitzar posteriorment el disseny de nous punts de partició de la inteïna.

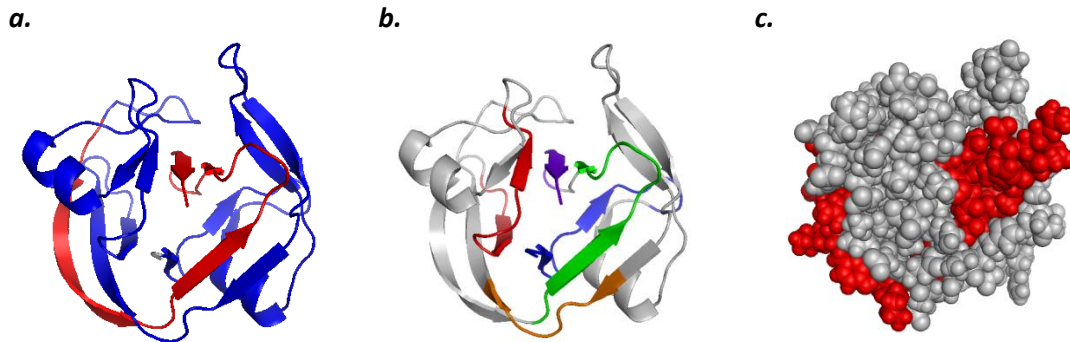


Fig. 1.16. Estructura tridimensional de la *Npu DnaE*. a) Representació esquemàtica de l'estructura tridimensional de la inteïna on en blau es mostra la subunitat I_N i en vermell la subunitat I_C . b) Representació esquemàtica de les diferents regions de la inteïna on en blau es mostren els residus pertanyents a la regió A, en taronja la regió N2, en vermell la regió B, en verd la regió F i en lila la regió G. Cal destacar que totes aquestes regions es troben molt properes entre si formant el centre actiu de la inteïna. c) Representació de la superfície electrostàtica de la inteïna on en gris es mostra la subunitat I_N i en vermell la subunitat I_C . Cal destacar que part de la subunitat I_C es troba majoritàriament enterrada a la regió més hidrofòbica de la inteïna.

1.4.2 Cinètica i afinitat d'unió de la inteïna

Npu DnaE és una de les inteïnes partides que presenten un $t_{1/2}$ més baix d'entre totes les inteïnes que es coneixen fins a l'actualitat. La seva constant de primer ordre (K_{obs}) a 37°C és de $1.1 \pm 0.2 \times 10^{-2} \text{ s}^{-1}$ que equival a un $t_{1/2}$ de 60s utilitzant *head protein D* (gpD) com a exteïna N i tioredoxina com a exteïna C (Zettler et al., 2009) quan la majoria d'inteïnes presenten un $t_{1/2}$ d'hores (Iwai et al., 2006). Aquest $t_{1/2}$ és molt inferior a altres inteïnes molt utilitzades en el camp de la biologia molecular com són la *Ssp DnaE* que presenta un $t_{1/2}$ de 35 min a 23°C (Nichols et al., 2003) o la *Sce VMA* amb un $t_{1/2}$ de 6 min a 25°C (Brenzel et al., 2006). Més recentment, s'ha descrit que aquesta inteïna presenta un $t_{1/2}$ de només 18.9s a 37°C i pH 7.2 utilitzant Ubiquitina com a exteïna N i SUMO com a exteïna C (Shah et al., 2012).

A partir dels experiments de relaxació nuclear de Ooemig i col·laboradors, s'han dissenyat nous punts de tall en aquesta inteïna i s'ha observat que el mecanisme de tall i unió és actiu quan es troba partida entre els residus 123 i 124 i entre els residus 131 i 132. En el cas de la inteïna partida en el residu 123 el $t_{1/2}$ observat és de 120 min (Aranko et al., 2009) i en el cas de la inteïna partida en el residu 131 el $t_{1/2}$ observat és de 220 min (Ooemig et al., 2009). Recentment, s'ha descobert que la inteïna també presenta activitat quan es troba partida entre els residus 12 i 13 amb un $t_{1/2}$ de 260 min i entre els residus 36 i 37 amb un $t_{1/2}$ de 18 min (Lee et al., 2012). El fet que aquesta inteïna continuï sent funcional en aquestes punts de partició permet obtenir fragments de I_C més curts encara que el salvatge. Descobrir que partint la inteïna *Npu DnaE* en altres punts i obtenir fragments de I_C funcionals (de 15 i 6 residus), ha permès sintetitzar artificialment aquests fragments i poder-los utilitzar en EPL, ja que, la subunitat I_C de *Npu DnaE* composta per 36 residus és massa llarga per ser sintetitzada químicament juntament amb una exteïna que incorpori les modificacions post-traduccionals desitjades (Aranko et al., 2009). Recentment, l'equip de Iwai i col·laboradors han

1. Introducció

buscat nous punts de partició de la inteïna. A partir de l'estructura tridimensional de la inteïna i dels experiments de relaxació per RMN es van buscar quins eren els nous possibles punts de tall. Els punts de tall seleccionats van ser entre els residus 11 i 12 (N11), 35 i 36 (N35), 124 i 125 (C14) com també, ja descrit anteriorment 131 i 132 (C6). En tots aquests punts de tall, la inteïna es va mostrar activa, però només en els casos N35 i sobretot en el cas de C14 l'eficiència de tall i unió era elevada i comparable amb la inteïna quan presenta el seu punt de tall i unió salvatge. En el cas de N11 i C6 l'eficiència era molt baixa i, a més, les escissions N i C eren significatives. Es van provar altres punts de tall per obtenir fragments de I_C més curts de 6 residus, no obstant, amb els fragments de I_C més curts de 6 residus només es produïa escissió N. Aquests resultats van suggerir que com a mínim es necessiten 6 residus del fragment I_C per dur-se a terme de forma correcta el mecanisme de tall i unió (Aranko et al., 2014).

Només les inteïnes *gp41-1*, *NrdJ-1*, *1MPDH-1* i *gp41-8* (Dassa et al., 2009) presenten uns $t_{1/2}$ inferiors a *Npu DnaE*. *Gp41-1* presenta un $t_{1/2}$ de 30s a 37°C, però passa a ser de només 3.8s a 45°C, *gp41-8* presenta un $t_{1/2}$ de 15s a 37°C, *NrdJ-1* presenta un $t_{1/2}$ de 7s a 37°C i *1MPDH* que presenta un $t_{1/2}$ de 8s a 37°C (Carvajal-Vallejos et al., 2012). Recentment s'ha descobert la trans inteïna *AceL-TerL* amb un $t_{1/2}$ de 7.2 min a 8°C, no obstant, presenta una forta dependència de la temperatura i a 37°C el $t_{1/2}$ ascendeix fins a 324 min (Thiel et al., 2014).

Des del punt de vista d'afinitat, els fragments de *Npu DnaE* presenten una elevada afinitat d'uníó, la qual es tradueix en una K_d de 2.9 ± 1.8 nM utilitzant ubiquitina com a exteïna N i SUMO com a exteïna C (Shah et al., 2011). Aquesta K_d és similar a l'observada en el complex FKBP12-rapamicina-FRB que presenta una K_d de 3nM (Chen et al., 1995) i que és àmpliament utilitzat per l'associació de proteïnes (Massoud et al., 2010). En el cas d'altres inteïnes partides, els fragments de la *Ssp DnaE* presenten també una elevada afinitat d'uníó però no tant com la de *Npu DnaE*, en aquest cas, la K_d observada és de 26 ± 12 nM utilitzant la GFP com a exteïna (Shi and Muir, 2005) i en el cas de la mini inteïna *Ssp DnaB*, la K_d d'aquesta augmenta ja fins a 900 nM (Appleby-Tagoe et al., 2011). Sue i col·laboradors van calcular l'afinitat que presenten els fragments de la *Npu DnaE* quan aquesta es troba partida entre els residus 36 i 37 de la I_N . La K_d en aquest cas va ser de 1239 ± 200 nM, un valor molt superior al que presenta aquesta inteïna quan es troba partida per la seva regió natural (Lee et al., 2012).

1.4.3 Robustesa de la inteïna

A més de presentar un dels $t_{1/2}$ més baixos descrits fins a l'actualitat i una afinitat d'uníó molt elevada, *Npu DnaE* presenta una elevada robustesa en dur a terme la reacció de tall i unió i una elevada capacitat de formar el producte independentment de l'exteïna. Entre 6 i 37°C la seva K_{obs} no es veu alterada i en presència d'un agent desnaturalitzant com és la urea, la K_{obs} no comença a

augmentar fins a 6M d'urea. Sota aquestes condicions, el producte de tall i unió format és del 65-70% respecte al 80-85% observat en condicions òptimes (Zettler et al., 2009). A més, recentment s'ha descrit que *Npu DnaE* presenta una T_m de 84 °C a pH 7, la qual cosa indica que és una proteïna bastant resistent a altes temperatures (Aranko et al., 2013).

1.4.4 Promiscuïtat de la inteïna

Malgrat la independència que presenta *Npu DnaE* a l'hora de dur a terme el tall i unió de l'exteïna que porta unida, estudis de mutagènesi dirigida demostren que la substitució de la Cys per una Ser a la posició +1 provoca una abolicció completa del procés de tall i unió i que els tres residus de l'extrem N-terminal de la C-exteïna han de ser Cys-Phe-Asn per tal d'obtenir un tall i unió eficient (Lockless and Muir, 2009; Shah et al., 2012). Substituint la Phe +2 per un Trp o Met i conservant la Cys +1 i Asn +3 es poden obtenir uns resultats en l'eficiència del mecanisme de tall i unió semblants als obtinguts amb el triplet Cys-Phe-Asn. Altres mutacions en aquests 3 residus empitjoren o bloquegen l'eficiència de la inteïna (Cheriyana et al., 2013).

En el cas de l'extrem N-terminal, la regió no requereix de gran especificitat i la inteïna accepta gran varietat de seqüències de les diferents exteïnes. No obstant, sí que és indispensable que el primer residu de la inteïna sigui una Cys (Cheriyana et al., 2013).

A més, una característica que presenta *Npu DnaE* juntament amb tota la família d'inteïnes partides *DnaE* és que l'escissió C només es pot produir en cas que hi hagi hagut primer l'escissió N excepte en el cas de la mini inteïna *Ssp DnaX* (Qi et al., 2011). Algunes inteïnes se'ls ha mutat la Cys 1 per una Ala per tal de suprimir l'escissió N sense interferir l'escissió C, en el cas d'aquestes inteïnes, aquesta mutació també suprimeix l'escissió C (Martin et al., 2001; Zettler et al., 2009). Només la mutació D118G és capaç de bloquejar el mecanisme de tall i unió en aquesta inteïna i induir l'escissió C sota condicions reductores i sense l'escissió N prèvia (Ramirez et al., 2013; Wood et al., 1999). Aquesta *Npu DnaE* amb la mutació D118G va ser utilitzada per elaborar un protocol de purificació de proteïnes extremadament ràpid basat en el fenomen de l'escissió C. Aquesta variant presenta una elevada velocitat de formació de l'escissió C en presència d'agents reductors, on en només 30 min s'ha format el 90% de l'escissió a 22°C. L'escissió C, és a la vegada independent del residu +1, ja que es va observar en tots els aminoàcids que ocupaven aquesta posició excepte per Pro i Thr. L'escissió C es pot inhibir en aquesta variant mitjançant $ZnCl_2$ (0.5 mM) sota condicions no reductores (Guan et al., 2013).

1.4.5 El procés associatiu dels fragments I_N i I_C

El procés pel qual els dos fragments de inteïna associen no és del tot conegut. No obstant, s'ha postulat que les interaccions iòniques tenen un paper clau en l'associació de les inteïnes partides (Dassa et al., 2007). Per altra banda i utilitzant la *Npu DnaE* com a model, es va voler demostrar com les forces electrostàtiques entre els fragments de la inteïna participaven també en el procés associatiu. Per estudiar el paper d'aquestes forces electrostàtiques es van buscar basant-se en l'esctructura de la inteïna, totes les parelles de residus + i - que es troben en les dues subunitats i es van crear mutants dels fragments de I_N i I_C. Per una banda, els residus àcids del fragment I_N seleccionats es van mutar a residus bàsics i per l'altra banda, els residus bàsics del fragment I_C seleccionats es van mutar a residus àcids. Es van fer els assajos d'activitat amb totes les combinacions possibles i es va veure que només el mecanisme de tall i unió era actiu quan hi havia complementarietat de forces electrostàtiques (figura 1.17). És a dir, es va observar activitat quan el fragment I_N amb càrrega positiva neta es barrejava amb el fragment I_C de càrrega negativa neta i quan el fragment I_N amb càrrega negativa neta es barrejava amb el fragment I_C de càrrega positiva neta (inteïna salvatge).

Quan els dos fragments de la inteïna presentaven càrrega negativa o positiva neta no s'observava activitat. A més, l'afinitat que presentaven el fragment I_N de càrrega positiva neta amb el fragment de I_C de càrrega positiva neta era semblant a l'observada en els dos fragments salvatges (Shah et al., 2011).

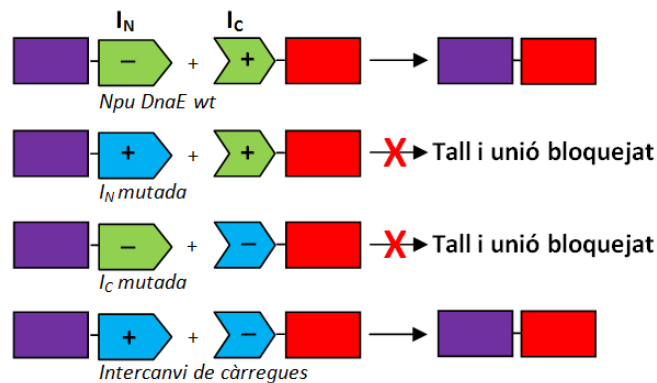


Fig. 1.17. Esquema de l'especificitat en l'intercanvi de càrregues en l'activitat de *Npu DnaE*. Esquema reproduït de (Shah et al., 2011).

Recentment, l'equip de Muir i col·laboradors han mostrat com el fragment I_N de la inteïna *Npu DnaE* es troba parcialment plegat, mentre que el fragment I_C es troba completament desplegat. Mitjançant RMN van observar que el fragment I_N presentava una primera meitat (NpuN₁) corresponent als residus 1 a 50 amb una elevada concentració de residus hidrofòbics i una baixa càrrega neta mitjana. Això indicava que aquesta primera meitat es trobava força estructurada. Per altra banda, l'altra meitat del fragment (NpuN₂) corresponent als residus 51 a 102 presentava una elevada quantitat de

1. Introducció

residus aniónics i una alta càrrega neta mitjana que faria que aquesta part es trobés desestructurada (Uversky et al., 2000). A més, NpuN₂ i el fragment I_C presentaven una elevada complementarietat electrostàtica. Estudiant la interacció que presentaven els diferents fragments mitjançant cromatografia per exclusió de mida i làser multi àngle de dispersió de llum (SEC-MALS), es va veure que no hi havia interaccions entre NpuN₁ i el fragment I_C i que en canvi, sí que n'hi havia entre NpuN₂ i el fragment I_C. Tornant als estudis de RMN, van veure com NpuN₂ sola no presentava estructura, mentre que, quan aquesta es trobava en presència del fragment I_C adoptava estructures d'hèlix α i conformacions β en les mateixes regions que es van observar en la inteïna nativa. Amb els 13 primers residus del fragment I_C juntament amb NpuN₂ (residus 51 a 115) es va observar que la proteïna es trobava estructurada, indicant que els 13 primers residus de la I_C eren suficients pel plegament de NpuN₂. Aleshores, es va fer un seguiment del procés associatiu mitjançant fluorescència. Per dur a terme això, es va mutar el triptòfan 47 present a NpuN₁ per una tirosina i la tirosina 58 present a NpuN₂ es va mutar per un triptòfan. Aquestes mutacions no van afectar significativament la capacitat d'associació ni l'activitat de la inteïna. Amb els estudis previs mitjançant RMN i el seguiment del procés associatiu mitjançant fluorescència es va acabar proposant que el procés associatiu es dona per un mecanisme de captura i col·lapse on en primer lloc es produeix una interacció (captura) per part dels fragments I_C i NpuN₂. Aquesta captura té lloc gràcies a la complementarietat electrostàtica que presenten els 2 fragments. Després d'aquesta captura, les regions flexibles de NpuN₂ es col·lapsen en un estat ordenat que quedaria estabilitzat per les interaccions hidrofòbiques entre I_C i la regió estructurada pertanyent a NpuN₁. A partir d'aquest col·lapse i estabilització de la inteïna, s'assoleix la conformació nativa i aquesta adquiriria la capacitat per dur a terme el tall i unió de les exteïnes (Shah et al., 2013b). Aquest mecanisme de captura i col·lapse es troba esquematitzat a la figura 1.18.

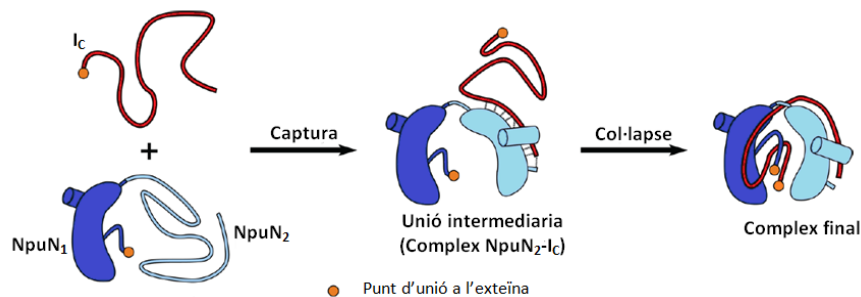


Fig. 1.18. Esquema del mecanisme de captura i col·lapse proposat per Muir i col·laboradors. El fragment N-terminal de la I_C és capturat per la segona part de la I_N (NpuN₂) compactant aquesta part en una regió estructurada. Aquest intermediari després col·lapse el que resta de la I_C en la primera part de la I_N (NpuN₁) assolint la conformació nativa. Esquema reproduït de (Shah et al., 2013b).

2. Objectius

OBJECTIUS

L'estudi de les inteïnes ha permès desenvolupar noves tècniques en el camp de la biotecnologia i biologia molecular basades en l'activitat que presenten aquest tipus de proteïnes. De totes les inteïnes descobertes fins a l'actualitat, una de les més interessants i utilitzada és la *trans* inteïna *Npu DnaE* de *Nostoc punctiforme*. Aquesta inteïna, tal i com s'ha descrit en la introducció, és una de les que presenta una velocitat de reacció més elevada, una bona solubilitat i una considerable estabilitat i és activa independentment de l'exteïna que porta unida.

En aquesta tesi es va proposar realitzar un seguit d'estudis relacionats amb l'estabilitat, la cinètica i l'estructura per tal d'entendre millor aspectes mecanístics, el procés d'associació dels fragments I_N i I_C , i l'efecte d'aquesta associació en el plegament global de la inteïna. Aquest conjunt d'estudis hauria de permetre aprofundir la comprensió de les excel·lents qualitats que presenta aquesta inteïna i obrir la possibilitat de desenvolupar noves aplicacions biotecnològiques utilitzant aquesta inteïna. Per tant, els objectius concrets d'aquesta tesi són:

1. Dissenyar i construir uns vectors per a la expressió i producció de la inteïna *Npu DnaE* activa tant quan aquesta es troba formant part d'una única cadena polipeptídica (*cis*), com en forma de subunitat I_N i subunitat I_C separades (*trans*) i de la subunitat I_N aïllada.
2. Dissenyar i optimitzar protocols de purificació de la inteïna *Npu DnaE* activa en les diferents versions esmentades anteriorment.
3. Caracteritzar l'estabilitat tèrmica a diferents valors de pH que presenten les diferents variants de *Npu DnaE*.
4. Estudiar l'efecte de la substitució per Ser en les posicions C1 i C+1 en l'activitat de la inteïna.
5. Dissenyar i construir vectors per l'expressió i purificació de la subunitat I_N amb fragments de I_C de longitud creixent corresponent a diferents cadenes β d'aquesta subunitat per tal d'obtenir les variants $I_{N|C_sc114}$, $I_{N|C_sc123}$ i $I_{N|C_sc131}$.
6. Caracteritzar per tècniques espectroscòpiques el plegament que presenten la subunitat I_N aïllada i les diferents variants $I_{N|C_sc114}$, $I_{N|C_sc123}$ i $I_{N|C_sc131}$ per tal d'estudiar l'efecte que té la subunitat I_C en el plegament global de la inteïna.

3. Material i Mètodes

3.1 MATERIAL BIOLÒGIC

3.1.1 Soques bacterianes

En la realització d'aquest treball s'han utilitzat les següents soques de l'espècie *Escherichia coli*, el genotip dels quals es troba detallat a continuació:

3.1.1.1 *E. coli* DH5 α

Aquesta soca ha estat utilitzada per a l'obtenció de les diferents construccions realitzades a partir dels vectors pET22b(+), pRSET/EmGFP, pSKBAD2, pSKDuet01 i pBADm. S'ha utilitzat aquesta soca perquè presenta una elevada taxa de transformabilitat, a més del caràcter genotípic *hsdR17*(rk-, mk+), que li proporciona la capacitat de metilar el DNA propi i alhora l'incapacita de destruir qualsevol DNA forà. Aquesta característica és essencial per tal d'obtenir el DNA que serà utilitzat com a motlle en les reaccions de mutagènesi dirigida per oligonucleòtid, el qual ha d'estar metilat en la seqüència reconeguda per l'enzim de restricció *DpnI*. (Hanahan, 1983)

Genotip: F-, ϕ 80 Δ lacZ Δ M15, Δ (lacZYA-argF), U169, *deoR*, *recA1*, *endA1*, *hsdR17*(rk-, mk+), *phoA*, *supE44*, λ -, *thi-1*, *gyrA96*, *relA1*

3.1.1.2 *E. coli* BL21(DE3)

Aquesta soca ha estat utilitzada per a l'expressió de les diferents construccions obtingudes a partir dels vectors pET22b(+), pRSET/EmGFP, pSKBAD2, pSKDuet01 i pBADm. Ha estat obtinguda de Novagen (EUA). Presenta la particularitat que en el gen *int* del genoma bacterià s'hi ha inserit el profag DE3 (derivat del fag λ). Conseqüentment, el gen *int* bacterià resta inactiu. La inactivació del gen *int* fa que el profag DE3 no es pugui escindir del genoma ni tornar-se a integrar, de manera que resta com a lisogen estable. El profag DE3 consisteix en un fragment de DNA que conté el promotor *lacUV5*, un fragment inicial del gen *lacZ* i el gen que codifica per la RNA polimerasa de T7. En aquestes condicions, l'únic promotor capaç de conduir la transcripció de la RNA polimerasa de T7 és el promotor *lacUV5*, el qual és induïble amb IPTG. En afegir IPTG a un cultiu d'aquesta soca en fase exponencial de creixement, es transcriu la RNA polimerasa de T7, la qual transcriu el gen que ha estat clonat darrera del promotor reconegut per aquesta polimerasa en els plasmidis pET22b(+), pRSET/EmGFP, pSKBAD2, pSKDuet01. (Studier and Moffatt, 1986)

Genotip: F-, *ompT*-, *hsdS*(rB- mB-), *dcm*+, *gal*, λ (DE3), *lon*-

3.1.2 Vectors procariotes

3.1.2.1 Vectors derivats de pRSET/EmGFP

Per al clonatge i expressió de la inteïna *Npu DnaE* descrita en aquest treball, s'ha utilitzat el vector pRSET/EmGFP procedent de Invitrogen (USA). (Schoepfer, 1993)

El vector presenta un origen de replicació d'*E. coli* i el promotor T7 reconegut per la RNA polimerasa del T7. Seguit del promotor hi ha una seqüència que codifica per 6 histidines que permeten facilitar la purificació de la proteïna a partir de columnes d'afinitat Ni-NTA. El vector presenta també la seqüència d'EmGFP, una proteïna fluorescent, seguit d'una seqüència terminadora. Tant en l'inici del gen de EmGFP i en la seqüència terminadora hi ha la zona de clonatge múltiple. Com a mecanisme de selecció el vector és portador del gen de resistència a l'ampicil·lina. El vector no és portador de l'operador lac. D'aquesta manera, la proteïna clonada en el vector s'expressa de forma constitutiva en la cèl·lula sempre i quan aquesta presenti la RNA polimerasa de T7.

3.1.2.2 Vectors derivats de pET22b(+)

Per a l'expressió de I_N i I_C associades descrit en aquest treball s'ha utilitzat el vector pET22b(+) procedent de Novagene (EUA) (Better et al., 1988).

Aquest vector presenta un origen de replicació d' *E. coli* (ori) i un promotor del tipus T7lac i és portador del gen de resistència a l'ampicil·lina. El vector també conté la seqüència pelB de 63 pb que codifica per un pèptid senyal d'exportació perioplasmàtica. Darrera d'aquesta seqüència presenta una zona policonnectora amb deu possibles dianes, on es poden inserir els gens forans, expressant-se la proteïna per la qual codifiquen fusionada al pèptid pelB. Aquest pèptid està dissenyat per exportar les proteïnes heteròlogues al periplasma de la cèl·lula bacteriana.

3.1.2.3 pSKDuet01

Per al clonatge i expressió de la I_N de *Npu DnaE* descrita en aquest treball, s'ha utilitzat el vector pSKDuet01 procedent de Addgene (USA). (Iwai et al., 2006)

El vector presenta un origen de replicació d'*E. coli* i el promotor T7 reconegut per la RNA polimerasa del T7 induïble per IPTG. Seguit del promotor hi ha una seqüència codificant de 6 histidines, el domini GB1 (domini B1 d'una proteïna que uneix les IgG) i la seqüència de la I_N de la inteïna *Npu DnaE*. A la vegada, el vector és portador del gen de resistència a la kanamicina. El sistema d'expressió d'aquest vector segueix el mateix model que el descrit en el vector pET22b(+).

3.1.2.4 pSKBAD2

Per al clonatge i expressió de la I_C de *Npu DnaE* descrita en aquest treball, s'ha utilitzat el vector pSKBAD2 procedent també de Addgene (USA). (Iwai et al., 2006)

Aquest vector presenta un origen de replicació d'*E. coli* i el promotor araBAD, promotor de l'operó d'arabinosa, induïble per arabinosa. Seguit d'aquest promotor hi ha la seqüència de la I_C de la inteïna *Npu DnaE* fusionada al domini GB1. El vector presenta també el gen de resistència a l'ampicil·lina.

3.1.2.5 Vectors derivats de pBADM

Per al clonatge i expressió de les diferents I_NI_C_scX descrites en aquest treball, s'ha utilitzat una sèrie de vectors procedents de pBADM procedents de EMBL (Alemanya).

Aquesta sèrie de vectors dissenyats per facilitar el procés d'expressió i purificació de proteïnes, presenten un origen de replicació d'*E. coli* i el promotor araBAD, promotor de l'operó d'arabinosa induïble per arabinosa. Seguit d'aquest hi ha una proteïna de solubilització fusionada a una seqüència codificant de 6 histidines. Tot seguit hi ha una diana reconeguda per la proteasa TEV (*Tobacco etch virus*) i una seqüència policonectora amb diferents dianes on es poden inserir els gens forans. La proteïna d'interès serà expressada juntament amb la diana reconeguda per la TEV i la proteïna de solubilització que contindrà una cadena rica en histidines per facilitar-ne la purificació a través d'una columna de Ni-NTA. Aquesta sèrie de vectors, tots ells portadors del gen de resistència a l'ampicil·lina i induïbles per arabinosa, és composta per 4 vectors diferents en funció de la proteïna de solubilització que contenen:

- pBADM-20(+): Conté la TrxA (*Thioredoxin*)
- pBADM-30(+): Conté la GST (*Glutathione S-Transferase*)
- pBADM-41(+): Conté la MBP (*Maltose Binding Protein*)
- pBADM-60(+): Conté la NusA

3.2 MÈTODES MICROBIOLÒGICS

3.2.1 Medis de cultiu

Pel creixement i propagació de les diferents soques bacterianes s'han utilitzat els medis de cultiu Luria-Bertrani (LB) i medi mínim. El medi LB es tracte d'un medi ric d'ús comú i es pot preparar tant per utilitzar-lo en cultius líquids o sòlids. El medi mínim és utilitzat només per la producció de proteïnes marcades amb ¹⁵N i ¹³C. La composició per un litre de cada un d'ells és la següent:

3.2.1.1 Medi LB

Triptona	10 g
Extracte de llevat	5 g
NaCl	10 g

Enrasar a un volum final de 1 litre amb aigua destil·lada. Cal ajustar el pH a 7.2 amb NaOH, esterilitzant posteriorment a l'autoclau durant 20 min a 121°C. Per l'obtenció de medi sòlid per plaques de cultiu s'afegeixen 15 g d'agar per litre de medi (1.5% final).

3.2.1.2 Medi mínim

Sobre 877 ml d'aigua milliQ autoclavada i en condicions estèrils afegir seguint l'ordre següent:

MgSO ₄ 1M	1 ml
CaCl ₂ 1M	200 µl
Hidroclorur de tiamina 40mg/ml	1 ml
Elements traça	10 ml
10XM9	100 ml
Glucosa (Font ¹³ C)	2 g (dissolts en 5 ml d'aigua milliQ)
NH ₄ HCl (Font ¹⁵ N)	0.5 g (dissolts en 5 ml d'aigua milliQ)
Medi LB	1 ml

Preparació elements traça: 800 ml d'aigua milliQ + 5 g de Na₂EDTA + 0.5 g de FeCl₃ + 0.05 g de ZnCl₂ + 0.01 g de CuCl₂ + 0.01 g de CoCl₂ + 0.01 g de H₃BO₃ + 1.6 g de MnCl₂·6H₂O, ajustar a pH 7 després de l'addició de cada un, enrasar fins a 1 l, autoclavar i guardar a 4°C

Preparació de 100 ml de 10XM9: 6 g de Na₂HPO₄ + 3 g de KH₂PO₄ + 0.5 g de NaCl, enrasar fins a 100 ml, ajustar a pH 7 – 7.4 amb NaOH i autoclavar.

3.2.2 Antibiòtics

Quan es treballa amb soques transformades amb vectors, els medis de cultiu s'han de complementar amb els antibiòtics pels quals el vector confereix resistència. D'aquesta manera creixen només les cèl·lules portadores del vector i la soca no perd informació genètica de la qual és portadora. En la complementació dels medis es parteix d'una solució mare esterilitzada d'antibiòtic. Si cal complementar medi sòlid, l'antibiòtic s'afegeix al medi ja fred, just abans de distribuir-lo a les plaques.

L'antibiòtic de selecció de les cèl·lules que contenen el plasmidi pRSET/EmGFP, pET22b(+), pSKBAD2 i pBADM és l'ampicil·lina, mentre que en el cas de pSKDuet01 és la kanamicina. Es parteix d'una solució mare de 100mg/ml en aigua destil·lada esterilitzada per filtració (0,22 µm de diàmetre de porus) i es guarda alíquotada a -20°C. Per al seu ús, la solució mare es descongela i s'addiciona al medi en condicions estèrils. La concentració de treball de l'antibiòtic en els diferents medis de cultiu ha estat de 100 µg/ml.

3.2.3 Manteniment de les soques

Per al manteniment a curt termini de les soques de *E. coli*, portadores o no de plasmidis, es cultiven en plaques de Petri amb medi LB i es guarden a 4°C en posició invertida. Les plaques han d'estar etiquetades convenientment (soca, plasmidi i data de sembra) i seques (si és necessari es poden deixar obertes dins de la cabina de flux laminar per tal d'assecar-les). Un cop sembrades, les plaques s'incuben en posició invertida a 37°C o/n fins que les colònies siguin visibles i després es guarden a 4°C protegides amb parafilm.

El procediment dut a terme per a l'obtenció dels glicerinat consisteix en inocular una colònia de la soca desitjada en 10 ml de medi LB, incubar a 37°C en agitació (250 rpm) o/n, prendre una alíquota de 850 µl del cultiu, transferir-la a un tub Eppendorf estèril i afegir 150 µl d'una solució de glicerol 80% esterilitzada a l'autoclau. S'homogeneïtza i es guarda a -80°C.

Per inocular un cultiu a partir d'un d'aquests estocs congelats, es toca amb una nansa de Kolle, esterilitzada a la flama, la superfície del glicerinat (superficialment descongelat) i es realitza una estria en medi sòlid. Posteriorment es selecciona una colònia i s'inocula en un nou medi líquid. El glicerinat es retorna al congelador sense que s'hagi deixat descongelar completament.

3.2.4 Creixement dels cultius

Tot el material que s'utilitza per a la manipulació de microorganismes ha de ser estèril. El material i la majoria de reactius s'esterilitzen a l'autoclau amb l'excepció dels antibiòtics, que s'esterilitzen per filtració a través d'un filtre de 0.22 µm. En general, les condicions d'esterilització a l'autoclau són de 20 min a 121°C.

Els cultius de menys de 10 ml s'han preparat sempre a partir d'una colònia única aïllada en placa de Petri o bé d'un glicerinat. Per a cultius de volum superior sempre s'ha fet créixer abans un pre-inòcul de volum adequat, referit normalment com a cultiu *overnight* (o/n), per després afegir-lo al medi de cultiu a una relació 1:100 (inòcul:medi). El temps d'incubació dels cultius o/n ha estat sempre entre 12 i 15 h. No s'han realitzat incubacions de més durada, ja que pot veure's reduïda la població de cèl·lules portadores del vector d'interès, en favor de la població de cèl·lules sense vector. El volum del medi de cultiu ha estat sempre com a màxim 1/5 part de la capacitat del flascó. Tots els cultius s'han incubat a 18 o 37°C i en agitació constant entre 180 i 250 rpm.

Totes les manipulacions s'han de fer a la cabina de flux laminar. Aquesta s'esterilitza per irradiació durant 10 min amb llum ultraviolada i mentre no s'atura el flux es pot considerar estèril. També es convenient, si la cabina s'ha utilitzat moltes vegades, rentar la superfície de treball amb etanol 70%.

També és convenient treballar amb la flama encesa i flamejar qualsevol tub, pot o ampolla que s'obri. La nansa de Kolle s'esterilitza flamejant-la al bunsen i la de Digrafsky, flamejant-la amb l'ajut d'etanol 95%.

En el cas dels cultius en placa la durada de la incubació va entre 10 i 16 hores. Cal observar la presència de colònies clares evitant la formació de colònies satèl·lit. Aquestes apareixen després de períodes perllongats d'incubació i són conseqüència de processos de degradació de l'antibiòtic al voltant de les colònies portadores del vector.

3.3 TÈCNIQUES DE DNA RECOMBINANT

3.3.1 Manipulació del DNA

A l'hora de treballar amb àcids nucleics és important l'esterilitat del material i la cura amb què ha de treballar qui els manipula, per tal d'evitar contaminacions amb nucleases que podrien degradar la mostra o amb altres fragments de DNA que podrien donar lloc a resultats no desitjats. Per això, cal esterilitzar a l'autoclau les puntes de pipeta, tubs Eppendorf i qualsevol altre material que hagi de tenir un contacte directe amb el DNA. Pel mateix motiu també cal autoclavar totes les solucions que tinguin contacte amb el DNA i si això no és possible, cal esterilitzar-les per filtració. Així mateix, és recomanable manipular les solucions amb guants i en determinades ocasions es recomana l'ús de quelants d'ions divalents com l'EDTA que impossibilitin l'acció de les nucleases.

3.3.2 Obtenció de DNA plasmídic

L'obtenció de DNA plasmídic o minipreparacions de DNA a partir de les soques bacterianes a les quals ha estat transformat, permeten obtenir de manera ràpida DNA plasmídic. Aquest DNA és útil per a realitzar posteriors digestions amb enzims de restricció, clonatges, lligacions, PCR, seqüenciacions i transformacions.

En aquest treball s'ha utilitzat un Kit comercial, *QIAprep® Spin Miniprep Kit* (Qiagen), seguint el protocol següent:

Transferir 1.5 ml d'un cultiu de nit de cèl·lules transformades amb el vector d'interès a un tub Eppendorf. Centrifugar a 15000 rpm durant 30 s i eliminar el sobrenedant per aspiració. Repetir el procés tres vegades. Resuspendre el sediment amb 250 µl de solució de resuspensió. En aquest pas es produeix el trencament de les cèl·lules i l'eliminació del RNA. Aplicar 250 µl de solució de lisi cel·lular. La solució esdevindrà viscosa i lleugerament clara per la desnaturalització del DNA. Afegir 350 µl de solució neutralitzadora. La solució es tornarà tèrbola i blanquinosa a causa de la precipitació del DNA genòmic. Centrifugar 10 min a 15000 rpm. Aplicar el sobrenedant a les

columnnes d'adsorció de sílica i centrifugar durant 60 s a 15000 rpm. La sílica de les columnnes reté el DNA. Rentar el DNA amb 0.75 ml de solució de rentat que conté etanol. Es fan dos centrifugacions de 60 s per eliminar restes d'etanol. Assecar durant 15 min a 37°C per acabar d'eliminar l'etanol, el qual podria interferir en les reaccions enzimàtiques posteriors. S'elueix el DNA amb 50 µl de solució d'elució centrifugant 1 min a 15000 rpm. El DNA purificat es pot guardar a -20°C.

3.3.3 Electroforesi de DNA en gel d'agarosa

Aquesta tècnica, s'ha utilitzat per analitzar preparacions de DNA plasmídic, productes de digestió amb enzims de restricció i productes de PCR, a banda de ser l'etapa preliminar per a la purificació de fragments de DNA.

Per a la preparació dels gels s'ha utilitzat agarosa de baixa electroendosmosi (Ecogen, Espanya). Pel què fa al percentatge d'agarosa per a la preparació del gel, ha oscil·lat sempre entre un 0.8% i un 1.2%, en funció de la mida dels fragments que es volen separar. Com més gran sigui el fragment de DNA menys percentatge cal utilitzar d'agarosa. El tampó d'electroforesi utilitzat és el TAE (Tris-Acetate 40 mM, EDTA 2 mM, pH 8.0). En aquestes condicions el DNA es mou en base a la seva càrrega i se separa en funció de la longitud i conformació de les molècules. Les molècules de DNA es visualitzen tenyint el gel amb un agent que s'intercala entre les bases del DNA (Bromur d'etidi, afegint al gel a una concentració final de 0.5 µg/ml) i que emet fluorescència quan és transil·luminat amb llum UV.

3.3.4 Purificació de fragments de DNA a partir de gels d'agarosa

Per dur a terme la purificació de fragments de DNA s'ha utilitzat el Kit comercial QIAquick® Gel Extraction Kit (Qiagen, Alemanya), que permet purificar fragments entre 70 pb i 10 Kb, seguint el protocol de la casa comercial.

Tallar la banda de DNA minimitzant al màxim la mida del fragment escindit, posar la banda de gel en un tub Eppendorf i pesar-la. Solubilitzar el fragment amb 300 µl de tampó de solubilització per cada 100 mg de gel. Incubar 15 min a 50°C – 60°C o fins que el fragment de gel estigui totalment dissolt, carregar a la columna d'adsorció de sílica proporcionada per la casa comercial amb la solució i centrifugar 60 s a 15000 rpm, eliminar el sobrenedant i addicionar 750 µl de tampó de rentat (conté etanol). Centrifugar dues vegades a 15000 rpm durant 60 s, evaporar les restes d'etanol mantenint 15 min a 37°C i eluir el DNA amb 50 µl de tampó d'elució. El DNA purificat es pot guardar a -20°C.

3.3.5 Estimació espectrofotomètrica de concentracions de DNA

Per a fer una estimació de la quantitat de DNA present en una dissolució, s'ha mesurat l'absorbància a 260 nm d'una dilució de la solució de DNA de la qual es vol determinar la concentració. S'accepta

de forma general que una unitat de densitat òptica a 260 nm equival a una concentració de 50 µg/ml de DNA de doble cadena i a 20 µg/ml d'un oligonucleòtid.

3.3.6 Utilització d'enzims que modifiquen el DNA

3.3.6.1 Tractament amb enzims de restricció

Durant aquest treball s'han utilitzat enzims de restricció de tipus II per obtenir fragments de DNA amb extrems cohesius, els quals es poden purificar i unir per donar lloc a molècules de DNA quimèriques.

L'activitat dels enzims de restricció i la seva especificitat es troba directament relacionada amb el pH, la força iònica i la temperatura a la qual la reacció es duu a terme. Així, les condicions òptimes d'incubació per a cada digestió dependran dels requeriments específics de l'enzim a fi d'assolir el seu màxim d'activitat i un correcte reconeixement de la seva diana específica. A continuació es detallen les condicions generals que s'han emprat:

El volum d'enzim de restricció addicionat a la reacció no ha sigut mai superior al 10% del volum final de la digestió. Això és degut a què els enzims de restricció són subministrats en solucions al 50% de glicerol i per a molts d'ells, una concentració de glicerol superior al 5% en la barreja de reacció els fa perdre l'especificitat de reconeixement. Aquesta pèrdua d'especificitat també es pot manifestar, si la concentració d'enzim respecte la de DNA és massa elevada (> 25 U/µg de DNA).

Les barreges de restricció s'han incubat a 37°C entre 1 i 3 h.

En el cas de reaccions amb dos enzims de restricció diferents, l'addició dels enzims era simultània quan tots dos presentaven activitat òptima en les mateixes condicions, és a dir, en el mateix amortidor. Una vegada realitzada la digestió caldrà separar els fragments obtinguts mitjançant una electroforesi en gel d'agarosa seguida d'una purificació d'aquests fragments.

3.3.6.2 Reacció de lligament amb la DNA lligasa de T4

Barrejar els fragments de DNA que es volen lligar i que prèviament s'hauran purificat d'un gel d'agarosa amb el volum necessari d'amortidor de la DNA lligasa de T4 (Roche, Alemanya). Aquest amortidor el proporciona concentrat deu vegades la casa comercial juntament amb l'enzim, i és recomanable guardar-lo en alíquotes de 5-10 µl a -20°C. En el cas d'extrems cohesius és aconsellable utilitzar una relació entre vector i inserit de 1:4.

Afegir la lligasa de DNA de T4 (Roche) en una proporció de 0.1 - 0.5 unitats per µg de DNA. Incubar la reacció de lligament a 18°C durant 15 - 20 h. Si es treballa amb concentracions més elevades de

l·ligasa (fins a 5 U/ μ g de DNA), es pot optar, sobretot en l·ligaments d'extremes cohesius, per incubacions de 2 - 4 h a temperatura ambient, cosa que redueix el risc de degradació dels components de la reacció.

En el cas d'un l·ligament vector-inserit que no requereixi manipulacions posteriors, aquesta mescla pot utilitzar-se directament per a transformar cèl·lules competents. En el cas que no es pensi utilitzar immediatament, es pot guardar a -20°C.

3.3.7 Mutagènesi dirigida per oligonucleòtid: Mètode “QuikChange™”

El mètode *QuikChange*, desenvolupat per Stratagene (EUA), permet introduir una mutació específica en una seqüència de DNA coneguda, utilitzant DNA de doble cadena com a motlle. El procés consta de dues etapes: una amplificació no exponencial amb un termociclador per tal de produir DNA mutat i una digestió amb l'enzim de restricció *DpnI* per tal d'eliminar el DNA parental.

Per a aquest mètode és necessari un vector de doble cadena portador del gen a modificar i dos encebadors sintètics que continguin la seva seqüència amb la mutació desitjada. Aquests dos encebadors han de ser complementaris entre ells i d'aquesta manera es generen, utilitzant la reacció en cadena de la polimerasa, plasmidis portadors dels gens amb la mutació desitjada. El producte és tractat amb l'enzim de restricció *DpnI*, la funció del qual és digerir el DNA que contingui una cadena parental (metilat o hemimetilat) i així seleccionar el DNA sintetitzat que conté la mutació (no metilat). *DpnI* és un enzim de restricció de tipus II que reconeix la diana GATC únicament quan la guanina es troba metilada a la posició 5'. La reacció de digestió es realitza a 37°C 1h. El producte digerit pot ser transformat directament sobre cèl·lules d'*E. coli*.

3.3.7.1 Disseny dels oligonucleòtids

De manera general, quan es vol substituir una única base per una altra, els oligonucleòtids utilitzats per a la mutagènesi han de tenir una longitud aproximada de 21 bases, amb la mutació centrada. En el cas d'insercions, delecions o canvis de més d'una base consecutiva, els oligonucleòtids han de tenir com a mínim una longitud de 15 bases a cada costat de la regió de mutació. La longitud de cada braç s'ha d'ajustar de manera que la temperatura teòrica d'hibridació (o temperatura de fusió, T_m) sigui superior a 42°C, d'acord amb la fórmula següent:

$$T_m = 81.5 + 0.41(\%GC) - 675/N - \% \text{ de desaparellament}$$

on N és el nombre de nucleòtids totals de l'oligonucleòtid, %GC és el percentatge de guanina+citosina i % desaparellament és el percentatge de nucleòtids que no hibriden amb la seqüència patró.

3. Material i Mètodes

El protocol de *QuikChange* recomana que els oligonucleòtids tinguin una T_m per sobre de 70°C, per tal de poder ajustar la temperatura d'hibridació de la PCR.

Els oligonucleòtids han estat subministrats, de manera liofilitzada, per Roche Diagnostics S.L. (Roche Molecular Biochemicals, Alemanya). Aquests liofilitzats s'han resuspès amb 200 µl d'aigua milliQ, s'ha determinat la seva concentració espectrofotomètricament, i s'han preparat alíquotes de cada oligonucleòtid a la concentració de treball per a cada metodologia.

Reacció d'amplificació

<i>Pfu Buffer</i> (x10)	5 µl
dNTP mix	1 µl
dsDNA motlle (~50ng)	2 µl
Oligonucleòtid 3' (25ng/µl)	5 µl
Oligonucleòtid 5' (25ng/µl)	5 µl
Aigua milliQ estèril	32 µl
<i>Pfu Turbo</i> [®] DNA polimerasa	1 µl

Programa (Termociclador)

3 minuts a 95°C (desnaturalització inicial)	} 20 cicles
1 minut a 95°C (desnaturalització)	
1 minut a 55°C (hibridació)	
12 minuts a 68°C (extensió)	
Indefinidament a 4°C	

3.3.8 Obtenció de cèl·lules competents d'*E. coli*

El procediment dut a terme amb les màximes garanties d'esterilitat ha estat el següent:

Posar 10 ml de medi LB en una ampolla de 75 ml o en un tub Falcon, i inocular-hi la soca desitjada d'*E. coli* procedent del socari. Aquestes soques estan glicerinades.

Incubar durant tota la nit a 37°C a 180 rpm.

L'endemà, reinocular a una dilució 1/100 a partir del cultiu de nit (100 µl de cultiu de nit en 10 ml de medi) i incubar a 37°C a 200 rpm fins a assolir una DO_{550} de 0.6 – 0.7 (aproximadament 2 h).

Aliquotar en tubs Eppendorf o bé transferir a un tub Falcon de manera estèril.

Centrifugar a 4000 rpm a 4°C durant 10 min. Descartar el sobrenedant de manera estèril. Resuspendre el sediment amb el mateix volum de $CaCl_2$ 100mM fred i estèril.

Incubar en gel durant un mínim de 30 min.

Centrifugar novament a 4000 rpm a 4°C durant 10 min. Descartar el sobrenedant de manera estèril.

Resuspendre el sediment amb 1/10 del volum de CaCl₂ 100mM fred i estèril. Si s'escau, aliquotar el volum en tubs Eppendorf de 100 µl.

A partir d'aquí es poden transformar les cèl·lules competents o emmagatzemar-les a -80°C. En cas de guardar-les, en el pas final s'ha d'utilitzar una solució de CaCl₂ 100mM amb glicerol per obtenir-ne una concentració final del 15%.

3.3.9 Transformació de cèl·lules competents d'*E. coli*

Per a la introducció del DNA en les cèl·lules competents s'ha realitzat el següent protocol:

Afegir estèrilment, a cada alíquota de 100 µl de cèl·lules competents, 1 µL de DNA plasmídic si prové de minipreps o 10 µL de DNA si prové de la reacció amb *DpnI* posterior al QuikChange o d'un lligament (un màxim de 50 ng de DNA en un volum que no sobrepassi 1/10 del volum de la fracció o alíquota). Agitar suaument, colpejant amb el dit, per tal d'homogeneïtzar la solució.

Incubar en gel durant 1 - 2 h (mínim 20 min) i posar els tubs Eppendorf al bany a 42°C durant 1-2 min. Incubar els tubs en gel durant 10 min.

Realitzar una recuperació fenotípica, afegint 300 µl (transformants de minipreps) o 600 µL (transformants de QuikChange i lligaments) de medi LB líquid a cada tub Eppendorf. Incubar a 37°C en agitació suau durant 1 h.

Centrifugar 5 min a 4000 rpm, eliminar el volum de medi necessari fins un volum final de 200 µl i fer una sembra en plaques d'agar LB suplementades amb l'antibiòtic pel qual el vector que s'està manipulant proporcioni resistència (marcador de selecció). Incubar les plaques en una estufa a 37°C o/n, fins que s'observin les colònies transformants (12 - 15 h).

3.3.10 Seqüenciació de DNA

La seqüenciació del DNA de les construccions realitzades en aquest treball ha estat el mètode que ha permès validar els diferents clons obtinguts. S'ha seguit l'anomenat mètode del dideoxiribonucleòtid (Sanger and Coulson, 1978; Sanger et al., 1977).

S'ha emprat el kit comercial *ABI PRISM® dRhodamine Terminator Cycle Sequencing Ready Reaction Kit* (Applied Biosystems, EUA), el qual utilitza *AmpliTaq® DNA polymerase*. Aquest sistema es basa en reaccions de seqüenciació cíclica amb dideoxiribonucleòtids marcats amb un fluoròfor. L'enzim *AmpliTaq® DNA polymerase* és una variant de la polimerasa de *Thermus aquaticus* sense activitat

3. Material i Mètodes

correctora 3'→5', de manera que s'evita l'eliminació dels ddNTPs que s'incorporen a la cadena i aturen la síntesi.

La tècnica es basa en la síntesi d'una cadena complementària a la cadena motlle, per part de la *AmpliTaq® DNA polymerase*, fins que s'incorpora un dideoxiribonucleòtid marcat fluorescentment (ddNTP) al qual li manca el grup OH en 3' de la ribosa, de manera que s'atura la síntesi, ja que la polimerasa no pot seguir incorporant dNTPs. D'aquesta manera es van generant fragments de diferents mides aleatòriament, els quals són posteriorment separats i identificats mitjançant un seqüenciador automàtic. La identificació dels fragments es pot fer gràcies a la presència de fluoròfors units a cada un dels ddNTPs. Hi ha quatre tipus de ddNTP, un per cada nucleòtid, els quals porten incorporat un fluoròfor diferent, per facilitar la seva identificació.

Per portar a terme aquesta tècnica, cal obtenir prèviament gran quantitat del fragment de DNA a seqüenciar, és a dir del gen teòricament mutat. Aquest DNA s'aconsegueix utilitzant la reacció en cadena de la polimerasa PCR i utilitzant els encebadors universals T7term i T7prom o d'altres dissenyats específicament, els quals són complementaris a les regions del promotor i finalitzador dels gens clonats en els vectors utilitzats en aquest treball.

Un cop amplificat el fragment de DNA a seqüenciar, es realitza una segona PCR amb els ddNTPs i l'enzim *AmpliTaq® DNA polimerase*.

A continuació es detallen les condicions per a dur a terme la primera i la segona PCR.

3.3.10.1 Primera PCR (amplificació de l'inserit)

Barreja de Reacció:

Aigua milli-Q estèril	35.5 µl
<i>Vent Buffer</i> (x10)	5 µl
Oligonucleòtid <i>T7Prom</i> (5µmol/µl)	2.5 µl
Oligonucleòtid <i>T7Term</i> (5µmol/µl)	2.5 µl
dsDNA motlle (~50ng)	2.5 µl
<i>dNTPs</i> 10mM (0.4 nM final)	2 µl
<i>Deep Vent</i> DNA polimerasa	0.5 µl

Programa del termociclador

3 minuts a 95°C (desnaturalització inicial)	} 30 cicles
1 minut 15 segons a 94°C (desnaturalització)	
1 minut a 52°C (hibridació)	
1 minut 15 segons a 72°C (extensió)	
5 minuts a 72°C (extensió final)	
Indefinidament a 4°C	

El DNA amplificat es precipita i es purifica a partir d'un gel d'agarosa. El DNA s'elueix en 40 µl d'aigua milliQ estèril escalfada prèviament a 65°C.

3.3.10.2 Precipitació del producte de PCR

Recuperar el producte de PCR i afegir-hi 0.1 volums d'acetat sòdic 3 M pH 5.5 i 3 volums d'etanol absolut refredat a -20°C. Agitar en vòrtex i guardar a -80°C durant 1 h o a -20°C o/n. Passat aquest temps, centrifugar durant 25 min a 15000 rpm i eliminar el sobrenedant. Afegir-hi 500 µl d'etanol 70% fred (-20°C) i centrifugar durant 5 min a 15000 rpm, aspirar el sobrenedant i deixar assecar el sediment a 37°C durant 10 - 15 min. Finalment, resuspendre el sediment amb 20µl d'aigua milliQ estèril.

3.3.10.3 Segona PCR (reacció de seqüenciació)

Utilitzant el kit comercial *ABI PRISM® dRhodamine Terminator Cycle Sequencing Ready Reaction Kit* (Applied Biosystems), com a motlle el dsDNA obtingut en la primera PCR i com a encebador T7Prom o T7Term.

Barreja de Reacció:

dsDNA motlle	5 µl
Aigua milli-Q estèril	1 µl
Sequencing Buffer X5	2 µl
Encebador (5 pmol/µl)	0.5 µl
Ready mix	1.5 µl

El *Ready Mix* conté ddNTPs, dNTPs, Tris-HCl (pH 9.0), MgCl₂, *AmpliTaq® DNA polimerase* i pirofosfatasa alcalina.

Programa del termociclador

3 minuts a 94°C (desnaturalització inicial)	} 25 cicles
10 segons a 96°C (desnaturalització)	
5 segons a 50°C (hibridació)	
4 minuts a 60°C (extensió)	
10 minuts a 96°C (desnaturalització)	
Indefinidament a 4°C	

3.3.10.4 Precipitació de la reacció de seqüenciació

Utilitzant els encebadors per separat s'aconsegueix la seqüenciació de cada una de les cadenes, la qual cosa permet seqüenciar tot el gen. Els diferents fragments marcats amb ddNTPs es precipiten afegint 0.1 volums d'acetat de sodi 3 M pH 5.5 i 3 volums d'etanol absolut al producte de la reacció. Es vorteja suaument i es deixa 15 min a temperatura ambient. Es transfereix a un tub Eppendorf.

Centrifugar 15 min a 15000 rpm, aspirar el sobrenedant, afegir-hi 250 µl d'etanol 70% fred i vortejar suaument. Centrifugar 5 min a 15000 rpm, aspirar el sobrenedant i assecar 15 min a l'estufa a 37°C. Resuspendre amb 10 µl de formamida, vortejar suaument i deixar 10 min a temperatura ambient. Fer un pols de centrífuga i portar al servei de seqüenciació.

3.4 MÈTODES D'EXPRESSIONI I PURIFICACIÓ DE PROTEÏNES RECOMBINANTS

3.4.1 Expressió de la inteïna *Npu DnaE*, $I_N - I_C$, I_N -GB1 i I_C -GB1

Les condicions d'expressió de *Npu DnaE*, $I_N - I_C$, I_N -GB1 i I_C -GB1 han estat les mateixes en tots els casos variant l'antibiòtic (Ampicil·lina per *Npu DnaE*, $I_N - I_C$, I_C -GB1 i kanamicina en el cas de I_N -GB1) i inductor requerit (IPTG $I_N - I_C$ i I_N -GB1 i L-Arabionsa en el cas de I_C -GB1). S'ha utilitzat sempre la soca d'*E. coli* BL21(DE3).

A partir d'un glicerinat o d'una colònia aïllada d'una sembra en placa d'*E. coli* BL21(DE3) transformat amb el vector corresponent es realitza un inòcul en 10 ml de medi LB complementat amb ampicil·lina o kanamicina 100 µg/ml. S'incuba durant 10 min i es reinocula a una dilució 1/100 sobre 40 ml de medi LB complementat amb ampicil·lina 100 µg/ml. D'aquesta manera s'assegura que la fase estacionària no és massa llarga i que la quantitat de cèl·lules viables és suficientment gran. S'incuba durant 10 - 12 h a 37°C i sota agitació constant a 180 rpm.

Després d'aquest temps d'incubació, s'inoculen 4 flascons Erlenmeyer amb 250 ml de medi líquid LB complementat amb ampicil·lina 100 µg/mL amb una dilució 1/100 del volum provinent del cultiu de nit. Es deixa créixer el cultiu a 37°C sota agitació constant de 250 rpm durant 2 h, fins que s'assoleix un valor d'absorbància a 550 nm de 0.6 – 0.8.

En aquest moment s'afegeix al cultiu IPTG (Fermentas, EUA) a una concentració final de 1 mM o L-Arabionsa (Sigma Aldrich, EUA) al 0.2%. Es deixa créixer el cultiu induït durant 3 - 4 h, en les mateixes condicions i, tot seguit, se centrifuga a 7000 rpm durant 7 min per tal de recollir les cèl·lules. Aquestes es guarden a -20°C fins al moment de començar el procés de purificació.

Per tal d'analitzar la producció de proteïna recombinant durant la inducció, es prenen mostres d'1 ml de cultiu en el moment anterior a la inducció i al finalitzar el temps d'inducció. Les mostres se centrifuguen a 15000 rpm durant 1 min, s'elimina el sobrenedant per aspiració i es resuspèn el sediment cel·lular en 100 µl de tampó d'aplicació de proteïnes x1 amb β - mercaptoetanol. L'anàlisi de l'expressió es duu a terme electroforèticament, utilitzant gels de poliacrilamida SDS tenyits amb Coomassie blue per visualitzar proteïnes. Aquest procés es realitza en cadascun dels protocols d'expressió de proteïnes recombinants realitzats en aquest treball.

3.4.2 Expressió de I_N, les variants I_NI_C_sc i de la proteasa TEV

3.4.2.1 Procediment general

Les condicions d'expressió de I_N, variants de I_NI_C_sc i la proteasa TEV han estat les mateixes en tots els casos variant l'antibiòtic (Ampicil·lina I_N i I_NI_C_sc i kanamicina en el cas de la proteasa TEV) i inductor requerit (L-Arabionsa en el cas de I_N i I_N_sc i IPTG en el cas de la proteasa TEV) i utilitzant sempre la soca d'*E. coli* BL21(DE3).

A partir d'un glicerinat o d'una colònia aïllada d'una sembra en placa d'*E. coli* BL21(DE3) transformat amb el vector corresponent es realitza un inòcul en 10 ml de medi LB complementat amb ampicil·lina o kanamicina 100 µg/ml. S'incuba durant 10 min i es reinocula a una dilució 1/100 sobre 40 ml de medi LB complementat amb ampicil·lina 100 µg/ml. D'aquesta manera s'assegura que la fase estacionària no és massa llarga i que la quantitat de cèl·lules viables és suficientment gran. S'incuba durant 10 - 12 h a 37°C i sota agitació constant a 180 rpm.

Després d'aquest temps d'incubació, s'inoculen 4 flascons Erlenmeyer amb 250 ml de medi líquid LB complementat amb ampicil·lina 100 µg/ml amb una dilució 1/100 del volum provinent del cultiu de nit. Es deixa créixer el cultiu a 37°C sota agitació constant de 250 rpm durant 2 h, fins que s'assoleix un valor d'absorbància a 550 nm de 0.8 - 1. Llavors es baixa la temperatura de l'incubador a 18°C.

Una vegada l'incubador està a 18°C s'afegeix al cultiu IPTG (Fermentas, EUA) a una concentració final de 1 mM o L-Arabionsa (Sigma Aldrich, EUA) al 0.2%. Es deixa créixer el cultiu induït durant 15-18 h, en les mateixes condicions i, tot seguit, se centrifuga a 7000 rpm durant 7 min per tal de recollir les cèl·lules. Aquestes es guarden a -20°C fins al moment de començar el procés de purificació.

3.4.2.2 Producció de les variants I_NI_C_sc marcades amb ¹³C¹⁵N

Per tal d'obtenir proteïnes marcades amb isòtops pel treball amb ressonància magnètica nuclear (NMR) es va seguir un procediment equivalent al descrit en el punt previ, substituint el medi de creixement LB per medi mínim complementat amb ¹⁵NH₄Cl com a única font de nitrogen i ¹³C₆H₁₂O₆ com a única font de carboni (Cambridge Isotope Labs, USA).

3.4.3 Purificació de la *NpuDnaE* i de I_N-I_C

El procés de purificació s'inicia amb l'expressió heteròloga de la proteïna. Les cèl·lules recollides per centrifugació es resuspenen en un tampó i tot seguit es procedeix al seu trencament per mitjà de tres passades a la premsa de French. A continuació se separa la fracció soluble de la insoluble per centrifugació i s'aïlla la fracció soluble que és on hi ha la major part de la proteïna. La fracció soluble es dialitza contra l'amortidor d'equilibrat d'una columna de bescanvi aniònic Hi-Trap Q HP d'HPLC

(GE Healthcare, EUA) de la qual s'elueix mitjançant un gradient de clorur de sodi. Es recull el pic principal del cromatograma que correspon a la proteïna pura i homogènia.

De manera detallada, el protocol utilitzat és el següent:

1. Les cèl·lules provinents de la centrifugació del litre de cultiu, es resuspenen en 80 ml de Tris-HCl 50mM, EDTA 10mM pH 8.
2. Passar la suspensió cel·lular 3 vegades per la Premsa de French a 13000 psi. Un bon indicador del bon funcionament de la tècnica és la pèrdua de viscositat de la suspensió.
3. Centrifugar les mostres a 11000 rpm durant 40 min a 4°C. D'aquesta manera separem la fracció soluble que quedarà en el sobrenedant de la fracció insoluble que quedarà en el precipitat.
4. Es recupera el sobrenedant i es dialitza contra 5 l de Tris-Acetata 50mM, EDTA 10mM, DTT 1mM a pH 8.
5. Es filtra la mostra amb un filtre de 0.22 µm.
6. Aplicar la mostra a una columna Hi-Trap Q HP (GE Healthcare, EUA) d'intercanvi aniònic. El tampó d'equilibrat ha estat Tris-Acetata 50mM, EDTA 10mM, DTT 1mM a pH 8, i la proteïna s'ha eluït utilitzant un gradient de clorur de sodi amb el tampó Tris-Acetata 50mM, EDTA 10mM, DTT 1mM, NaCl 1M a pH 8. L'elució de la proteïna se segueix amb una mesura continua d'absorbància a 280nm i es recull el volum corresponent al pic principal del cromatograma.
7. Dialitzar la mostra enfront d'aigua milliQ per tal d'eliminar les sals de la mostra. S'han emprat sacs de diàlisi *Spectra/Por*[®] MWCO: 3.500.
8. Congelar la mostra a -80°C i liofilitzar-la.
9. La mostra liofilitzada s'emmagatzema a -20°C en un recipient tancat. D'aquesta manera la proteïna es manté en condicions òptimes durant mesos.

3.4.4 Purificació de I_N i de les variants I_NI_C_sc

El procés de purificació s'inicia amb l'expressió heteròloga de la proteïna. Les cèl·lules recollides per centrifugació es resuspenen en un tampó i tot seguit es procedeix al seu trencament per mitjà de tres passades a la premsa de French. A continuació se separa la fracció soluble de la insoluble per centrifugació i s'aïlla la fracció soluble que és on hi ha la major part de la proteïna. La fracció soluble es passa per una columna d'afinitat Ni-NTA de la qual la proteïna d'interès s'elueix mitjançant un gradient de imidazol. Aleshores es realitza una digestió amb la proteasa TEV per tal de separar la proteïna d'interès de la proteïna de solubilització. La mostra es passa de nou per una columna d'afinitat Ni-NTA on la proteïna de solubilització quedarà retinguda a la reïna i la proteïna d'interès s'eluirà en el través. El través d'aquesta columna d'afinitat Ni-NTA es dialitza contra l'amortidor

3. Material i Mètodes

d'equilibrat d'una columna de bescanvi aniònic Hi-Trap Q HP d'HPLC (GE Healthcare, EUA) de la qual s'elueix mitjançant un gradient de clorur de sodi. Es recull el pic principal del cromatograma que correspon a la proteïna d'interès. Aquest pic es dialitza en aigua milliQ i es liofilitza. El liofilitzat es ressuspen amb un amortidor, el qual s'ha utilitzat per equilibrar una columna de gel filtració Superdex 75 10/300 GL (GE Healthcare, EUA) que s'utilitza per acabar de purificar i homogeneïtzar la proteïna.

De manera detallada, el protocol utilitzat és el següent:

1. Les cèl·lules provinents de la centrifugació del litre de cultiu, es ressuspenen en 80 ml de Tris-HCl 25mM, NaCl 150mM pH 8.
2. Passar la suspensió cel·lular 3 vegades per la Premsa de French a 13000 psi. Un bon indicador del bon funcionament de la tècnica és la pèrdua de viscositat de la suspensió.
3. Centrifugar les mostres a 11000 rpm durant 40 min a 4°C. D'aquesta manera separem la fracció soluble que quedarà en el sobrenedant de la fracció insoluble que quedarà en el precipitat.
4. Es recupera el sobrenedant i s'afegeix la reïna de Ni-NTA prèviament equilibrada amb Tris-HCl 25mM, NaCl 150mM pH 8. Es deixa en agitació suau a 4°C o/n. D'aquesta manera la proteïna quedarà unida a la reïna de Ni-NTA.
5. Es posa la mostra juntament amb la reïna dins una columna per la qual hi passarem 3V d'una solució de Tris-HCl 25mM, NaCl 150mM, 5mM d'imidazol a pH 8 per fer rentats, 2V d'una solució de Tris-HCl 25mM, NaCl 150mM, 25mM d'imidazol a pH 8 per eliminar possibles proteïnes que s'hagin unit inespecíficament a la reïna i per últim es passen 4V d'una solució de Tris-HCl 25mM, NaCl 150mM, 250mM d'imidazol a pH 8 que es recull ja que és on hi ha la proteïna d'interès.
6. Es dialitza la fracció recollida enfront Tris-HCl 25mM pH 8.
7. S'ajusta la concentració de proteïna de la mostra a 1.5 mg/ml i s'afegeix la proteasa TEV a una proporció 1:50 la qual deixarem digerir la mostra a 25°C o/n.
8. Es realitza una altra columna d'afinitat de Ni-NTA repetint els passos 4 i 5 d'aquest protocol. Aquesta vegada però, la proteïna s'elueix en el través de la cromatografia que caldrà recollir.
9. La mostra es dialitza contra 3 l de Tris-HCl 25mM, EDTA 1mM, DTT 1mM pH 8.
10. Es filtra la mostra amb un filtre de 0.22 µm.
11. Aplicar la mostra a una columna Hi-Trap Q HP (GE Healthcare, EUA) d'intercanvi aniònic. El tampó d'equilibrat ha estat Tris-HCl 25mM, EDTA 1mM, DTT 1mM a pH 8, i la proteïna s'ha eluït utilitzant un gradient de clorur de sodi amb el tampó Tris-HCl 50mM, EDTA 1mM, DTT 1mM, NaCl 1M a pH 8. L'elució de la proteïna se segueix amb una mesura continua d'absorbància a 280nm i es recull el volum corresponent al pic principal del cromatograma.

3. Material i Mètodes

12. Dialitzar la mostra enfront d'aigua milliQ per al d'eliminar les sals de la mostra.
13. Congelar la mostra a -80°C i liofilitzar-la.
14. La mostra es resuspen amb Tris-HCl 100mM, NaCl 100mM, EDTA 1mM, DTT 1mM pH 8 a una concentració de 1.5 – 2 mg/ml.
15. Aplicar la mostra a una columna de gel filtració Superdex 75 10/300 GL (GE Healthcare, EUA) on el tampó d'equilibrat ha estat Tris-Acetate 50mM, EDTA 10mM, DTT 1mM a pH 8. L'elució de la proteïna se segueix amb una mesura continua d'absorbència a 280nm i es recull el volum corresponent al pic principal del cromatograma.
16. Dialitzar la mostra enfront d'aigua milliQ per al d'eliminar les sals de la mostra.
17. Congelar la mostra a -80°C i liofilitzar-la.
18. La mostra liofilitzada s'emmagatzema a -20°C en un recipient tancat. D'aquesta manera la proteïna es manté en condicions òptimes durant mesos.

3.4.5 Purificació de I_N-GB1 i I_C-GB1

El procés de purificació s'inicia amb l'expressió heteròloga de la proteïna. Les cèl·lules recollides per centrifugació es resuspenen en un tampó i tot seguit es procedeix al seu trencament per mitjà de tres passades a la premsa de French. A continuació se separa la fracció soluble de la insoluble per centrifugació i s'aïlla la fracció soluble que és on hi ha la major part de la proteïna. La fracció soluble es dialitza contra l'amortidor d'equilibrat d'una columna de Ni-NTA de la qual s'elueix mitjançant un amortidor d'elevada concentració d'imidazol. Es recull la mostra provinent de l'elució amb imidazol que és on hi haurà la mostra pura i homogènia.

De manera detallada, el protocol utilitzat és el següent:

1. Les cèl·lules provinents de la centrifugació del litre de cultiu, es resuspenen en 80 ml de Tris-HCl 25mM, NaCl 150mM pH 8.
2. Passar la suspensió cel·lular 3 vegades per la Premsa de French a 13000 psi. Un bon indicador del bon funcionament de la tècnica és la pèrdua de viscositat de la suspensió.
3. Centrifugar les mostres a 11000 rpm durant 40 min a 4°C. D'aquesta manera separem la fracció soluble que quedarà en el sobrenedant de la fracció insoluble que quedarà en el precipitat.
4. Es recupera el sobrenedant i s'afegeix la reïna de Ni-NTA prèviament equilibrada amb Tris-HCl 25mM, NaCl 150mM pH 8. Es deixa en agitació suau a 4°C o/n. D'aquesta manera la proteïna quedarà unida a la reïna de Ni-NTA.
5. Es posa la mostra juntament amb la reïna dins una columna per la qual hi passarem 3V d'una solució de Tris-HCl 25mM, NaCl 150mM, 5mM d'imidazol a pH 8 per fer rentats, 2V d'una solució de Tris-HCl 25mM, NaCl 150mM, 25mM d'imidazol a pH 8 per eliminar

possibles proteïnes que s'hagin unit inespecíficament a la reïna i per últim es passen 4V d'una solució de Tris-HCl 25mM, NaCl 150mM, 250mM d'imidazol a pH 8 que es recull ja que és on hi ha la proteïna d'interès.

6. Es dialitza la fracció recollida enfront l'amortidor que utilitzarem per tal de dur a terme els assajos pertinents. La mostra també es pot dialitzar enfront d'aigua milliQ per tal de liofilitzar la mostra.

3.4.6 Purificació de la proteasa TEV

El procés de purificació de la proteasa TEV segueix el mateix protocol que el descrit en l'apartat 2.4.5. No obstant cal tenir en compte que l'amortidor d'elució cal afegir-hi un 10% de glicerol. La mostra un cop purificada s'aliquota en tubs Eppendorf de 2 ml i es guarda a una concentració de 2 mg/ml -80°C amb Tris-HCl 25mM, NaCl 150mM, 50% de glicerol pH 8. No es recomana descongelar la mostra més de dues vegades ja que pot perdre eficàcia en el procés de digestió.

3.5 ANÀLISI I CARACTERITZACIÓ PROTEICA

3.5.1 Determinació de la massa molecular

La massa molecular de les proteïnes recombinants estudiades en aquest treball s'ha comprovat per espectrometria de masses MALDI-TOF (*Matrix-Assisted Laser Desorption/Ionization Time-Of-Flight mass spectrometry*) en un equip BRUKER. Aquesta anàlisi s'ha dut a terme als Serveis Tècnics de Recerca de la UdG.

3.5.2 Determinació de la quantitat de proteïna

3.5.2.1 Mètode espectrofotomètric

La concentració de les diferents proteïnes es va determinar per espectrofotometria, a partir del valor d'absorbància a 280 nm, segons la fórmula:

$$\text{Abs} = [\text{proteïna}] \cdot \epsilon_{280} \cdot l$$

On Abs és l'absorbància a 280 nm, [proteïna] és la concentració de proteïna en M, ϵ_{280} és el coeficient d'extinció molar a 280nm utilitzant els coeficients d'extinció molar calculats seguint el mètode de Pace (Pace et al., 1995) i l és la longitud de pas de la cubeta utilitzada en cm.

3.5.3 Electroforesi discontinua en gel de poliacrilamida

L'electroforesi discontinua en gel de poliacrilamida en presència d' SDS (SDS-PAGE) permet la separació de proteïnes segons la seva massa molecular gràcies a l'acció d'un camp elèctric. Ha estat el mètode analític utilitzat de manera rutinària per a la determinació del contingut proteic de cadascuna de les fraccions cel·lulars i el seguiment dels processos de purificació, així com per a la determinació del grau de puresa de les mostres obtingudes al final del procés.

En aquest treball s'ha utilitzat el mètode d'electroforesi discontinua descrit per Lane & Harlow (Lane and Harlow, 1982). Aquest mètode es caracteritza per la utilització d'un gel superior o gel apilador (amb una concentració d'acrilamida insuficient per separar les proteïnes) que té com a funció concentrar les proteïnes abans que entrin en el gel inferior o gel separador (amb una concentració superior d'acrilamida) el qual permetrà separar les proteïnes.

En el cas de la preparació d'un gel SDS-PAGE al 15% d'acrilamida i de 0.75mm de gruix, el procediment és el següent:

1. Preparació del gel separador al 15% d'acrilamida

Acrilamida 30% - Bisacrilamida 0.8%	1.5 ml
Tampó gel separador (x4) (Tris-HCl 1.5M, SDS 0.4%, pH 8.8)	1 ml
Aigua milliQ	1.5 ml
Persulfat amònic 10% (p/v)	30 µl
TEMED	3 µl

2. Preparar el kit d'electroforesi i abocar la dissolució del gel separador en el suport evitant la formació de bombolles. A continuació, afegir una fina capa d'aigua milliQ per facilitar la polimerització i evitar malformacions del front.

3. Preparació del gel apilador

Acrilamida 30% - Bisacrilamida 0.8%	0.25 ml
Tampó gel apilador (x4) (Tris-HCl 0.5M, SDS 0.4%, pH 6.8)	0.5 ml
Aigua milliQ	1.25 ml
Persulfat amònic 10% (p/v)	30 µl
TEMED	3 µl

4. Eliminar la capa d'aigua i abocar el gel apilador. Col·locar el motlle formador de butxaques i deixar polimeritzar.

5. Preparar la cubeta d'electroforesi i l'amortidor d'elució:

Amortidor d'elució 10x	100 ml
(Tris-HCl 0.25M, glicina 1.92M, SDS 1%, pH 8.52)	
Aigua milliQ	900 ml

3. Material i Mètodes

6. Preparar les mostres aplicant amortidor d'aplicació de mostres x4 (Tris-HCl 50mM, glicerol 20%, SDS 8%, blau de bromofenol 0.02%) en una relació 3:1. Si es vol que l'amortidor tingui caràcter reductor s'afegeix β -mercaptoetanol (15% v/v) al tampó d'aplicació.
7. Escalfar les mostres 5 minuts a 90°C i carregar-les a les butxaques del gel.
8. Connectar la font d'alimentació a un voltatge constant de 150-200v i deixar avançar la separació fins que el tampó d'aplicació que marca el límit de la separació, assoleixi l'extrem inferior del gel (aproximadament una hora).
9. Tenyir el gel de poliacrilamida durant 10 minuts amb blau de Comassie: Comassie Blue 0.1% (p/v), metanol 45% (v/v), àcid acètic 10% (v/v).
10. Eliminar l'excés de colorant amb diversos rentats amb àcid acètic al 7%.

3.5.4 Western blot

Tècnica electroforètica basada en la transferència de proteïnes carregades d'un gel d'electroforesi a una membrana de material sintètic, mitjançant l'aplicació d'un camp elèctric perpendicular a les dues parts. A continuació es detecten les proteïnes transferides a la membrana mitjançant la utilització d'anticossos específics.

El procés per a la realització d'un western blot és el següent:

1. Electrotransferència de les mostres a partir d'un gel SDS-PAGE (veure apartat 2.5.3).
2. Activar la membrana de PVDF (Millipore, EUA) submergint-la en metanol 100% durant 15 segons i després rentant-la amb aigua milliQ durant 2 minuts.
3. Submergir la membrana activada i les esponges en tampó d'electrotransferència (Tris 25mM, glicina 192mM, metanol 20% v/v, pH 8.3) durant almenys 5 minuts. Es submergeixen també els papers de filtre com a mínim 30 segons.
4. Muntar el sistema d'electrotransferència segons el següent ordre a partir del càtode o pol negatiu (eliminant les possibles bombolles): esponja, paper de filtre, gel de poliacrilamida, membrana activada, paper de filtre i esponja.
5. Posar el sistema dins la cubeta d'electrotransferència, col·locar també la unitat refrigerant i una barra magnètica, i omplir la cubeta amb tampó d'electrotransferència.
6. Connectar la font d'alimentació i deixar l'electrotransferència 2-3 hores a 100V.
7. Finalitzada la transferència, comprovar que els marcadors s'han transferit bé.
8. Submergir la membrana de PVDF en tampó TBST (Tris-HCl 10mM, NaCl 100mM, Tween 0.1%, pH 7.5) durant 5 minuts a temperatura ambient i amb agitació.
9. Eliminar el tampó i afegir 20 ml de tampó de bloqueig (1% BSA en TBST), incubar durant 1h a temperatura ambient i amb agitació.

3. Material i Mètodes

10. Eliminar el tampó de bloqueig i rentar la membrana amb tampó TBST. Repetir el procés 3 vegades.
11. Preparar 10 ml d'una dilució 1:6000 de l'anticòs primari *Anti-GFP Mouse IgG* (Roche, Alemanya) o 1:20000 *Tetra-His Antibody mouse IgG* (Qiagen, Alemanya) en tampó de bloqueig (1% BSA en TBST) i incubar a 4°C amb agitació o/n.
12. Eliminar l'anticòs primari i rentar 3 vegades amb TBST durant 5 minuts a temperatura ambient i amb agitació.
13. Preparar 20 ml d'una dilució 1:20000 de l'anticòs secundari *Goat Anti-Mouse IgG Peroxidase Conjugate* (Calbiochem, Alemanya) en tampó de bloqueig (1% BSA en TBST) i incubar 1h a temperatura ambient i amb agitació.
14. Eliminar l'anticòs secundari i rentar 3 vegades amb TBST durant 5 minuts a temperatura ambient i amb agitació.
15. Barrejar 2 ml de la dissolució A (substrat de la peroxidasa, conté luminol) amb 2 ml de la dissolució B (peròxid d'hidrogen) del kit quimioluminiscent Immobilon™ Western (Millipore, EUA).
16. Abocar la dissolució de detecció sobre la membrana i incubar durant 5 minuts a temperatura ambient sense agitació.
17. Assecar la membrana amb paper de filtre i col·locar-la entre dos plàstics transparents.
18. Revelar amb una càmera CCD.

3.5.5 Determinació de l'estabilitat tèrmica

L'estabilitat conformacional d'una proteïna pot ser determinada mitjançant el seguiment espectrofotomètric del desplegament induït per temperatura. Els canvis en l'entorn dels residus aromàtics durant aquest procés de desplegament poden ser monitoritzats mitjançant mesures d'absorbància a 280 nm, permetent-se la determinació de la temperatura mitja de fusió (T_m).

Per això, es dissol la proteïna a estudiar fins una concentració de 0.5 mg/ml en l'amortidor escollit i es fa un seguiment de l'absorbància a 280 nm, entre 20°C i 90°C, amb increments de temperatura de 2°C i un temps d'equilibrat de 5 - 10 min a cada temperatura, en un espectrofotòmetre equipat amb una cel·la termostatitzada.

Una vegada assolida la temperatura màxima, amb l'absorbància constant, es retorna a l'estat inicial amb salts de temperatura de 5°C per tal de comprovar la reversibilitat del procés. En la desnaturalització proteica s'estableix un equilibri entre les formes plegada i desplegada. Si en refredar la mostra se segueix un camí equivalent al de desplegament es pot entendre el procés com una reacció en dos estats permetent el càlcul de paràmetres termodinàmics (Font et al., 2006; Mozhaev et al., 1996; Torrent et al., 1999). En el cas que una corba d'estabilitat tèrmica no sigui

reversible, aquesta no es pot ajustar a un model de dos estats i no es poden calcular els paràmetres termodinàmics de ΔG i ΔH .

En aquests estudis s'espera obtenir una corba sigmoïdal representativa del procés de desnaturalització proteica, amb una regió de pre-transició, una de post-transició i una regió on la variació de l'absorbància és elevada i que correspon a la transició deguda al pas de l'estat natiu al desnaturalitzat. Aquesta corba s'ajusta emprant el programa Sigmaplot (Jandel, Regne Unit) a l'equació:

$$\text{Abs} = \left(\frac{n-d+(q-m) \cdot t}{1+e^{-ab}} + d + m \cdot t \right)$$

$$a: ((\Delta H_m/R)((1/(t+273.15))-(1/T_m+273.15)))-(\Delta C_p/R)$$

$$b: (((T_m+273.15)/(t+273.15))-1+\ln((t+273.15)/(T_m+273.15)))$$

R: Constant dels gasos

ΔC_p : Capacitat calorífica en KJ/Kmol

n: punt de tall de la pre-transició

d: punt de tall de la post-transició

q: pendent pre-transició

m: pendent post-transició

ΔH_m : Canvi d'entalpia a la temperatura de fusió

t: temperatura °C

T_m : Temperatura de fusió

A partir d'aquest ajust s'obté la temperatura de fusió (T_m) que determina l'estabilitat de la proteïna en front de la temperatura.

3.5.6 Estudi del procés de tall i unió

Per l'estudi del procés de tall i unió, cada una de les inteïnes purificades (sortint de columna Ni-NTA) van ser dialitzades enfront el tampó de tall i unió Tris-HCl 25mM, NaCl 150mM, EDTA 1mM, DTT 1mM pH 8 i es va ajustar la concentració de cada proteïna a 1mg/ml. Per iniciar la reacció, la I_N i la I_C es van barrejar en volums iguals (25 μ l) en tubs Eppendorf i es van incubar a les temperatures pertinents (de 20 a 60°C) durant el temps de reacció a assajar. Passat el temps de reacció, la mostra es barrejava amb tampó de càrrega 2x (Tris-HCl 25mM, glicerol 10%, SDS 4%, blau de bromofenol 0.01%, β -mercaptoetanol 15% v/v) per tal d'aturar la reacció. Per a cada una de les reaccions es va agafar un temps 0 barrejant I_N i I_C a iguals quantitats. Les mostres es van escalfar durant 10 minuts a 90°C i es van carregar en un gel SDS-PAGE. Els gels es van tenyir amb blau de Comassie i es van escanejar.

3.5.7 Càlcul de paràmetres cinètics

El càlcul de paràmetres cinètics del procés de tall i unió es va realitzar a partir de la quantificació del producte de tall unió a partir dels gels SDS-PAGE obtinguts. Cada columna del gel es va analitzar utilitzant el programa de quantificació Gel Analyzer (Dr. Istvan Lazar, <http://www.gelanalyzer.com/>). Per tal de normalitzar l'error de càrrega de la mostra, la intensitat de cada banda en una columna es va expressar com a la fracció de la intensitat total de la columna analitzada. Es van fer gràfics de les intensitats en funció del temps i la corba obtinguda es va ajustar a una equació exponencial de primer ordre utilitzant el programa SigmaPlot (Jandel, Regne Unit). L'equació de la corba obtinguda per la formació del producte de tall i unió és la següent:

$$Y = Y_{\max} \cdot (1 - e^{-k_{\text{obs}} \cdot t})$$

On Y es la intensitat normalitzada de cada banda, t és el temps en minuts, Y_{\max} és el factor escalar de la formació de producte i K_{obs} és la constant de primer ordre de la reacció de tall i unió.

El temps de vida mitja ($t_{1/2}$), temps en el qual tarda a formar-se la meitat del producte total observat, es va calcular a partir del valor de la constant de primer ordre K_{obs} :

$$t_{1/2} = \ln 2 / k_{\text{obs}}$$

3.5.8 Anàlisi de l'estructura secundària per dicroisme circular

Si bé la llum no polaritzada consta d'ones que vibren en tots els plans perpendiculars a la direcció del desplaçament, la llum polaritzada plana està composta per ones que vibren en un únic pla i el seu camp elèctric té una orientació fixa. En la llum polaritzada circular, el camp elèctric pot rotar tant en el sentit de les agulles del rellotge (llum polaritzada circular cap a la dreta) com en el sentit contrari (llum polaritzada circular cap a l'esquerra). La majoria de les molècules biològiques són asimètriques i presenten una diferència d'absorció de la llum polaritzada circularment cap a la dreta, respecte de la llum polaritzada circularment cap a l'esquerra, anomenada dicroisme circular (CD). El dicroisme circular es defineix com:

$$\Delta A = (A_L - A_R) / A$$

On A_L és l'absorbància (per part de la mostra analitzada) de la llum polaritzada circularment cap a l'esquerra, A_R és l'absorbància corresponent a la llum polaritzada circularment cap a la dreta i A és l'absorbància de la llum no polaritzada. Així doncs, ΔA pot ser positiva o negativa.

3. Material i Mètodes

L'absorció desigual de la llum polaritzada circular (activitat òptica) de les proteïnes en la regió de l'ultraviolat llunyà o regió amida (170 – 250 nm) depèn principalment dels enllaços peptídics. A més, els ponts disulfur donen lloc a petites bandes CD al voltant de 250nm. Així doncs, les bandes CD de la regió amida contenen informació sobre els enllaços peptídics i sobre l'estructura secundària de la proteïna.

Es va emprar un espectropolarímetre equipat amb una cel·la termostatitzada. L'espectre de diroisme circular es van enregistrar entre 200 i 260 nm (UV llunyà) on el compartiment de la mostra es purgava ininterrompudament amb nitrogen.

Les condicions que es van utilitzar per enregistrar els espectres van ser:

Sensibilitat:	20 mdeg
Resolució:	0.5 nm
Amplada de banda:	1.2 nm
Velocitat de lectura:	50 nm/min
Acumulació d'espectres:	6

Les proteïnes es van dissoldre en tampó fosfat 10 mM, DTT 1mM pH 7 a una concentració de 1mg/ml i es va emprar una cubeta de 0.1 cm de pas òptic. La cubeta es va rentar prèviament amb una solució d'àcid nítric al 50% i es van esbandir amb aigua milliQ.

Per cada espectre enregistrat es va restar el blanc corresponent i els espectres obtinguts es van analitzar amb el programa Spectra Analysis (JASCO Co., USA). L'aparell donava les mesures inicials d'activitat òptica en el·lipticitat, θ_{obs} (mdeg). Després d'introduir les mesures del pas de llum de la cubeta (en cm) i la molaritat de la mostra (M), es mesura l'el·lipticitat molar $[\theta]$ ($\text{deg}\cdot\text{cm}^2\cdot\text{dmol}^{-1}$). L'activitat òptica de les proteïnes en la regió amida s'expressa sempre com a el·lipticitat mitjana per residu $[\theta]_{MRW}$ ($\text{mdeg}\cdot\text{cm}^2/\text{dmol}$). El MRW d'una proteïna és igual a la seva massa molecular dividida pel nombre d'aminoàcids. Així doncs, el càlcul de l'el·lipticitat mitjana per residu s'obté a partir de l'equació:

$$[\theta]_{MRW} = \frac{\theta_{obs} \cdot MRW}{10 \cdot l \cdot c}$$

On θ_{obs} és l'el·lipticitat mesurada en mdeg, MRW és la massa molecular mitjana per residu (Da), c és la concentració de proteïna en mg/ml i l és el pas òptic de la cubeta en cm.

A partir dels valors de el·lipticitat mitjana per residu s'ha fet una predicció de l'estructura secundària de la proteïna a partir dels programes *K2D3* (Louis-Jeune et al., 2011) i *Dicroprot* (Deleage and Geourjon, 1993).

3.6 ESPECTROSCÒPIA DE RESSONÀNCIA MAGNÈTICA NUCLEAR (NMR)

El fenomen de la ressonància magnètica té el seu origen en la interacció entre el moment magnètic d'un nucli atòmic amb un camp magnètic extern. La causa d'aquest moment magnètic es troba en el moment quàntic mecànic angular (moment angular d'spin), que tenen els nuclis atòmics amb un número imparell de protons i neutrons.

Es pot imaginar el nucli com una petita partícula carregada que gira al voltant del seu propi eix, de manera que passa a comportar-se com un petit electromagnet. En aplicar un camp magnètic extern fort aquests nuclis s'hi alinearan, assolint un estat energètic superior. La NMR estudia aquests nuclis magnètics alineant-los a un camp constant i pertorbant aquest alineament utilitzant un altre camp.

De manera general, les mostres de proteïna preparades per NMR es van preparar a una concentració entre 1 i 5 mM en KH_2PO_4 10mM, DTT 1mM, 10% D_2O pH 7. Tots els espectres d'NMR han estat recollits a 5 i 25°C en un espectròmetre Bruker BioSpin AV-800 equipat amb una criosonada, processats amb Topspin (Bruker BioSpin) i analitzats i integrats amb SPARKY (T. D. Goddan and D. G. Kneller; SPARKY 3, University of California, San Francisco).

3.6.1 Assignació de les variants I_NI_C_sc

Es va realitzar l'assignació de les ressonàncies de ^1H , ^{13}C i ^{15}N de les variants I_NI_C_sc marcades amb ^{13}C i ^{15}N . Es van utilitzar un conjunt d'espectres bidimensionals homo i heteronuclears: $^1\text{H}^{13}\text{C}$ HSQC (*heteronuclear single quantum coherence*), $^1\text{H}^{15}\text{N}$ HSQC, TOCSY (*total correlation spectroscopy*) i NOESY (*nuclear Overhauser enhancement spectroscopy*).

El procés d'assignació va ser facilitat per les assignacions prèvies de la proteïna sencera *Npu DnaE* (Heinamaki et al., 2009), dipositades al *Biological Magnetic Resonance Bank* amb el codi 16009.

3.7 ANÀLISI BIOINFORMÀTICA

La composició aminoacídica de les diferents variants de la *Npu DnaE* analitzades de manera bioinformàtica per l'estudi del plegament en aquest treball, van ser analitzades mitjançant el *Composition profiler Software* (Vacic et al., 2007) el qual inclou les bases de dades de Disprot 3.4 i PDB_Select_25. La predicció del plegament de les diferents variants de *Npu DnaE* va ser feta utilitzant el software PONDR® amb l'algoritme VL-XT (Romero et al., 1997). D'aquesta predicció es van poder distingir les variants ordenades de les desordenades segons l'escala de *Kyte-Doolittle*

(Uversky et al., 2000). Les anàlisis de càrrega-hidrofobicitat i de funció de distribució acumulativa (CDF) es van fer també utilitzant PONDR®.

3.8 ANÀLISI ESTADÍSTICA

Es va fer quantificacions del producte total del procés de tall i unió de diferents variants de *Npu DnaE* assajades en aquest treball. Es van fer 3 rèpliques diferents per cada experiment, i a partir dels valors del producte total obtingut, es va fer el test de Anderson-Darling per comprovar la normalitat i el test de Levene per comprovar l'homogeneïtat de variàncies. També es va aplicar una prova t per 2 mostres suposant variàncies desiguals, on cada variable era el percentatge total del producte de tall i unió de dos assajos a comparar. El nivell de significació va ser de $p < 0.05$.

Es va utilitzar el software Microsoft Office Excel 2007 per Windows (Microsoft, USA). Per a la realització del test de normalitat es va utilitzar l'aplicació desenvolupada per Kevin Otto (University of Missouri) i per el test de normalitat es va utilitzar l'aplicació desenvolupada pels serveis estadístics del *College of Liberal Arts & Science* (University of Florida).

4. Resultats i Discussió

4.1 EXPRESSIÓ, PURIFICACIÓ I CARACTERITZACIÓ DE L'ESTABILITAT TÈRMICA DE LA INTEÏNA *Npu DnaE*, DE LES SUBUNITATS I_N – I_C ASSOCIADES I DE I_N AÏLLADA

Una de les línies de recerca del nostre grup se centra en la purificació i caracterització de diferents inteïnes, per tal de poder-les utilitzar com a possibles eines biotecnològiques. Aconseguir un protocol de purificació òptim i una bona caracterització, és essencial per permetre l'estudi i comprensió de les inteïnes tant a nivell molecular com estructural i funcional.

L'interès en la inteïna *Npu DnaE* rau en el fet que presenta unes característiques gairebé ideals per a ser explotada en el camp biotecnològic. Tal i com s'ha detallat a la introducció: se sap que presenta la capacitat de formar el seu producte de tall i unió de manera molt ràpida i eficient i la seva activitat no es veu alterada per l'efecte de la temperatura ni per l'efecte d'un agent químic desnaturalitzant com la urea. Totes aquestes característiques la converteixen en una inteïna molt robusta (Martin et al., 2001; Nichols and Evans, 2004).

Heinamaki i col·laboradors van purificar *Npu DnaE* com una única cadena polipeptídica però resultava inactiva ja que presentava la mutació C1A (Heinamaki et al., 2009). No obstant, un dels primers requisits a l'hora de dissenyar un protocol d'expressió i purificació de *Npu DnaE* en el nostre grup, era que fos activa, tant quan formés part d'una única cadena polipeptídica com quan es trobés en forma de les subunitats I_N i I_C. Estudiar la inteïna quan es troba separada en les dues subunitats o estudiar només la subunitat I_N també havia de permetre comprendre aspectes d'estabilitat i plegament.

4.1.1 Construcció, expressió i purificació de *Npu DnaE* formant una única cadena polipeptídica

4.1.1.1 Construcció de *Npu DnaE* formant una única cadena polipeptídica

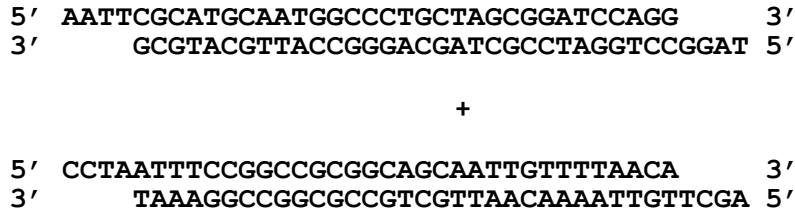
Per tal d'obtenir la inteïna *Npu DnaE* unida mitjançant un enllaç peptídic entre la subunitat I_N i la I_C, es van dissenyar un seguit de passos que van conduir a la construcció d'un vector que contenia la seqüència de la inteïna flanquejada per la seqüència de la GFP la qual constituïa l'exteïna.

Primer de tot es van dissenyar els oligonucleòtids per tal d'obtenir una seqüència de clonatge múltiple (MCS) per tal d'introduir, en el vector pRSET/EmGFP (Invitrogen), les dianes de restricció necessàries per a la clonació de la I_N i la I_C. Per a la síntesi de la seqüència MCS es van dissenyar quatre oligonucleòtids complementaris entre ells, excepte en els extrems, de tal manera que després

4. Resultats i Discussió

de realitzar una hibridació entre ells s'obtinguessin uns extrems que permetessin la seva unió per complementaritat al vector pRSET/EmGFP.

Els oligonucleòtids dissenyats per obtenir la seqüència MCS van ser els següents:



En el disseny i síntesi de les seqüències de tots els oligonucleòtids descrits en aquest treball es va tenir en compte tant el manteniment de la pauta de lectura en la seqüència com la freqüència d'ús del codó en *E. coli*.

L'amplificació de I_N i de I_C de la *Npu DnaE* es va realitzar mitjançant PCR. La subunitat I_N es va amplificar a partir del vector pskDuet01 (Addgene, USA) i la subunitat I_C a partir del vector pskBAD2 (Addgene, USA). Els encebadors utilitzats per a l'amplificació de la I_N i la I_C es detallen a continuació:

Amplificació de la I_N :



Amplificació de la I_C :



Els encebadors es van dissenyar de tal manera que el gen de la subunitat I_N tingués els extrems cohesius següents: a l'extrem 5' portés incorporada la diana de restricció de *SphI* i a l'extrem 3' la diana de restricció de *BamHI*. El mateix es va fer amb el gen de la subunitat I_C fent que a l'extrem 5' hi hagués la diana *EclXI* i a l'extrem 3' la diana *BamHI*.

A partir del mateix vector pRSET/EmGFP es va fer l'amplificació dels gens de la GFP_C i la GFP_N. La GFP continguda en el vector presentava 238 residus, dels quals la GFP_N contenia els residus 1 a 157 i la GFP_C els residus 158 al 238. Es va escollir partir la GFP entre els residus 157 i 158 ja que en estudis previs s'havia demostrat que la partició de la proteïna en aquesta regió no alterava la funció de la proteïna (Lindman et al., 2009; Paramban et al., 2004).

4. Resultats i Discussió

Es van dissenyar uns encebadors per amplificar la GFP_N que permetessin obtenir una diana *NdeI* a l'extrem 5' i una diana *SphI* a l'extrem 3' i en el cas de la GFP_C s'havia d'incorporar una diana *MfeI* a l'extrem 5' i una diana *XhoI* a l'extrem 3'.

Els encebadors dissenyats per a l'amplificació dels gens de la GFP_N i GFP_C foren els següents:

GFP_N:

5' **TAATACGACTCACTATAGGG** 3' (T7 Prom)
3' **GATATAGTGGCGGCGGCTCCGTACGTTGCCGTAGTTCC** 5'

GFP_C:

5' **TATGCTAGTTATTGCTCAG** 3' (T7 Term)
3' **GTGGAACACTACGGCTTTGTTAACAACAGCCGCCACTATA** 5'

Un cop amplificats aquests fragments es va dur a terme la lligació amb el vector pRSET/EmGFP digerit prèviament amb *NdeI* i *XhoI*. Dels clons obtinguts, es van seqüenciar i seleccionar els que tinguessin la seqüència correcta. Aleshores, per mutagènesi dirigida utilitzant QuikChange™ es va eliminar la seqüència MCS que es trobava entre les subunitats I_N i I_C. Els encebadors utilitzats per a l'eliminació de la seqüència MCS són els següents:

5' **GGTTGATAATTTGCCGAATATCAAAATAGCCACACGTAAATATTTAC** 3'
3' **CCAACACTATTAACGGCTTATAGTTTTATCGGTGTGCATTTATAAATG** 5'

Els clons que es van obtenir a partir del QuikChange es van seqüenciar i seleccionar els que contenien la seqüència correcta i es va obtenir així la *Npu DnaE* formant una única seqüència polipeptídica i flanquejada per la GFP.

La construcció d'aquest vector va permetre detectar de manera senzilla si el mecanisme de tall i unió es trobava actiu o no. En transformar cèl·lules d'*E. coli* BL21(DE3) amb el vector, en expressar-se, si el mecanisme de tall i unió era actiu, la GFP es reconstituïria i les cèl·lules adquiririen una tonalitat verda. A la figura 4.1 es mostra una esquema de l'estratègia seguida per la construcció del vector pRSET-GFP_N-*Npu DnaE*-GFP_C.

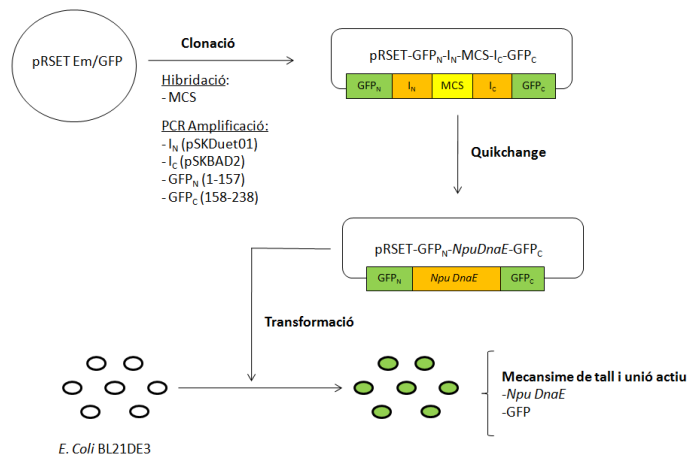


Fig. 4.1. Esquema de l'estratègia d'obtenció del vector pRSET-GFP_N-*Npu DnaE*-GFP_C.

4.1.1.2 Producció i purificació de *Npu DnaE* formant una única cadena polipeptídica

Abans d'iniciar el procés de producció i purificació de *Npu DnaE* formant una única cadena polipeptídica, es va analitzar el grau d'expressió de diferents clons. El fet que les cèl·lules expressaven *Npu DnaE* activa significava que es duia a terme el mecanisme de tall i unió i per tant es reconstituïa i s'expressava la GFP. Per tant, les cèl·lules que expressaven la inteïna funcional, adquirien una coloració verdosa. Es van analitzar electroforèticament cultius de diferents clons i es va seleccionar el que presentava una millor expressió. Amb el clon seleccionat es va obtenir un rendiment de 178 mg de *Npu DnaE* per litre de cultiu.

Es va fer un seguiment del procés de purificació per tal d'analitzar en quines fraccions es trobava *Npu DnaE*. Un cop passades les cèl·lules per la French Press després de la producció, es va voler determinar si la inteïna es trobava en la fracció soluble o insoluble. Després de centrifugar la mostra provinent de la French Press per tal de separar la fracció soluble de la insoluble, es va carregar la mostra provinent del sobrenedant (fracció soluble) i la mostra provinent del sediment (fracció insoluble) en un gel SDS PAGE, el qual va demostrar que la inteïna es trobava a la fracció soluble tal i com es mostra en la figura 4.2.

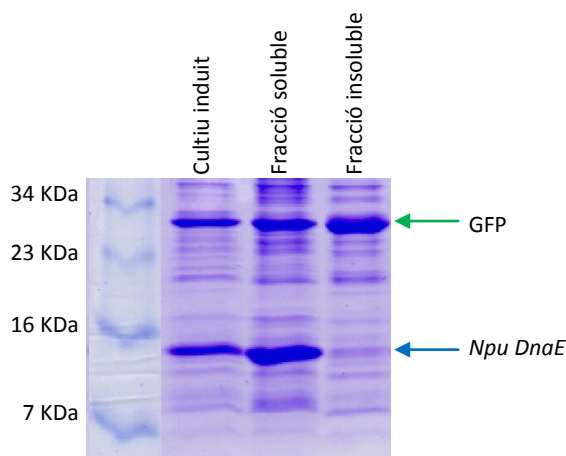


Fig. 4.2. Anàlisi per SDS-PAGE per tal de determinar si *Npu DnaE* es localitzava a la fracció soluble o a la insoluble de la mostra provinent de French Press. La fletxa verda indica la banda pertanyent a la GFP amb una massa molecular teòrica de 27,49 KDa i la fletxa blava indica la banda pertanyent a la *Npu DnaE* amb una massa molecular teòrica de 15,8 KDa.

Es va seguir amb el procés de purificació mitjançant HPLC de bescanvi aniònic a partir de la fracció soluble. Els cromatogrames obtinguts es van enregistrar a 280nm i es van recollir mostres de cadascun dels pics, per posteriorment, analitzar-los en un gel SDS-PAGE per tal de determinar en quin pic eluïa la *Npu DnaE* tal i com es mostra en la figura 4.3.

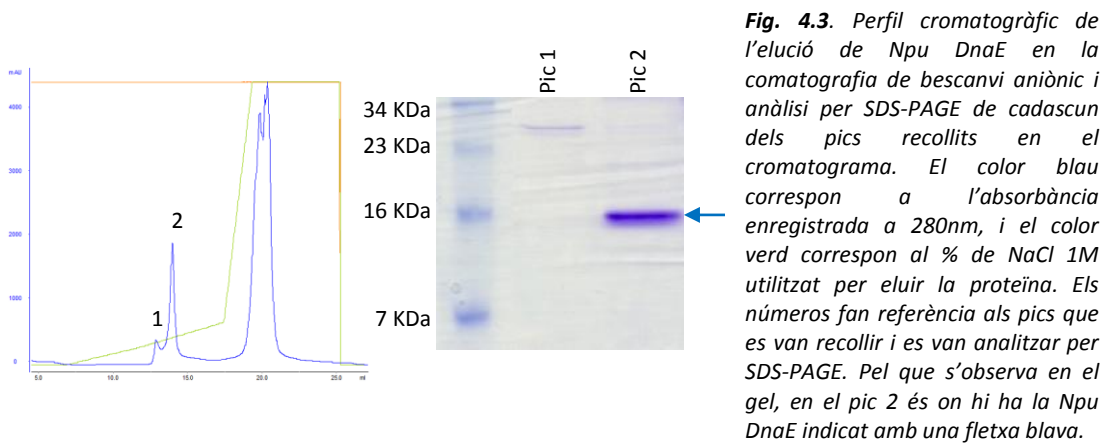


Fig. 4.3. Perfil cromatogràfic de l'elució de *Npu DnaE* en la comatografia de bescanvi aniónic i anàlisi per SDS-PAGE de cadascun dels pics recollits en el cromatograma. El color blau correspon a l'absorbància enregistrada a 280nm, i el color verd correspon al % de NaCl 1M utilitzat per eluir la proteïna. Els números fan referència als pics que es van recollir i es van analitzar per SDS-PAGE. Pel que s'observa en el gel, en el pic 2 és on hi ha la *Npu DnaE* indicat amb una fletxa blava.

Una vegada obtinguda la *Npu DnaE* purificada, es guardà liofilitzada i es va analitzar per espectrometria de masses MALDI-TOF. La seva massa molecular observada va ser de 15832 Da, un valor que no va diferir significativament de l'esperat en el càlcul teòric. Aquest fet confirmava que la proteïna es trobava sense la formació de ponts disulfur ni tampoc es van observar pics d'agregació o dimerització en l'espectre de MALDI-TOF obtingut.

4.1.2 Expressió i purificació de les subunitats I_N - I_C de *Npu DnaE* associades

Per a la producció i purificació de les subunitats I_N i I_C formant complex es va partir d'un vector que ja havia estat construït en el nostre grup, el vector pET22b(+)-cHPR (Novagene, USA). Aquest vector és utilitzat per la ciclació de la ribonucleasa pancreàtica humana (cHPR). cHPR és un zimogen de la ribonucleasa pancreàtica humana amb una seqüència de bloqueig al centre actiu, la qual és reconeguda específicament per la proteasa de tipus 1 del virus de la immunodeficiència humana (HIV-1 PR).

Per tal d'obtenir cHPR, la seqüència del zimogen en el vector es troba flanquejada en el seu extrem N-terminal per la seqüència de la I_C de la *Npu DnaE* i per la seqüència de la I_N de la *Npu DnaE* pel seu extrem C-terminal. Quan s'expressa la construcció i es duu a terme el mecanisme de tall i unió, s'allibera l'exteïna de manera ciclada (cHPR) i com a subproducte, les subunitats I_N i I_C associades formant la *Npu DnaE*. A la figura 4.4 es mostra un esquema de la construcció i de com es duu a terme el mecanisme de tall i unió.

Abans d'iniciar el procés de producció i purificació de I_N i I_C associades, es va analitzar el grau d'expressió de diferents clons. Per això es van analitzar electroforèticament cultius induïts d'aquests clons i es va seleccionar aquell que presentava millor expressió. Amb el clon seleccionat es va obtenir un rendiment de 64 mg de I_N i I_C associades per litre de cultiu.

4. Resultats i Discussió

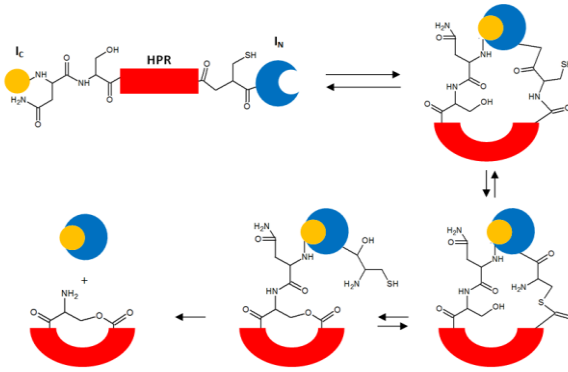


Fig. 4.4. Mecanisme de tall i unió per l'obtenció de cHPR i de I_N i I_C associades formant la *Npu DnaE*.

Per a la purificació de I_N i I_C es van seguir els mateixos passos que en el cas de la purificació de la *Npu DnaE* (apartat 3.4.3 de material i mètodes). A la figura 4.5 es mostra un cromatograma de bescanvi aniónic d'HPLC enregistrat a 280nm de la fracció soluble del cultiu. Es va recollir mostra de cadascun dels pics obtinguts en el cromatograma i es va carregar en un gel SDS-PAGE per tal de determinar en quin dels pics s'elua la I_N i la I_C .

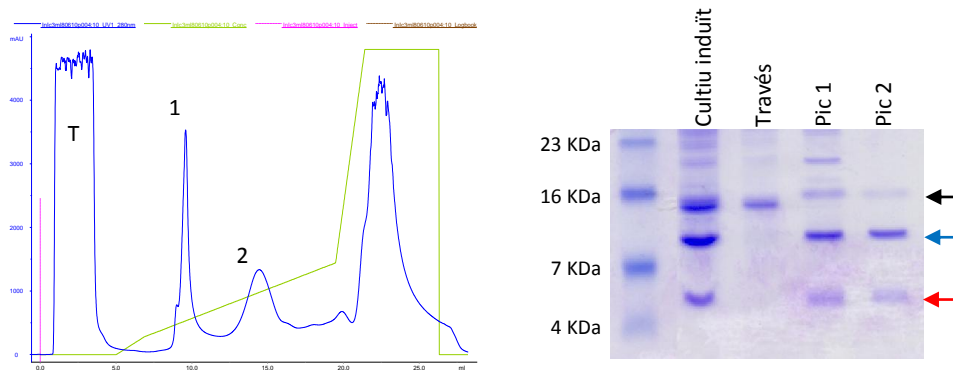


Fig. 4.5. Perfil cromatogràfic de l'elució de I_N i I_C associades en la cromatografia de bescanvi aniónic i anàlisi per SDS-PAGE de cadascun dels pics recollits en el cromatograma. El color blau correspon a l'absorbència enregistrada a 280nm, i el color verd correspon al % de NaCl 1M utilitzat per eluir la proteïna. Els números fan referència als pics que es van recollir i es van analitzar per SDS-PAGE. Pel que s'observa en el gel, en el través és on elueix la cHPR, en els pic 1 i 2 és on hi ha la I_N (fletxa blava) i la I_C (fletxa vermella) i el complex I_N i la I_C formant la *Npu DnaE* (fletxa negra).

A partir dels resultats del gel SDS-PAGE de la figura 4.5 es va poder determinar que el mecanisme de tall i unió era actiu, ja que en el carril on hi havia la mostra del cultiu induït es poden diferenciar tres bandes pertanyents a la cHPR, la I_N i la I_C . Aquest fet indicava doncs, que la I_N i la I_C es trobaven actives. A més, es va determinar que cHPR eluïa en el través i les subunitats I_N i I_C eluïen de manera conjunta tant en el pic 1 com el 2. No obstant, en el pic 2 aquestes eluïen de manera més pura i homogènia, mentre que en el pic 1 hi havia presència d'altres proteïnes. No s'ha sabut trobar una raó per la qual I_N i I_C eluïen juntes i en 2 pics diferents en el cromatograma.

Per aquest motiu es va decidir aprofitar només la mostra provinent del pic 2 i aquesta es va guardar liofilitzada i es va analitzar per espectrometria de masses MALDI-TOF. La seva massa molecular

observada va ser de 11856 Da en el cas de la subunitat I_N i de 4126 Da en el cas de la subunitat I_C, uns valors que no va diferir significativament dels esperats en el càlcul teòric. Aquest fet confirmava que les subunitats es trobaven sense la formació de ponts disulfur ni tampoc es van observar pics d'agregació o dimerització en els espectres de MALDI-TOF obtinguts.

4.1.3 Construcció, expressió i purificació de la subunitat I_N de *Npu DnaE*

Per tal de purificar la subunitat I_N es va decidir primerament fer-ho amb una cromatografia de gel filtració a partir de la mostra provinent del pic 2 de la purificació de I_N i I_C associades. Els resultats obtinguts però no van ser satisfactoris ja que, les interaccions entre I_N i I_C són tant fortes que ambdues subunitats eluïen en el mateix pic a la columna de gel filtració.

Per tant, es va posar a punt un altre protocol per tal d'expressar i purificar la subunitat I_N, en el qual aquesta s'expressés de manera aïllada de la subunitat I_C.

4.1.3.1 Construcció de la subunitat I_N de *Npu DnaE*

El fet que la I_N s'expressés de manera aïllada feia pensar que aquesta es podria expressar de manera insoluble en la cèl·lula tal i com s'ha descrit per algunes subunitats I_N d'altres inteïnes (Zheng et al., 2012). L'interès que la subunitat I_N es trobés en la fracció soluble rau bàsicament en facilitar el seu procés de purificació i que no es trobés agregada. L'estratègia a seguir per tal que la subunitat I_N s'expressés de forma soluble va ser fusionar-la a una proteïna de solubilització. La sèrie de vectors pBADm (EMBL) tenen diferents proteïnes de solubilització seguides per la diana de la proteasa TEV i estan dissenyats per purificar proteïnes solubles de manera ràpida. Dels 4 vectors que componen la sèrie pBADm es va escollir el pBADm41 que presenta la MBP com a proteïna de solubilització (Kaput and Waugh, 1999). MBP no presenta cap cisteïna en la seva seqüència (Duplay et al., 1984) i per tant això disminuiria les possibles interaccions amb les cisteïnes de la subunitat I_N que podrien afectar el seu plegament i/o la seva solubilitat.

El primer pas a realitzar va ser l'obtenció del gen de la subunitat I_N (figura 4.6) que es va obtenir fusionat amb el gen del domini GB1 a partir del vector pskDuet 01. Digerint el vector pskDuet 01 amb *NcoI* i *HindIII* s'obtenia el gen de la subunitat I_N amb GB1 el qual es va lligar al vector pBADm41 digerit prèviament amb *NcoI* i *HindIII*. Els clons que es van obtenir a partir de la lligació es van seqüenciar i seleccionar els que contenien la seqüència correcta.

A partir d'aquí es va voler determinar la solubilitat de la construcció MBP-I_N-GB1. Per tant, es van fer produccions amb aquesta construcció a 37 i a 18°C i les fraccions soluble i insoluble de cadascuna es van analitzar per gel SDS-PAGE. En el gel es va observar que en ambdós casos la proteïna es trobava

4. Resultats i Discussió

tant en la fracció soluble com la insoluble, però a 18°C es trobava majoritàriament a la fracció soluble. Per tant, fent la producció d'aquesta construcció a 18°C s'obtenia un major quantitat de proteïna a la fracció soluble.

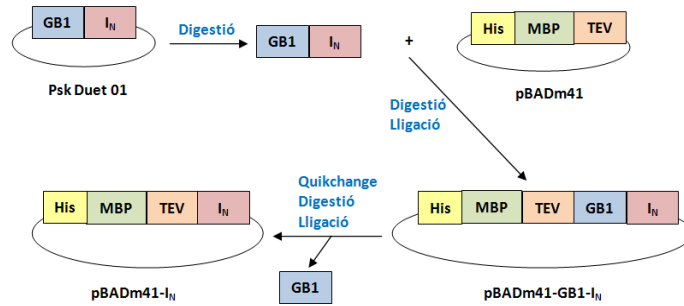


Fig. 4.6. Esquema de l'estratègia d'obtenció del vector pBADm41-I_N.

Per tal de determinar que la subunitat I_N es trobava activa en aquesta construcció, es van fer proves de tall i unió amb la subunitat I_C fusionada al domini GB1, construcció continguda en el vector psKBAD2. Abans de les proves de tall unió es va fer una digestió amb la TEV per tal d'eliminar la MBP de la subunitat I_N. A la figura 4.7 es mostren els resultats d'aquestes proves de tall i unió.

Els resultats del gel mostren que el mecanisme de tall i unió és actiu i que el producte de tall i unió comença a ser significatiu a partir del minut 3 de la reacció. Aquests resultats confirmen que la subunitat I_N es trobava activa dins la construcció realitzada.

El següent pas consistí en l'eliminació del domini GB1 que es trobava entre la MBP i la subunitat I_N (figura 4.6). Per fer-ho, es van dissenyar oligonucleòtids per tal d'introduir la diana *NcoI* a l'extrem C-terminal de la seqüència del domini GB1 pel mètode de QuikChange. D'aquesta manera el domini GB1 tindria la diana *NcoI* tant en el seu extrem N com C-terminal.

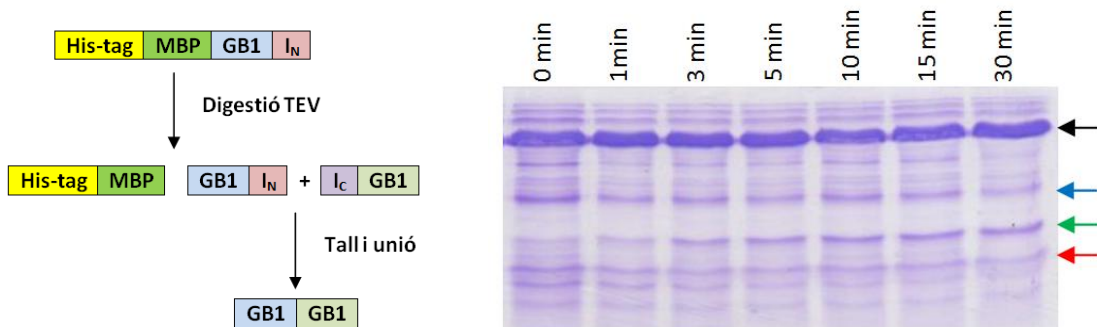


Fig. 4.7. Esquema i anàlisi per SDS-PAGE de les proves de tall i unió entre la construcció MBP-GB1-I_N digerida prèviament amb TEV i la construcció I_C-GB1. La fletxa negra indica la banda pertanyent a la His-MBP, la blava indica la banda pertanyent a GB1-I_N, la vermella indica la banda pertanyent a I_C-GB1 i la verda indica la banda pertanyent a la formació del producte de tall i unió GB1-GB1.

Els oligonucleòtids dissenyats per la introducció d'una diana *NcoI* a l'extrem C-terminal de la seqüència GB1 eren els següents:

5' CCAAAACCTACACCGTAACGGCCATGGGCTGTTTAAGCTATGAAACGG 3'
3' CCGTTTCATAGCTTAAACAGCCCATGGCCGTTACGGTGTAGGTTTTGG 5'

Els clons obtinguts a partir de QuikChange es van digerir amb *NcoI* i es va verificar emprant un gel d'agarosa que s'escindia la banda pertanyent al gen del domini GB1. Dels clons que s'havia escindit el gen de la GB1, es va fer una lligació i dels clons obtinguts es van seqüenciar i seleccionar els que contenien la seqüència correcta obtenint el vector pBADm41-MBP-I_N, amb la diana reconeguda per la proteasa TEV entre MBP i la subunitat I_N (figura 4.6). A partir d'aquest vector es va establir el procés de purificació de la subunitat I_N.

4.1.3.2 Producció i purificació de la subunitat I_N de *Npu DnaE*

Abans d'iniciar el procés de producció i purificació de la subunitat I_N de *Npu DnaE* es va analitzar el grau d'expressió de diferents clons. Es van analitzar electroforèticament cultius de diferents clons i es va seleccionar el que presentava una millor expressió. Amb el clon seleccionat es va obtenir un rendiment de 12 mg de la subunitat I_N per litre de cultiu.

Es va fer un seguiment del procés de purificació de la subunitat I_N per tal d'analitzar en quines fraccions de les cromatografies de bescanvi aniònic i gel filtració es trobava la subunitat I_N. Un cop passades les cèl·lules per la French Press després de la producció, la subunitat I_N juntament amb la proteïna de solubilització es trobava majoritàriament a la fracció soluble. A partir d'aquí es va realitzar una cromatografia de Ni-NTA (apartat 3.4.4 de material i mètodes) per tal d'aïllar MBP-I_N de la resta de proteïnes de la mostra. Un cop aïllada parcialment MBP-I_N, es va digerir amb la proteasa TEV per trencar l'enllaç peptídic entre la proteïna de solubilització i la subunitat I_N. Es va realitzar una segona cromatografia Ni-NTA per tal que la subunitat I_N eluís al través i la MBP quedés retinguda a la reïna ja que reté la cua d'histidines. I_N es va eluir al través però no de manera pura i per tant es va seguir amb el procés de purificació mitjançant HPLC de bescanvi aniònic carregant la fracció recollida de la segona cromatografia de Ni-NTA. Els cromatogrames obtinguts es van enregistrar a 280nm i es van recollir mostres de cadascun dels pics, per posteriorment, analitzar-los en un gel SDS-PAGE i determinar en quin pic eluïa la subunitat I_N, procés que es mostra en la figura 4.8.

Amb la cromatografia de bescanvi aniònic la subunitat I_N no es trobava pura del tot ja que apareixien altres bandes de proteïnes de pes molecular molt superior a la I_N. Per tant, es va continuar amb el procés de purificació mitjançant una cromatografia de gel filtració d'HPLC. Els cromatogrames obtinguts es van enregistrar a 280nm i es van recollir mostres de cadascun dels pics, per posteriorment, analitzar-los en un gel SDS-PAGE per tal de determinar en quin pic eluïa la subunitat I_N tal i com es mostra en la figura 4.9.

4. Resultats i Discussió

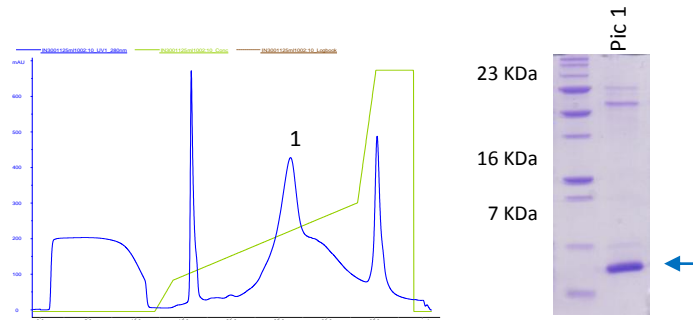


Fig. 4.8. Perfil cromatogràfic de l'elució de la subunitat I_N en la cromatografia de bescanvi aniónic i anàlisi per SDS-PAGE del pic recollit en el cromatograma. El color blau correspon a l'absorbància enregistrada a 280nm, i el color verd correspon al % de NaCl 1M utilitzat per eluir la proteïna. El número fa referència al pic que es va recollir i es van analitzar per SDS-PAGE. Pel que s'observa en el gel, en el pic 1 és on hi ha la I_N (fletxa blava).

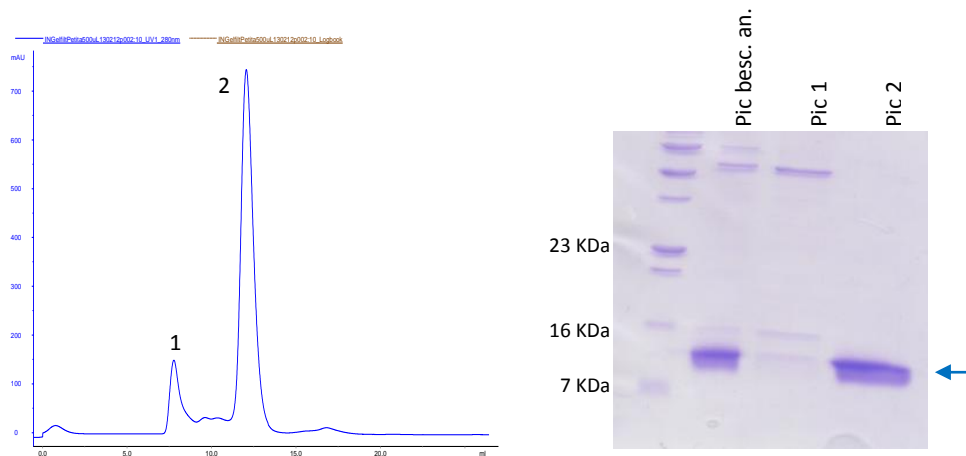


Fig. 4.9. Perfil cromatogràfic de l'elució de la subunitat I_N en la cromatografia de gel filtració i anàlisi per SDS-PAGE de cadascun dels pics recollits en el cromatograma on el color blau correspon a l'absorbància enregistrada a 280nm. Els números fan referència als pics que es van recollir i es van analitzar per SDS-PAGE, el qual també es va carregar mostra provinent del pic 1 de la cromatografia de bescanvi aniónic. Pel que s'observa en el gel, en el pic 2 és on hi ha la I_N (fletxa blava).

Mitjançant cromatografia de gel filtració es va obtenir finalment la subunitat I_N purificada. Aquesta es va liofilitzar i es va analitzar per espectrometria de masses MALDI-TOF. El valor de la massa molecular de la subunitat I_N observada va ser de 12172 Da i no diferia significativament de l'esperada a partir del càlcul teòric. Aquest fet indicava que la subunitat I_N es trobava sense la formació de ponts disulfur ni tampoc s'observava cap pic que pogués indicar l'agregació o dimerització de la proteïna.

4.1.4 Caracterització de la inteïna *Npu DnaE*, de les subunitats $I_N - I_C$ associades i de I_N aïllada

4.1.4.1 Estabilitat tèrmica de cada una de les variants de *Npu DnaE*

Per tal de poder calcular l'estabilitat de les diferents variants i comparar-les entre elles, es van enregistrar corbes de desnaturalització tèrmica a diferents valors de pH. En tots els casos s'observava una transició sigmoïdal i unes regions pre- i posttransició gairebé lineals (figura 4.10).

4. Resultats i Discussió

Aquesta transició correspon al procés de desnaturalització proteica. Convé ressaltar que les transicions espectrals per *Npu DnaE* i I_N - I_C associades que es van obtenir van ser semi reversibles i irreversibles en el cas de la subunitat I_N .

Cal esmentar que la reversibilitat del procés va ser superior quan la *Npu DnaE* es trobava formant part de una única cadena polipeptídica que no pas quan aquesta es trobava associada en forma de I_N i I_C . Aquest fet suggereix que en el procés de repliegament hi hauria una sinèrgia mútua entre la subunitat I_N i I_C que contribuirien al repliegament global de la proteïna. Quan la proteïna es troba formant una única cadena polipeptídica pot haver-hi una cooperació màxima ja que les dos subunitats es troben molt properes entre elles i per això s'observava un repliegament significatiu màxim a pH 8 al voltant del 75%, mentre que quan I_N i I_C es troben associades, el fet de dissociar i desnaturalitzar-se disminuïria la posterior capacitat d'associació entre I_N i I_C i per tant el procés de repliegament es veuria afectat. En aquest cas, s'observava un repliegament màxim a pH 6 i 8 al voltant de $\pm 35\%$. No obstant, en el cas de pH 6 la desviació estàndard és bastant alta indicant que el procés de repliegament a aquest pH no era tant fiable. En el cas de la subunitat I_N sola, al trobar-se poc estructurada per sí sola, el fet de desnaturalitzar-se li impedeix tornar-se a repliegament totalment ja que no té la subunitat I_C que l'ajuda en el repliegament. El percentatge de repliegament final de cada una de les variants de *NpuDnaE* assajades es troba recollit en la taula 4.1.

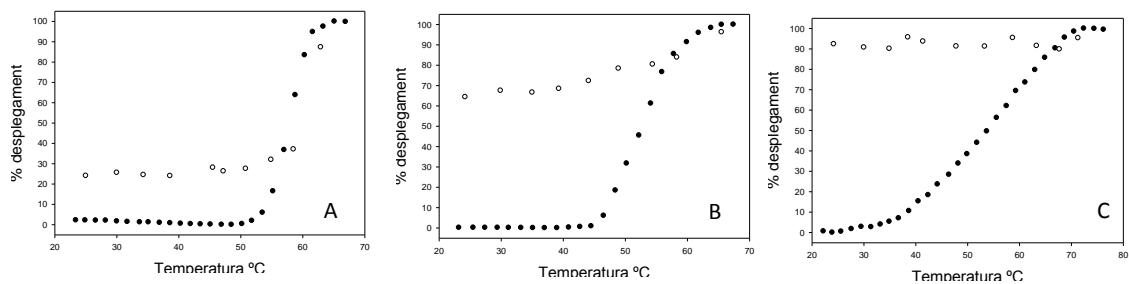


Fig. 4.10. Corbes normalitzades de desplegament (●) i repliegament (○) per temperatura de *Npu DnaE* (A), I_N i I_C associades (B) i la subunitat I_N (C). Els espectres es van realitzar entre 20 i 80°C amb increments de 2°C en el procés de desplegament i decrements de 5°C en el procés de repliegament. La concentració de proteïna va ser de 0,5 mg/ml en HEPES 50mM, NaCl 150mM, TCEP-HCl 1mM a pH 8.

Taula 4.1. Efecte del pH en el repliegament de les diferents variants de *Npu DnaE*

		pH*						
		5	6	7	7.5	8	8.5	9
% repliegament	<i>NpuDnaE</i>	46.4 ± 9.3	61.9 ± 2.2	70.5 ± 1.2	73.8 ± 5.0	76.4 ± 0.7	72.5 ± 5.4	68.1 ± 1.5
	I_N - I_C	-	35.9 ± 11	12.7 ± 7.2	8.5 ± 4.7	36.1 ± 0.6	22.2 ± 5.9	-
	I_N	-	-	-	-	-	-	-

*Tots els tampons utilitzats van ser a una concentració de 50mM amb NaCl 150mM, TCEP-HCl 1mM. Es va utilitzar acetat sòdic per pH 5, MES per pH 6, HEPES per pH 7, 7.5 i 8 i Tris per pH 8.5 i 9. En els casos que no hi ha valor indica que no es va observar repliegament.

4. Resultats i Discussió

Aquest procés de sinèrgia mutu en el qual les subunitats I_N i I_C contribueixen en el plegament global de la inteïna va ser observat també en el cas de la inteïna *Ssp DnaE* (Zheng et al., 2012). Per tal de verificar la irreversibilitat del procés de desplegament es van fer mesures d'absorbància de cadascuna de les mostres a diferents intervals de temps fins a 24h a 20°C després del desplegament per temperatura i no es van observar diferències respecte el valor inicial d'absorbància obtingut en acabar les corbes de replegament. Això indicava que totes les mostres mantenien el grau de desplegament obtingut en acabar la corba de replegament i que en el cas de la subunitat I_N , aquest desplegament era completament irreversible.

Les corbes de desnaturalització es van ajustar a un model termodinàmic de dos estats, el qual es va combinar amb funcions lineals per tal d'ajustar la pendent de la pre- i posttransició (Lange et al., 1996). A la taula 4.2 es presenten els valors de T_m que es van calcular a partir de les transicions de desplegament induïdes per temperatura i a diferents pHs de cadascuna de les variants estudiades.

Taula 4.2. Efecte del pH en el valor de T_m en les diferents variants de *Npu DnaE*

		pH*						
		5	6	7	7.5	8	8.5	9
T _m °C	<i>NpuDnaE</i>	49.3 ± 1.2	53.0 ± 0.2	54.5 ± 1.0	56.9 ± 0.3	57.5 ± 0.7	59.0 ± 0.2	61.7 ± 0.2
	I_N - I_C	-	49.8 ± 0.1	51.7 ± 0.1	52.4 ± 0.2	52.6 ± 0.1	52.4 ± 0.1	56.4 ± 0.7
	I_N	-	-	40.2 ± 2.8	49.3 ± 1.5	55.3 ± 0.7	-	-

*Tots els tampons utilitzats van ser a una concentració de 50mM amb NaCl 150mM, TCEP-HCl 1mM. Es va utilitzar acetat sòdic per pH 5, MES per pH 6, HEPES per pH 7, 7.5 i 8 i Tris per pH 8.5 i 9. En els casos que no hi ha valor de T_m indica que la proteïna es desnaturalitzava tant bon punt es ressuspenia amb el tampó. El rang de valors de pH assajats va ser de 3 a 12 però només es van obtenir valors de T_m entre 5 i 9.

Els valors de T_m obtinguts per les tres variants a pH 8 se situa a l'entorn dels 55°C. Recentment s'ha descrit que *Npu DnaE* presenta una T_m de 84°C en Tris 20mM pH 7 (Aranko et al., 2013). Aquesta diferència d'estabilitat observada podria ser deguda a què Aranko i col·laboradors van utilitzar espectrofluorimetria enlloc de espectrofotometria i una composició en els amortidors diferents pel càlcul de la T_m . Cal destacar que l'estabilitat tèrmica de la *Npu DnaE* i les subunitats I_N I_C associades es manté en el rang de pH entre 6 i 9 mentre que en el cas de la subunitat I_N tant els valors de T_m com la seva solubilitat es veuen significativament afectats pels canvis de pH. En les tres variants assajades es manté el fet que el valor de T_m incrementa com més bàsic és el pH. A pH superior a 9 però, les mostres ja precipitaven.

En aquests valors de pH lleugerament bàsics, similars als que es podrien trobar a l'interior de la cèl·lula, és on hi ha una major estabilitat de la inteïna i en aquest cas la Cys 1 i la Cys +1 podrien trobar-se lleugerament ionitzades. D'aquesta manera podrien actuar millor com a nucleòfils facilitant el procés de tall i unió. Al ser pH 8 on la inteïna es troba més estable, la majoria d'assajos que s'han realitzat en aquesta tesi doctoral han estat a aquest pH.

4.2 ESTUDI DEL PROCÉS DE TALL I UNIÓ

Actualment hi ha catalogades unes 600 inteïnes de les quals menys del 5% són inteïnes partides i la majoria d'aquestes pertanyen a la família "*cyanobacterial split DnaE inteïns*". Aquesta família d'inteïnes es caracteritza sobretot per la rapidesa en la qual efectuen el procés de tall i unió (Caspi et al., 2003; Perler, 2002). Dins d'aquesta família les dues inteïnes més conegudes són la *Ssp DnaE* i la *Npu DnaE*. Aquesta última és capaç de catalitzar el procés de tall i unió en trans amb un temps de vida mitja ($t_{1/2}$) d'un minut respecte d'altres inteïnes que poden presentar un $t_{1/2}$ d'hores (Zettler et al., 2009).

Malgrat l'elevada eficiència que presenta *Npu DnaE* en la catàlisi del procés de tall i unió, les seqüències flanquejants de la inteïna són essencials per dur a terme correctament aquest procés. Si bé la regió N-terminal de la exteïna no requereix d'una gran especificitat, la regió C-terminal requereix de la seqüència Cys+1 – Phe+2 – Asn+3 per una catàlisi eficient (Lockless and Muir, 2009). Modificacions en aquests tres residus altera significativament el procés de tall i unió. En diferents estudis s'ha descrit que la Cys +1 és un residu nucleofílic essencial pel mecanisme de tall i unió ja que es troba involucrat en la formació de l'intermediari éster del pas 1 del procés de tall i unió (figura 1.5 de la introducció), el residu +2 pot ser substituït només per Gly, Glu o Arg afectant significativament l'eficiència de la inteïna i la posició +3 només pot ser substituïda per Glu o Tyr afectant també la seva eficiència. Només la substitució a la posició +2 per Trp o Met sempre que es conservin les posicions +1 i +3 es pot aconseguir una eficiència i rapidesa en el procés de tall i unió similars als de la seqüència nativa Cys-Phe-Asn (Cheriyen et al., 2013; Lockless and Muir, 2009; Shah et al., 2012).

Npu DnaE és una inteïna molt interessant dins el camp de les aplicacions biotecnològiques a causa de la seva eficiència en el procés de tall i unió. Un dels principals inconvenients però alhora de utilitzar-la és que el residu +1 hagi de ser una Cys. Aquesta Cys es troba en realitat en la seqüència de l'exteïna i pot alterar de manera significativa la seva funció i plegament així com promoure interaccions no desitjades amb altres proteïnes. Substituir la Cys per un altre residu nucleofílic com la Ser i obtenir una bona eficiència en la catàlisi de la inteïna seria molt interessant des d'un punt de vista biotecnològic. En el cas de la *Ssp DnaE* la substitució d'una Ser a la posició +1 redueix fins a set vegades la velocitat de formació del producte de tall i unió, però no altera la formació de producte total a les 24h (Nichols and Evans, 2004).

En aquest treball s'ha estudiat quin efecte té la substitució de la Cys +1 i Cys1 per una Ser en l'eficiència de tall i unió en la *Npu DnaE* tant quan realitza el tall d'unió en cis com en trans. En el cas de l'estudi de tall i unió en cis es va utilitzar la GFP com a exteïna i en el cas del tall i unió en trans la GB1. Aquests estudis s'han completat mesurant l'eficiència del procés a diferents temperatures,

tenint present que l'increment de temperatura pot facilitar el procés de tall i unió influïent en l'estat de ionització de la Ser i la mobilitat dels diferents residus catalítics.

4.2.1 Estudi del procés de tall i unió en cis de la *Npu DnaE* C+1S utilitzant la GFP com a exteïna

4.2.1.1 Construcció, expressió i purificació de GFP_N-*Npu DnaE*-GFP_C C+1S

Per a l'obtenció d'aquesta construcció es va utilitzar el vector utilitzat per la purificació de la *Npu DnaE* formant una única cadena polipeptídica. Emprant QuikChangeTM es va introduir la mutació C+1S en la seqüència de la GFP_C. Els oligonucleòtids utilitzats foren els que es detallen a continuació:

5' GGCTTCATAGCTTCCAATACCTTCGGCATCAAGGTG 3'
3' CCGAAGTATCGAAGGTTATGGAAGCCGTAGTTCCAC 5'

Els clons que es van obtenir a partir del QuikChange es van seqüenciar i es van seleccionar els que contenien la seqüència correcta. El vector pRSET-GFP_N-*Npu DnaE*-GFP_C amb la mutació C+1S es va emprar per transformar cèl·lules BL21 (DE3) per tal de realitzar la producció de GFP_N-*Npu DnaE*-GFP_C C+1S seguint el mateix protocol que l'utilitzat per la producció de *Npu DnaE* formant una única cadena polipeptídica (apartat 3.4.1 de material i mètodes).

Acabada la producció, es van recollir les cèl·lules i es va separar la fracció soluble de la insoluble i es va observar que GFP_N-*Npu DnaE*-GFP_C es trobava a la fracció soluble i que només hi havia una banda pertanyent al precursor GFP_N-*Npu DnaE*-GFP_C, la qual cosa, indicava que amb la mutació C+1S el mecanisme de tall i unió era inapreciable a 37°C amb les tècniques emprades. De la fracció soluble es va realitzar una cromatografia d'afinitat Ni-NTA seguint el protocol estàndard (apartat 3.4.4 de material i mètodes) per tal de purificar el precursor i poder realitzar les diferents proves de tall i unió.

4.2.1.2 Estudi del procés de tall i unió de GFP_N-*Npu DnaE*-GFP_C C+1S

Per tal de realitzar l'estudi del procés de tall i unió de GFP_N-*Npu DnaE*-GFP_C C+1S en funció de la temperatura, el precursor purificat es va incubar a 20, 30, 40, 50 i 60°C durant diferents temps fins a un màxim de 24h. No es van realitzar proves a més de 60°C ja que la inteïna es trobaria a temperatures superiors a la seva T_m i per tant es trobaria desnaturalitzada. Una vegada el precursor va ser incubat a diferents temps i diferents temperatures, es van recollir i analitzar electroforèticament per SDS-PAGE cadascuna de les mostres. A la figura 4.11 es mostren els gels SDS-PAGE corresponents a 40, 50 i 60°C que són les temperatures les quals el procés de tall i unió es

4. Resultats i Discussió

feia evident malgrat que als 30°C ja es començava a observar una banda molt tènue corresponent al producte de tall i unió.

Per tal de verificar que el procés de tall i unió a aquestes temperatures era actiu i que la banda corresponent al fragment de tall i unió adjudicada era la correcta, es van realitzar dos western blots un amb anticòs anti-GFP per detectar el producte de tall i unió (GFP) i l'altre amb un anticòs anti-Histag ja que GFP_N presenta una cua d'histidines. Els resultats dels western blots es mostren a la figura 4.12. L'anticòs anti GFP reconeix la regió N-terminal de la GFP plegada, no obstant, no es va aconseguir observar la banda pertanyent a GFP_N en el western blot que es va emprar l'anticòs anti GFP.

Els resultats dels western blots van confirmar que el procés de tall i unió era actiu i que aquest era més efectiu com més elevada és la temperatura. Cal esmentar que a temps 0 s'observaven també bandes del producte de tall i unió. Aquest fet es deu a que el precursor, un cop purificat, es va dialitzar durant 24h a 4°C per eliminar l'imidazol provinent de la cromatografia de Ni-NTA. Durant aquestes 24h a 4°C ja hi va haver una certa activitat basal de la inteïna responsable de les bandes a temps 0.

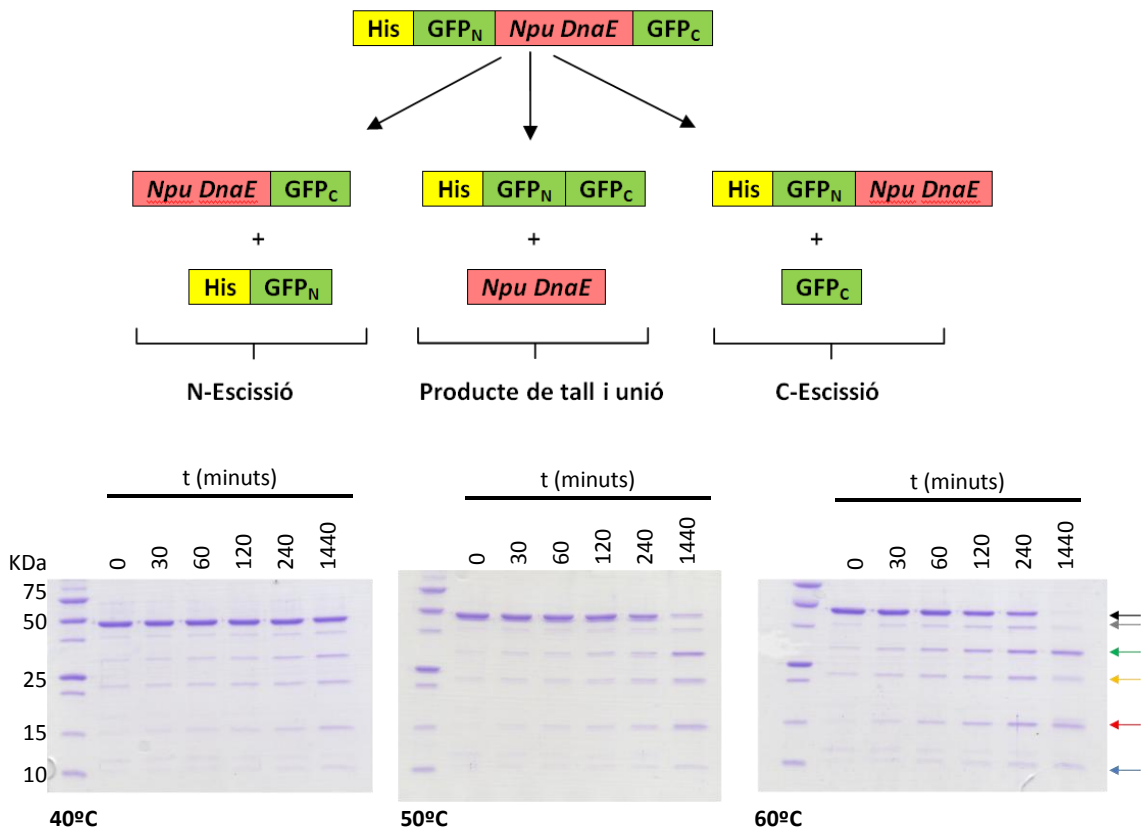


Fig. 4.11. Esquema i gels SDS-PAGE de les proves de tall i unió in vitro a 40, 50 i 60°C a diferents temps de GFP_N-Npu DnaE-GFP_C amb la mutació C+1S. La fletxa negra indica el precursor (47205 Da), la fletxa gris indica GFP_N-Npu DnaE procedent de la C-escissió (38182 Da), la fletxa verda indica la GFP producte del tall i unió (31408 Da), la fletxa groga indica Npu DnaE-GFP_C o GFP_N procedent de la N-escissió (24856 -22367 Da), la fletxa vermella indica la Npu DnaE i la fletxa blava indica la GFP_C procedent de la C-escissió.

4. Resultats i Discussió

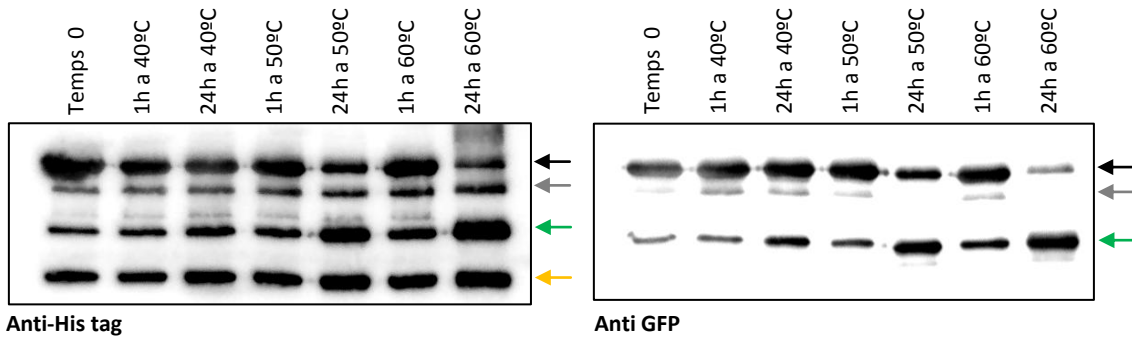


Fig. 4.12. Western blots de les proves de tall i unió *In vitro* a 40, 50 i 60°C a 1h i 24h de GFP_N-Npu DnaE-GFP_C amb la mutació C+1S. La fletxa negra indica el precursor, la fletxa gris indica GFP_N-Npu DnaE procedent de la C-escissió, la fletxa verda indica la GFP producte del tall i unió i la fletxa groga indica GFP_N procedent de la N-escissió.

Malgrat que el procés de tall i unió tenia lloc amb la mutació C+1S, s’observava la formació de bandes corresponents a reaccions no desitjades del procés de tall i unió com són la N-escissió i la C-escissió. Aquest fet no es va observar en el cas de la construcció en la qual el residu +1 era una Cys utilitzat per a la purificació de la *Npu DnaE* formant una única cadena polipeptídica. Encara que a 60°C és on s’observava una major quantitat de producte de tall i unió, calia veure a quina temperatura la formació del producte de tall i unió respecte l’aparició de N- i C-escissió era més eficient. A més, es van calcular les constants k_{obs} i $t_{1/2}$ a partir de la quantificació dels gels SDS-PAGE mitjançant el programa Gel Analyzer.

A la figura 4.13 es mostra en el gràfic (a) l’evolució en percentatge de la formació de producte de tall i unió al llarg de les hores per cada una de les temperatures assajades i en el gràfic (b) es mostra la formació de producte de tall i unió total a les 24h per cada una de les temperatures assajades.

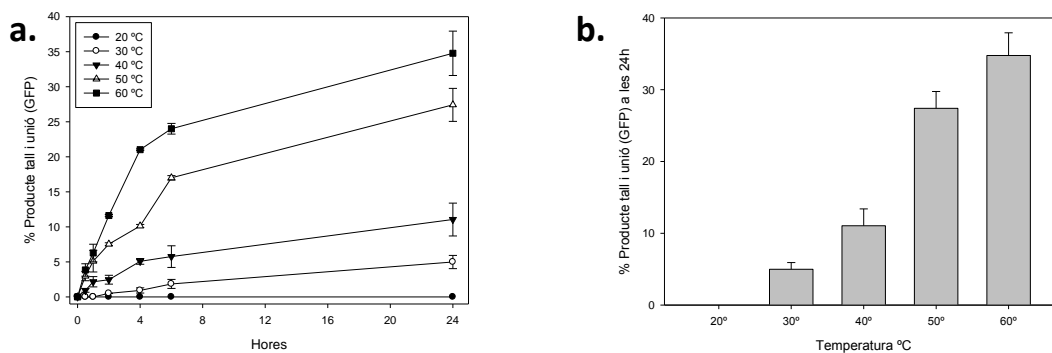


Fig. 4.13. (a) Formació del producte de tall i unió *in vitro* al llarg del temps a diferents temperatures amb la mutació C+1S. (b) Formació de producte de tall i unió total *in vitro* a les 24h per cada una de les temperatures assajades.

A 20°C el procés de tall i unió en aquest cas es trobava indetectable amb les tècniques emprades, mentre que a 30 i 40 °C hi havia formació de producte de tall i unió però aquest es trobava només al voltant del 10 % de la quantitat total de proteïna de la mostra. És a 50 i 60°C on s’observava una major activitat de la inteïna formant a les 24h al voltant del 30% de producte de tall i unió respecte la

4. Resultats i Discussió

quantitat total de proteïna i gairebé un 40% als 60°C. Els gràfics de la figura 4.14 mostren per cada una de les mostres a diferents temps i diferents temperatures el percentatge de cada una de les fraccions resultants del procés de tall i unió respecte la quantitat de proteïna total.

A la figura 4.15 es mostren per cada temperatura el percentatge a les 24h del producte de tall i unió, C-escissió i N-escissió. En les figures 4.14 i 4.15 s'observa ja a primera vista com a mesura que va incrementant la temperatura major era el producte de tall i unió com també les reaccions de N- i C-escissió. Malgrat que aquestes reaccions no desitjades també incrementaven amb la temperatura, el producte de tall i unió sempre era significativament superior a aquestes altres reaccions.

A partir de les quantificacions de cada una de les fraccions resultants del procés de tall i unió es va poder realitzar el càlcul de K_{obs} i $t_{1/2}$ del producte de tall i unió i de N- i C-escissió dels 30 als 60°C. No es va fer a 20°C ja que el sistema es trobava pràcticament bloquejat. A partir dels càlculs d'aquests paràmetres i del rendiment obtingut (producte de tall i unió total a les 24h) es podria veure a quina de les temperatures el procés de tall i unió era més eficient. Els paràmetres K_{obs} i $t_{1/2}$ es van calcular ajustant la formació de producte de tall i unió i N- i C-escissió en funció del temps a una corba exponencial de primer ordre. Els valors obtinguts es mostren a la taules 4.3 i 4.4.

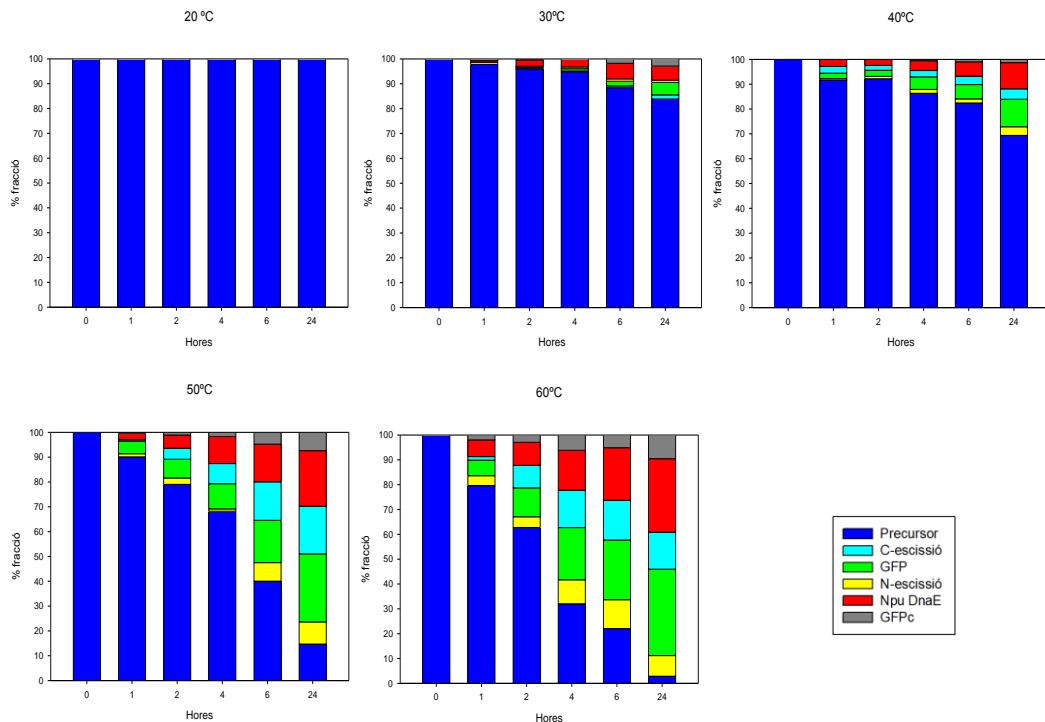


Fig. 4.14. Gràfics de les proves de tall i unió *In vitro* a 20, 30, 40, 50 i 60°C a diferents temps on es mostra el percentatge de cada una de les fraccions resultants del producte de tall i unió respecte la quantitat total de proteïna de cada una de les mostres.

4. Resultats i Discussió

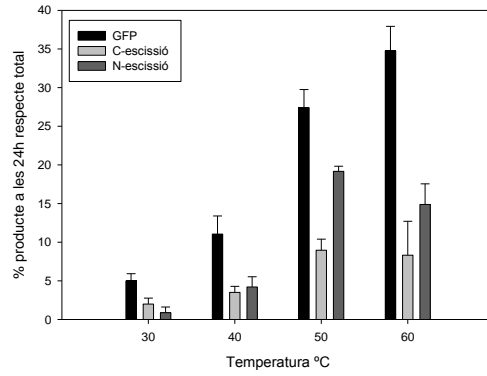


Fig. 4.15. Gràfic on es mostra per cada temperatura el percentatge a les 24h del producte de tall i unió, C-escissió i N-escissió.

Taula 4.3. Valors de la K_{obs} observada en les diferents reaccions de tall i unió in vitro de *NpuDnaE C+1S* en cis*

		Temperatura °C			
		30°	40°	50°	60°
K_{obs} (min^{-1})	Producte (GFP)	$5.0e^{-4} \pm 2.1e^{-4}$	$2.2e^{-3} \pm 2.6e^{-4}$	$2.4e^{-3} \pm 3.0e^{-4}$	$3.5e^{-3} \pm 1.8e^{-4}$
	N-escissió	$4.7e^{-3} \pm 1.3e^{-3}$	$7.2e^{-3} \pm 3.4e^{-3}$	$2.7e^{-3} \pm 6.8e^{-4}$	$6.3e^{-3} \pm 2.2e^{-3}$
	C-escissió	$7.5e^{-4} \pm 2.7e^{-4}$	$2.0e^{-3} \pm 3.2e^{-4}$	$3.6e^{-3} \pm 6.3e^{-4}$	$8.3e^{-3} \pm 2.8e^{-3}$

*L'error indicat és l'error estàndard obtingut en l'ajust de la corba a una equació exponencial de primer ordre, $n = 2-4$.

Taula 4.4. Valors de $t_{1/2}$ en les diferents reaccions de tall i unió in vitro de *NpuDnaE C+1S* en cis*

		Temperatura °C			
		30°	40°	50°	60°
$t_{1/2}$ (min)	Producte (GFP)	1377.8 ± 412.4	314.5 ± 32.4	295.3 ± 33.6	198.4 ± 9.5
	N-Escissió	148.0 ± 32.5	96.5 ± 30.9	261.0 ± 53.2	109.8 ± 28.6
	C-Escissió	925.9 ± 243.0	346.2 ± 47.6	190.9 ± 28.1	83.9 ± 21.1

*El temps de vida mitja es va calcular a partir del valor de la constant de primer ordre ($\ln 2/K_{obs}$). L'error indicat és l'error estàndard, $n = 2-4$.

Un cop obtinguts els valors de K_{obs} i $t_{1/2}$ es va poder fer finalment una avaluació global del procés. En primer lloc cal remarcar que a 60°C és on es trobava la major quantitat de producte de tall i unió format a les 24h i a la vegada és la temperatura a la qual s'observava un valor de $t_{1/2}$ més baix, la qual cosa indicava una velocitat del procés més elevada. El valor de $t_{1/2}$ a aquesta temperatura era d'uns 200 minuts i era 1.5 vegades inferior del $t_{1/2}$ a 50 i 40°C i 6.5 vegades inferior del $t_{1/2}$ a 30°C. En el cas de la C-escissió s'observava també una disminució del valor de $t_{1/2}$ a mesura que s'incrementava la temperatura obtenint un valor a 60°C 11 vegades inferior a l'obtingut a 30°C. Per tant la formació de C-escissió va lligada també a l'increment de temperatura. Malgrat que la velocitat de formació de C-escissió depengui també de la temperatura, la formació total de C-escissió es mantenia constant a partir de 50°C tal i com es mostra en el gràfic 4.15. A més, la quantitat total de C-escissió a les 24h era 4 vegades inferior respecte al producte de tall i unió. En el cas de la N-escissió els valors obtinguts de $t_{1/2}$ indiquen que la formació del producte no depenia de la temperatura ja que no hi ha correlació. A més, la quantitat de N-escissió formada a les 24h era bastant més important que la

quantitat de C-escissió formada. Malgrat això, sembla que la formació de N-escissió arribava al seu màxim als 50°C tal i com s'observa també en el cas del C-escissió i continua sent menor a la quantitat de producte de tall i unió format. Tenint en compte aquests resultats, podríem dir que a 60°C és on s'obtenien les millors condicions en quant a la quantitat de producte de tall i unió respecte a la quantitat de N- i C-escissió format, ja que a partir de 50°C aquestes reaccions no desitjades deixaven d'incrementar, mentre que, la quantitat de producte de tall i unió sí que incrementa a 60°C respecte dels 50°C.

S'ha descrit que el procés de tall i unió no requereix d'agents reductors. No obstant, la seqüència de *Npu DnaE* presenta 3 Cys que podrien formar ponts disulfur i aquests podrien impedir les reaccions del mecanisme de tall i unió (Callahan et al., 2011; Chen et al., 2012; Shi and Muir, 2005). Ramirez i col·laboradors van determinar que l'activitat de *Npu DnaE* requeria de la presència d'agents reductors. Purificant el domini I_N juntament amb el domini d'unió a quitina (ChBD) com a exteïna i el domini I_C amb la GFP com a exteïna i realitzant les proves de tall i unió a 22°C van observar que el producte de tall i unió apareixia en les mostres que contenien DTT (2mM) i que la inteïna era inactiva sense agent reductor (Ramirez et al., 2013). En el nostre cas es va voler determinar si a partir de la mutació C+1S s'observaven diferències en l'activitat de la inteïna en presència d'agent reductor i sense. Es van fer proves a diferents temperatures des de 20 a 60°C i no es van observar diferències entre les mostres amb i sense DTT. A la figura 4.16 es mostren els gels SDS-PAGE de 50 i 60°C que són les temperatures a les quals l'activitat de la inteïna es fa més evident.

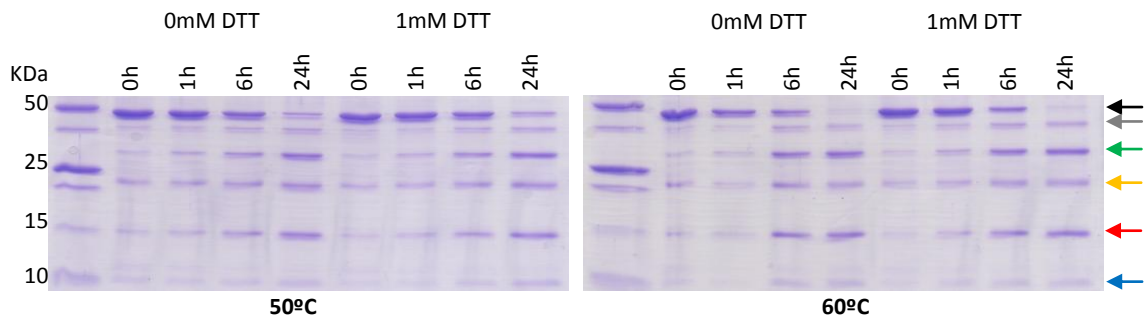


Fig. 4.16. Gels SDS-PAGE de les proves de tall i unió *In vitro* a 50 i 60°C a diferents temps de GFP_N-Npu DnaE-GFP_C amb la mutació C+1S amb i sense DTT. La fletxa negra indica el precursor (47205 Da), la fletxa gris indica GFP_N-Npu DnaE procedent de la C-escissió (38182 Da), la fletxa verda indica la GFP producte del tall i unió (31408 Da), la fletxa groga indica Npu DnaE-GFP_C o GFP_N procedent de la N-escissió (24856 -22367 Da), la fletxa vermella indica la Npu DnaE i la fletxa blava indica la GFP_C procedent de la C-escissió.

A partir dels resultats observats en els gels, no s'observaven diferències a simple vista en l'eficiència de tall i unió en presència o no de DTT. Aquests resultats indiquen doncs que el procés de tall i unió amb la mutació C+1S a GFP_N-Npu DnaE-GFP_C no requereix d'agents reductors. Per tal de verificar això, es va quantificar el producte total format a les 24h per cada una de les temperatures assajades en presència i absència de DTT. Els resultats es mostren a la figura 4.17.

4. Resultats i Discussió

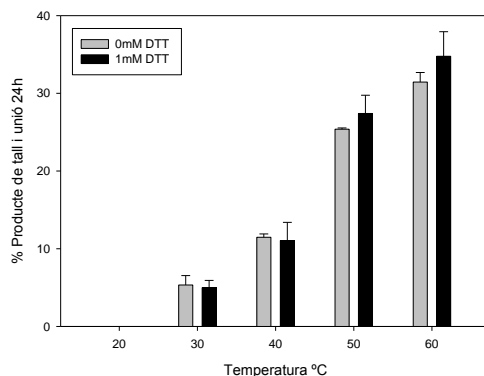


Fig. 4.17. Gràfic on es mostra per cada temperatura el percentatge a les 24h del producte de tall i unió de les proves realitzades a 0 i 1mM de DTT.

En el gràfic s'observa una formació del producte de tall i unió per cada temperatura bastant semblant amb o sense presència de DTT. Només en els casos de 50 i 60°C s'observava un lleuger increment del producte total de tall i unió format en presència de DTT respecte a les mostres que no en tenien. Per tal de veure si les petites diferències amb la formació del producte de tall i unió a les 24h amb DTT o sense eren significatives, es va realitzar el test estadístic paramètric t-student per dos mostres suposant variàncies desiguals amb un valor crític de 0.05. Abans de realitzar el test t-student es va comprovar que hi havia normalitat (test de Anderson-Darling) i homogeneïtat de variàncies (test de Levene). Si el valor-p obtingut en el test de t-student era superior a 0.05 significava que no hi havia diferències entre la formació de producte de tall i unió total amb presència de DTT o sense. A la taula 4.5 es mostren els valors-p obtinguts a cada una de les temperatures assajades.

Taula 4.5. Taula on es mostren els valors-p obtinguts en el test estadístic t-student per veure si hi havia diferències entre la quantitat total de producte de tall i unió format amb presència o sense DTT

Temperatura	Valor-p
30°C	0.29
40°C	0.37
50°C	0.46
60°C	0.18

Segons els resultats del test estadístic no hi havia diferències significatives en la formació de producte de tall i unió total amb presència de DTT o sense en el procés de tall i unió en cis amb la *Npu DnaE* com a inteïna i utilitzant la GFP com a exteïna amb la mutació C+1S.

4.2.2 Estudi del procés de tall i unió en trans de la *Npu DnaE* C+1S utilitzant GB1 com a exteïna

4.2.2.1 Construcció, expressió i purificació de GB1-I_N I_C-GB1 C+1S

Per a l'obtenció de les construccions GB1-I_N i I_C-GB1 C+1S es van utilitzar els vectors pskDuet01 i pskBAD2. La construcció GB1-I_N ja es trobava en el vector pskDuet01 de la manera com interessava per a la realització d'aquest estudi. En el cas de I_C-GB1 que es trobava en el vector pskBAD2, es va

haver de realitzar un QuikChange per tal d'introduir la mutació C+1S a la seqüència de la GB1. Els oligonucleòtids utilitzats van ser els que es detallen a continuació:

5' GGCTTCATAGCTTCTAATAGTTTCAATGGTACCC 3'
3' CCGAAGTATCGAAGATTATCAAAGTTACCATGGG 5'

Els clons que es van obtenir a partir del QuikChange es van seqüenciar i seleccionar els que contenien la mutació buscada. El vector pskBAD2 amb la mutació C+1S es va emprar per transformar cèl·lules BL21 (DE3) com també es va fer amb el vector pskDuet01 per tal de realitzar la producció de GB1-I_N i I_C-GB1 C+1S. Es va realitzar també la producció I_C-GB1 sense la mutació C+1S per tal de realitzar les proves de tall i unió entre GB1-I_N I_C-GB1 i GB1-I_N I_C-GB1 C+1S.

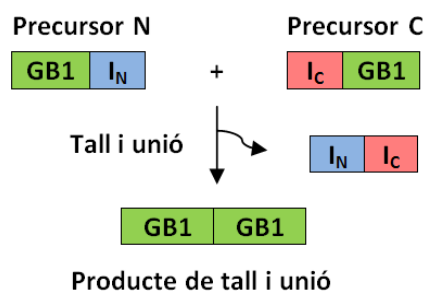
Acabada la producció, es van recollir les cèl·lules i es va separar la fracció soluble de la insoluble i es va observar que tots els precursors es trobaven a la fracció soluble. D'aquesta fracció es va realitzar una cromatografia d'afinitat Ni-NTA seguint el protocol estàndard (apartat 3.4.4 de material i mètodes) per tal de purificar els precursors i poder realitzar les diferents proves de tall i unió.

4.2.2.2 Estudi del procés de tall i unió de GB1-I_N I_C-GB1 C+1S

Per tal de realitzar l'estudi del procés de tall i unió de GB1-I_N I_C-GB1 C+1S i GB1-I_N I_C-GB1 en funció de la temperatura, els precursors purificats es van incubar a 20, 30, 40, 50 i 60°C durant diferents temps fins a un màxim de 24h. No es van realitzar tampoc les proves a més de 60°C ja que la proteïna es trobava a temperatures superiors a la seva T_m i per tant es trobaria desnaturalitzada. Una vegada els precursors van ser incubats a diferents temps i diferents temperatures, es van recollir i analitzar les diferents mostres electroforèticament per SDS-PAGE. A la figura 4.18 es mostren els gels SDS-PAGE corresponents a 20, 30, 40, 50 i 60°C, en aquest cas el procés de tall i unió es feia ja evident als 20°C.

Per verificar que el procés de tall i unió a aquestes temperatures era actiu i que la banda corresponent al fragment de tall i unió era la correcta, es va realitzar un western blot emprant un anticòs anti-Histag ja que el domini GB1 presenta una cua d'histidines. Els resultats del western blot es mostren a la figura 4.19.

a)



b)

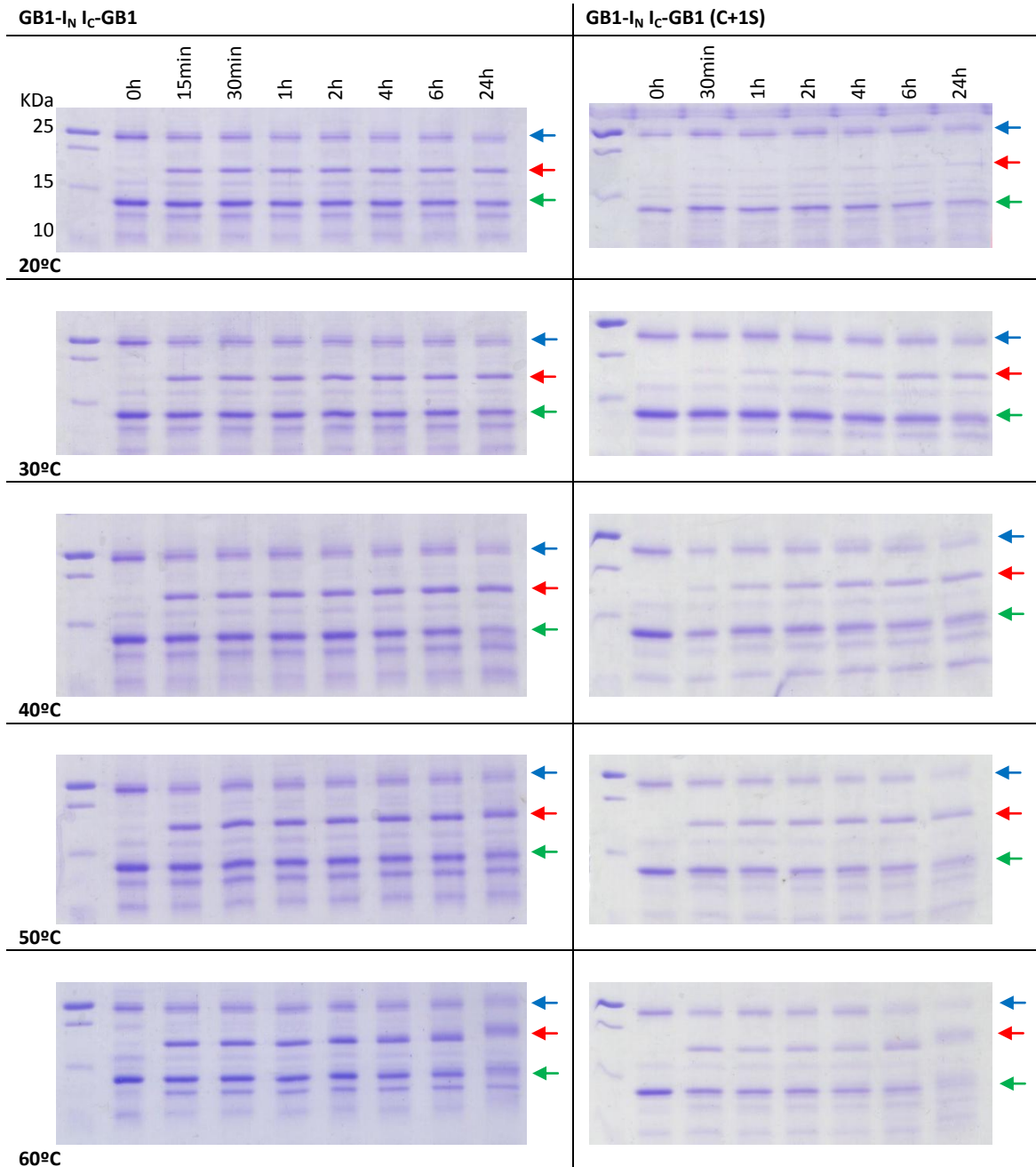


Fig. 4.18. (a) Esquema i (b) gels SDS-PAGE de les proves de tall i unió *In vitro* a 20, 30, 40, 50 i 60°C a diferents temps de GB1-I_N I_C-GB1 i GB1-I_N I_C-GB1 (C+1S). La fletxa blava indica el precursor N (20385 Da), la fletxa verda indica el precursor C (12258 Da) i la fletxa vermella indica el producte de tall i unió GB1-GB1 (12258 Da).

Els resultats dels western blots confirmaven també, que el procés de tall i unió era actiu i que aquest era més efectiu com més elevada era la temperatura fins arribar a 50°C. Les mostres que van estar a 60°C durant 24h presentaven certa degradació. Aquest fet es degut a que els precursors I_N i I_C es trobaven a una temperatura superior a la seva T_m que és de 52.6°C. Per aquest motiu i segons els resultats obtinguts a partir de les quantificacions dels gels amb el programa Gel Analyzer, a 60°C no hi havia un augment del producte de tall i unió format respecte a 50°C. A la figura 4.20 el gràfic (a) mostra l'evolució en percentatge de la formació de producte de tall i unió al llarg de les hores per

4. Resultats i Discussió

cada una de les temperatures assajades i en el gràfic (b) es mostra la formació de producte de tall i unió total a les 24h per GB1-I_N I_C-GB1 (wt) i GB1-I_N I_C-GB1 C+1S per cada una de les temperatures assajades.

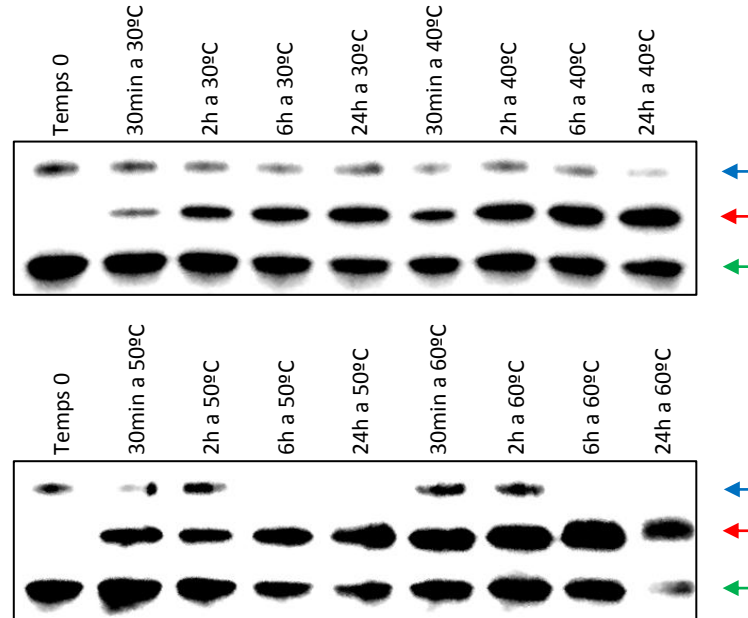


Fig. 4.19. Western blots (Anti-His tag) de les proves de tall i unió *in vitro* a 30, 40, 50 i 60°C a 30min, 2h, 6h i 24h de GB1-I_N I_C-GB1 (C+1S). La fletxa blava indica el precursor N, la fletxa verda indica el precursor C i la fletxa vermella indica el producte de tall i unió GB1-GB1.

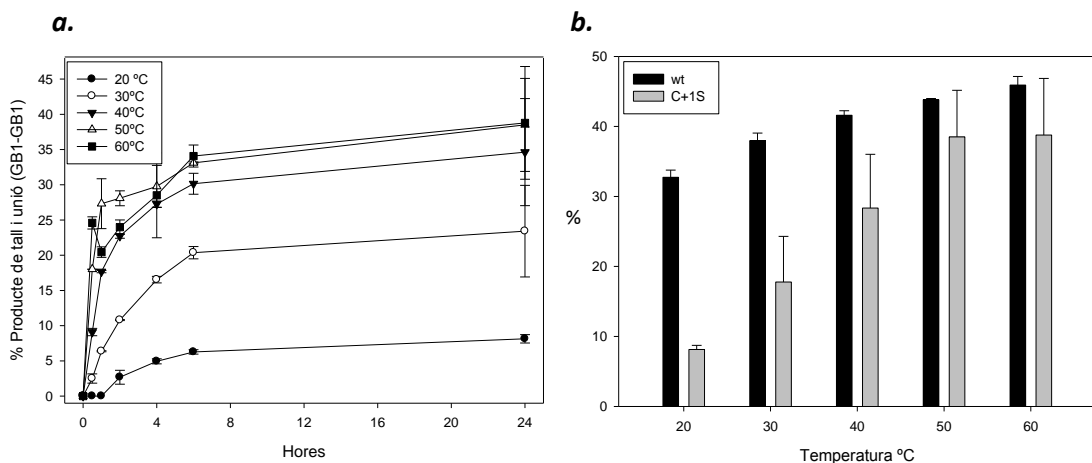


Fig. 4.20. (a) Formació del producte de tall i unió *in vitro* al llarg del temps a diferents temperatures de GB1-I_N I_C-GB1 C+1S. (b) Formació de producte de tall i unió total *in vitro* a les 24h de GB1-I_N I_C-GB1 C+1S i GB1-I_N I_C-GB1 (wt) per a cada una de les temperatures assajades.

A partir de les quantificacions de cada una de les fraccions resultants del procés de tall i unió es va poder realitzar el càlcul de K_{obs} i $t_{1/2}$ del producte de tall i unió. A partir dels càlculs d'aquests paràmetres i del rendiment obtingut (producte de tall i unió total a les 24h) es podia veure a quina de les temperatures el procés de tall i unió era més eficient. Els paràmetres K_{obs} i $t_{1/2}$ es van calcular ajustant la formació de producte de tall i unió en funció del temps a una corba exponencial de primer ordre. Els valors obtinguts es mostren a la taula 4.6.

4. Resultats i Discussió

Taula 4.6. Valors de K_{obs} (min^{-1}) i valors de vida mitja $t_{1/2}$ (min) observats en les diferents reaccions de tall i unió in vitro de NpuDnaE C+1S en trans*

		Temperatura °C									
		20		30		40		50		60	
		wt	C+1S	wt	C+1S	wt	C+1S	wt	C+1S	wt	C+1S
Producte	K_{obs}	0.10 ± 0.02	0.0026 ± 0.0007	0.09 ± 0.02	0.0064 ± 0.0008	0.10 ± 0.03	0.0106 ± 0.001	0.12 ± 0.03	0.0242 ± 0.004	0.14 ± 0.05	0.0315 ± 0.005
	$t_{1/2}$	6.86 ± 1.24	266.6 ± 39.98	7.28 ± 1.76	108.3 ± 8.51	6.86 ± 1.69	65.4 ± 5.05	5.83 ± 1.33	28.6 ± 3.23	5.13 ± 1.4	22.0 ± 2.41

*L'error indicat és l'error estàndard obtingut en l'ajust de la corba a una equació exponencial de primer ordre i la vida mitja es va calcular a partir del valor de la constant de primer ordre ($\ln 2/K_{obs}$). L'error indicat és l'error estàndard, $n = 2-4$.

De la mateixa manera que s'ha descrit per la inteïna Npu DnaE en cis amb la mutació C+1S, es va realitzar el test estadístic paramètric t-student per tal de veure si les diferències entre el producte total de tall i unió format a les 24h era diferent entre GB1-I_N I_C-GB1 C+1S i GB1-I_N I_C-GB1 (wt) per cada una de les temperatures. El test t-student utilitzat també va ser per dos mostres suposant variàncies desiguals amb un valor crític de 0.05. Si el valor-p obtingut era superior a 0.05 significava que no hi havia diferències entre la formació de producte de tall i unió total format entre GB1-I_N I_C-GB1 C+1S i GB1-I_N I_C-GB1 (wt). Abans de realitzar el test t-student també es va es va comprovar que hi havia normalitat (test de Anderson-Darling) i homogeneïtat de variàncies (test de Levene). A la taula 4.7 es mostra el valor-p obtingut a cada una de les temperatures assajades.

Taula 4.7. Taula on es mostren els valors-p obtinguts en el test estadístic t-student per veure si hi havia diferències entre la quantitat total de producte de tall i unió format entre GB1-I_N I_C-GB1 C+1S i GB1-I_N I_C-GB1 (wt)

Temperatura	Valor-p
20°C	8.62 E-06
30°C	0.02
40°C	0.04
50°C	0.15
60°C	0.52

Segons els resultats del test estadístic t-student, a 50 i 60°C no hi havia diferències significatives en la quantitat de producte de tall i unió format tant en el cas del salvatge com el mutat. Mentre que de 20 a 40°C sí que hi havia diferències significatives en la quantitat total del producte de tall i unió format entre GB1-I_N I_C-GB1 C+1S i GB1-I_N I_C-GB1 (wt).

En el cas de les proves de tall i unió utilitzant el domini GB1 com a exteïna no es va observar la formació de bandes pertanyents a la N- o C-escissió com les que s'havien observat en el cas de la GFP. Aquest fet es deu probablement a què l'eficiència del procés de tall i unió ve determinat en gran part per la exteïna tal i com s'ha demostrat en diferents estudis (Martin et al., 2001). Malgrat les diferències degudes a l'exteïna, amb la mutació C+1S s'obté a 50°C un producte de tall i unió del 30% amb la GFP i del 40% amb la GB1 i a 60°C és del 40% tant en el cas de la GFP com del domini

GB1 emprats com a exteïnes. Per tant, s'observen uns resultats bastant semblants en l'eficiència del producte total de tall i unió format i diferències significatives en el cas de productes secundaris.

Amb la mutació C+1S, *Npu DnaE* és activa tant quan realitza el tall i unió en cis com en trans malgrat que aquesta mutació incrementi molt el $t_{1/2}$ de formació del producte a temperatures baixes, i que, s'aconsegueix disminuir aquest $t_{1/2}$ mitjançant un increment de temperatura. Per altra banda, es va pensar que es podria aconseguir una inteïna activa mutant la Cys 1 per una Ser a temperatura fisiològica o inferior, atès que, algunes inteïnes són actives tenint una Ser a la posició 1 (Perler, 2002). Per aquest motiu, es va crear el mutant GB1-I_N C1S i el doble mutant C1S i C+1S per tal de veure si amb aquestes mutacions s'aconseguia una inteïna activa i, en cas que no ho fos, veure si mitjançant un increment de temperatura es podria revertir el bloqueig. Aquest treball es troba descrit en el següent apartat.

4.2.3 Estudi del procés de tall i unió en trans de la *Npu DnaE* utilitzant GB1 com a exteïna amb la doble mutació C1S i C+1S

4.2.3.1 Construcció, expressió i purificació de GB1-I_N I_C-GB1 C1S

Per a l'obtenció de les construccions GB1-I_N i I_C-GB1 C+1S es van utilitzar els vectors pskDuet01 i pskBAD2. La construcció GB1-I_N es trobava en el vector pskDuet01 i es va haver de realitzar un QuikChange per tal d'introduir la mutació C1S a la seqüència de la I_N. En el cas de I_C-GB1, es trobava en el vector pskBAD2 tal com interessava per a la realització d'aquest treball. Els oligonucleòtids utilitzats per introduir les mutacions van ser els que es detallen a continuació:

```
5' CCGTAACGGAAGGATCCTCTTTAAGCTATGAAACGG 3'  
3' GGCATTGCCTTCCTAGGAGAAATTCGATACTTTGCC 5'
```

Els clons que es van obtenir després del QuikChange es van seqüenciar i es van seleccionar els que contenien la seqüència correcta. El vector pskDuet01 amb la mutació C1S es va fer servir per transformar cèl·lules BL21 (DE3) així com es va fer amb el vector pskBAD2 per tal de realitzar la producció de GB1-I_N C1S i I_C-GB1 C+1S. Es va realitzar també la producció I_C-GB1 sense la mutació C+1S per tal de realitzar les proves de tall i unió entre GB1-I_N C1S I_C-GB1 i el doble mutant GB1-I_N C1S I_C-GB1 C+1S.

Acabada la producció, es van recollir les cèl·lules i es va separar la fracció soluble de la insoluble i es va observar que tots els precursors es trobaven a la fracció soluble. Aquesta fracció es va sotmetre a una cromatografia d'afinitat Ni-NTA seguint el protocol estàndard (apartat 3.4.4 de material i mètodes) per tal de purificar els precursors i poder realitzar les diferents proves de tall i unió.

4.2.3.2 Estudi del procés de tall i unió de GB1-I_N I_C-GB1 C1S i GB1-I_N C1S I_C-GB1 C+1S

Per l'estudi del procés de tall i unió de GB1-I_N C1S - I_C-GB1 i del doble mutant GB1-I_N C1S - I_C-GB1 C+1S en funció de la temperatura, els precursors purificats es van incubar a 20, 30, 40, 50 i 60 °C durant diferents temps fins a un màxim de 24h. No es van realitzar les proves a més de 60°C ja que la inteïna es trobaria a temperatures superiors a la seva T_m i per tant estaria desnaturalitzada. Una vegada els precursors van ser incubats a diferents temps i diferents temperatures, es van recollir i analitzar electroforèticament cadascuna de les mostres per SDS-PAGE. En els gels SDS-PAGE obtinguts no es va observar en cap cas la banda pertanyent al producte de tall i unió (GB1-GB1) tant per la mutació C1S com per la doble mutació C1S-C+1S.

Mitjançant un western blot es va verificar que el procés de tall i unió en aquests casos es trobava bloquejat. La tècnica es va realitzar utilitzant l'anticòs anti-Histag ja que el domini GB1 presentava una cua d'histidines. Els resultats del western blot es mostren a la figura 4.21.

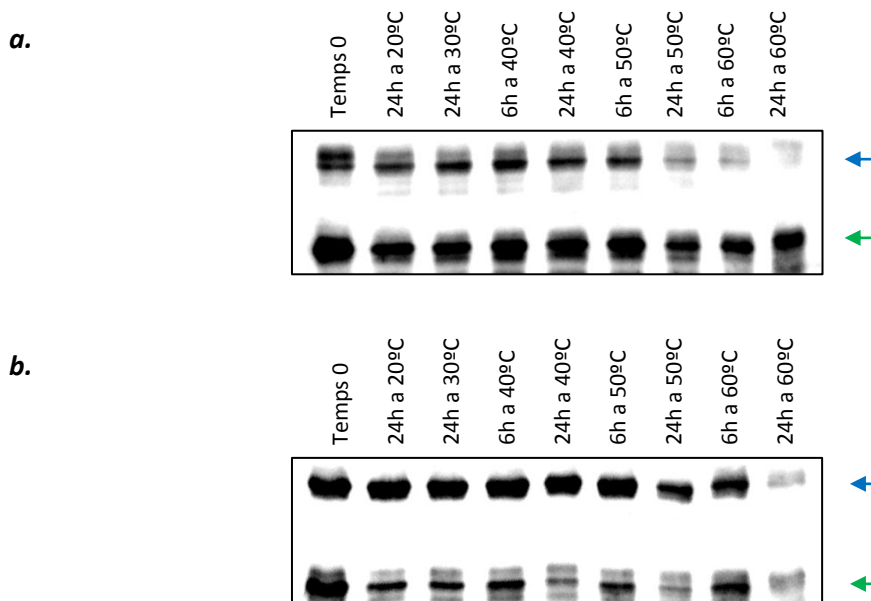


Fig. 4.21. Western blots (Anti-His tag) de les proves de tall i unió *In vitro* a 20, 30, 40, 50 i 60°C a, 6h i 24h. El western a correspon a les proves de tall i unió de GB1-I_N C1S - I_C-GB1 i el western b correspon a les proves de tall i unió de GB1-I_N C1S - I_C-GB1 C+1S. La fletxa blava indica el precursor N i la fletxa verda el precursor C.

En els resultats del western tampoc s'observava la banda que pertanyia al producte de tall i unió (GB1-GB1). Per tant, confirmaven el que s'havia observat en els gels SDS-PAGE, la mutació C1S inhibia el procés de tall i unió en la inteïna *NpuDnaE* quan aquesta es trobava en trans i amb la GB1 com a exteïna malgrat hi hagués un increment de temperatura. Amb la doble mutació C1S i C+1S el procés es trobava també bloquejat tal i com era d'esperar. Aquest fet demostra que la Cys 1 té un paper essencial en l'activitat de la inteïna i és indispensable per l'atac nucleofílic que es du a terme

entre els residus -1 i 1 per formar l'intermediari éster, i que, malgrat s'augmenti la temperatura, aquest atac no es du a terme quan aquesta Cys1 es substitueix per una Ser. No obstant, una mutació de la Cys +1 per una Ser que és essencial per dur a terme el procés de transesterificació, és capaç de disminuir l'activitat d'aquesta, però aquesta disminució es pot compensar augmentant la temperatura. Per tant, en aquest cas, la hipòtesi de partida la qual *Npu DnaE* podria ser activa amb la mutació C1S no es compleix.

4.3 ESTUDI ESTRUCTURAL MITJANÇANT RMN DEL PROCÉS ASSOCIATIU DE LA INTEÏNA *Npu DnaE*

El tall i unió de proteïnes en trans és un procés auto-catalític postraducciona l indispensable en els organismes unicel·lulars per tal d'aconseguir l'estructura i funció correcte de les proteïnes codificades en gens partits. Malgrat que les bases químiques d'aquest mecanisme es troben ja ben conegudes, no és gaire clar encara com aquestes inteïnes en trans s'associen, es complementen, catalitzen i regulen les diferents etapes del procés.

Muir i col·laboradors proposen, tal i com s'ha explicat a la introducció, un model de captura i col·lapse en el qual la primera regió de la subunitat I_N (residus 1-50) de la inteïna *Npu DnaE* en trans es trobaria parcialment plegada mentre que l'altre regió (residus 51-102) es trobaria completament desplegada. Per altra banda, la subunitat I_C es trobaria també desplegada. Mitjançant complementarietat electrostàtica que es dona entre la regió de la subunitat I_N desplegada amb la I_C , es produiria una "captura" dels dos fragments durant la qual el complex es col·lapsa en un estat ordenat que quedaria estabilitzat per les interaccions hidrofòbiques que s'establirien entre la subunitat I_C i la regió estructurada de la subunitat I_N . A partir d'aquest col·lapse i estabilització de la inteïna, s'assoleix la conformació nativa i aquesta adquiriria la capacitat de dur a terme el procés de tall i unió (Shah et al., 2013b) (figura 1.19 de la introducció).

De manera paral·lela i per tal d'estudiar les interaccions de I_N amb I_C , en el nostre grup s'havia dissenyat una estratègia diferent per tal de seguir el procés d'associació. En l'estructura de la *Npu DnaE* (Codi PDB 2KEQ) s'observa que la I_C es troba formada per 4 cadenes β ($\beta_{11} - \beta_{14}$) que participen en la formació de 3 fulles β diferents (figura 4.22). Per tant, es va decidir avaluar quins canvis conformacionals produeixen aquestes 4 cadenes β sobre el plegament global de *Npu DnaE*, partint del coneixament previ que la subunitat I_N sola presenta una estructura de glòbul fos tal i com s'explicarà en aquest apartat i que, a la vegada, s'ha observat en les anàlisis d'estabilitat tèrmica que hi ha una sinèrgia de plegament mútua entre les subunitats I_N i I_C , la qual també es troba descrita en la trans inteïna *Ssp DnaE*. Per realitzar aquest estudi s'ha posat a punt un protocol d'expressió i purificació del complex I_N-I_C en una única cadena polipeptídica amb diferents longituds de la

subunitat I_C (corresponents a les diferents cadenes β fusionades a la I_N anomenades I_NI_C_sc). Per tal d'avaluar quins canvis conformationals produeixen aquestes 4 cadenes β sobre el plegament global de *Npu DnaE*, s'han utilitzat mètodes espectroscòpics com el dicromisme circular (CD) i la ressonància magnètica nuclear (RMN). A més, s'ha realitzat una anàlisi bioinformàtica prèvia per tal de tenir una primera aproximació del plegament que presenten cada una de les variants I_NI_C_sc estudiades en aquest treball.

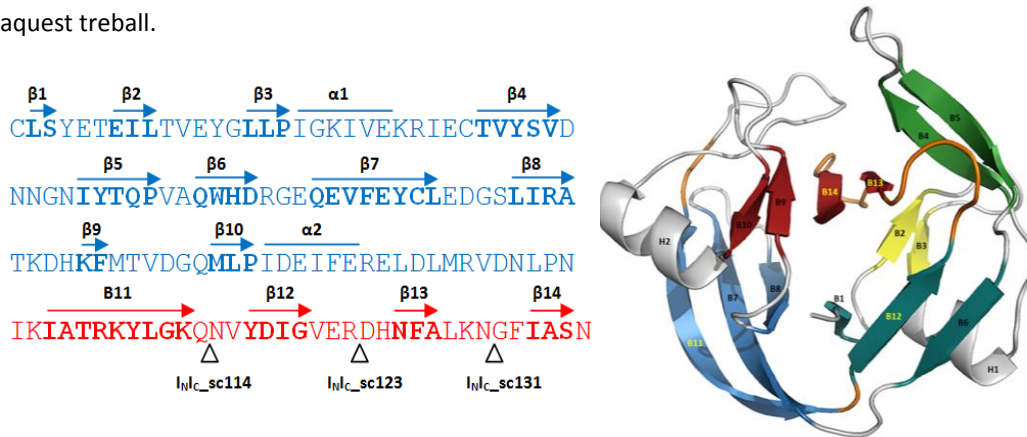


Fig. 4.22. Elements d'estructura secundària de *Npu DnaE*. En blau s'indica la regió pertanyent a la subunitat I_N i en vermell la regió pertanyent a la subunitat I_C. S'indica també la situació dels residus que delimiten les proteïnes de fusió I_NI_C_sc. A la dreta hi ha l'estructura tridimensional de la *Npu DnaE* on s'enumeren les diferents cadenes β i hèlix α.

4.3.1 Construcció, expressió i purificació de les variants I_NI_C_sc

4.3.1.1 Construcció de les variants I_NI_C_sc

Per a l'obtenció de les diferents variants de I_NI_C_sc es va partir del gen de la *Npu DnaE* formant una única seqüència que es trobava en el vector pRSET-GFP_N-*Npu DnaE*-GFP_C. El gen es va amplificar mitjançant PCR utilitzant un únic encebador per l'amplificació des de l'extrem N-terminal i diferents encebadors de l'extrem C-terminal. A cada un dels encebadors es va introduir un codó de finalització de la traducció en una posició que permetés l'obtenció de les diferents longituds desitjades de la subunitat I_C. Es van realitzar tres amplificacions diferents tal i com s'observa en les figures 4.22 i 4.23 per a l'obtenció de I_NI_C_sc114 (β 11), I_NI_C_sc123 (β 11 i β 12) i I_NI_C_sc131 (β 11, β 12 i β 13). Els números 114, 123 i 131 fan referència al nombre total de residus de cada una de les construccions. Els encebadors es van dissenyar de tal manera que tinguessin a l'extrem 5' una diana *NcoI* i a l'extrem 3' una diana *HindIII*. A continuació es detallen aquests encebadors:

Amplificació extrem N-terminal:

5' GGTCTATATCACCGCCGCATGGCATGCTTAAGC 3'

Amplificació extrem C-terminal:

I_NI_C_sc114: 5' CAACTCCAATGTCATAAGCTTATTGTTTGCCATA 3'

$I_{N|C_sc123}$: 5' CCATTTTTGAGTGCAAGCTTATGTTAGCGCTCAACTCC 3'

$I_{N|C_sc131}$: 5' GAAACAATTGGAAGCTAAAGCTTAATTTTTGAGTGC 3'

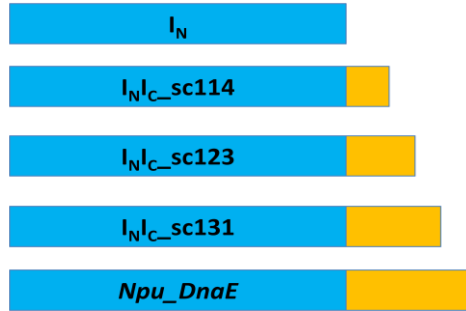


Fig. 4.23. Proteïnes de fusió $I_{N|C_sc}$ construïdes per a la realització d'aquest treball. En blau s'indica la regió de la subunitat I_N i en taronja les diferents longituds de la regió pertanyent a la subunitat I_C .

Una vegada realitzades les amplificacions es van digerir amb els enzims de restricció *NcoI* i *HindIII* i es va dur a terme una lligació amb el vector pBADm41 digerit prèviament amb els mateixos enzims. Dels clons obtinguts, es van seqüenciar i seleccionar els que tenien la seqüència buscada. D'aquesta manera es va obtenir un vector on cada una de les variants $I_{N|C_sc}$ es trobava unida a la MBP per tal de facilitar la seva solubilització. A la vegada, aquestes dues proteïnes es troben flanquejant la diana de la proteasa TEV (ENLYFQ↓G) (Carrington and Dougherty, 1988). A davant de la proteïna de solubilització hi havia també una seqüència de 6 His per tal de facilitar el procés de purificació. Un esquema de les construccions realitzades es troba detallada a la figura 4.24.



Fig. 4.24. Esquema de la construcció realitzada en el vector pBADm41 per tal de facilitar el procés de purificació de les variants $I_{N|C_sc}$. En vermell s'indica una seqüència de 6 His, en taronja la seqüència pertanyent a la proteïna de solubilització MBP, en verd la seqüència reconeguda per a la proteasa TEV i en blau la seqüència corresponent a cada una de les variants $I_{N|C_sc}$.

4.3.1.2 Expressió i purificació de les variants $I_{N|C_sc}$

Abans d'iniciar el procés de producció i purificació de les variants $I_{N|C_sc}$, es va analitzar el grau d'expressió de diferents clons. Per això es van analitzar electroforèticament cultius induïts de diferents clons i es van seleccionar aquells que presentaven millor expressió. Amb els clons seleccionats es van obtenir els rendiments indicats a la taula 4.8.

Es va fer un seguiment del procés de purificació per tal d'analitzar en quines fraccions es trobaven les diferents variants de $I_{N|C_sc}$. El procés de purificació de les variants $I_{N|C_sc}$ va ser molt similar al descrit per a la purificació de la subunitat I_N (apartat 3.4.4 de material i mètodes). Un cop passades les cèl·lules per la French Press, es va voler determinar si les variants es trobaven en la fracció soluble o insoluble. Per això, es va centrifugar la mostra provinent de la French Press i es va carregar

mostra provinent del sobrenedant (fracció soluble) i mostra provinent del sediment (fracció insoluble) en un gel SDS PAGE, el qual va demostrar que les variants es trobaven a la fracció soluble.

A partir d'aquí es va realitzar una cromatografia de Ni-NTA (apartat 3.4.4 de material i mètodes) per tal d'aïllar MBP- I_NI_C_sc de la resta de proteïnes de la mostra. Un cop aïllada parcialment MBP- I_NI_C_sc, es va digerir amb la proteasa TEV per trencar la unió entre la proteïna de solubilització i les variants I_NI_C_sc. Es va realitzar una segona cromatografia Ni-NTA per tal que I_NI_C_sc eluïssin en el través i la MBP quedés retinguda a la reïna. I_NI_C_sc van eluir en el través però no de manera pura i per tant es va seguir amb el procés de purificació mitjançant HPLC de bescanvi aniònic a partir del través de la segona cromatografia de Ni-NTA. La cromatografia es va realitzar amb la columna Hi-Trap Q HP (GE Healthcare, EUA) amb el tampó Tris-HCl 25mM, EDTA 1mM, DTT 1mM a pH 8 amb un gradient de NaCl. Es van recollir mostres de cadascun dels pics, per posteriorment, analitzar-los en un gel SDS-PAGE per tal de determinar en quin pic s'elueïen les variants I_NI_C_sc. Amb la cromatografia de bescanvi aniònic les variants I_NI_C_sc no es trobaven pures del tot ja que apareixien altres bandes de proteïna de pes molecular molt superior a I_NI_C_sc. Per tant, es va continuar amb el procés de purificació mitjançant una cromatografia de gel filtració Superdex 75 10/300 GL (GE Healthcare, EUA) per HPLC amb el tampó Tris-Acetats 50mM, EDTA 10mM, DTT 1mM a pH 8. Els cromatogrames obtinguts es van enregistrar a 280nm i es van recollir mostres de cadascun dels pics, per posteriorment, analitzar-los en un gel SDS-PAGE per tal de determinar en quin pic eluïa cada una de les variants I_NI_C_sc. A la figura 4.25 es mostren els cromatogrames obtinguts per a cada una de les variants de I_NI_C_sc.

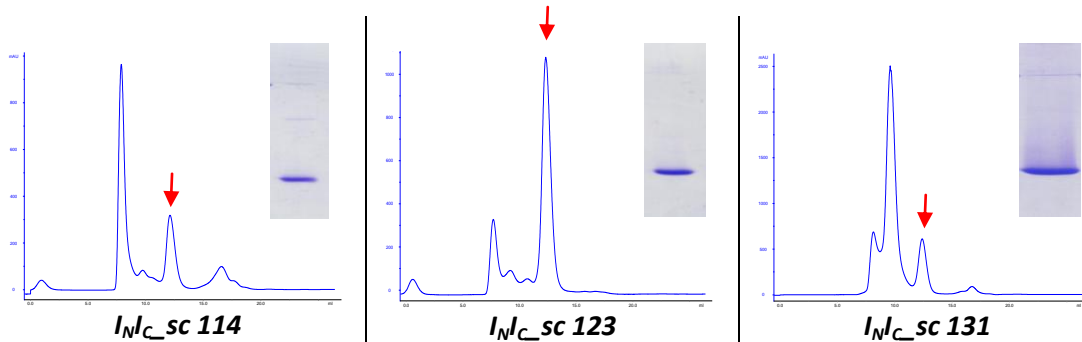


Fig. 4.25. Perfils cromatogràfics de l'elució de les diferents variants de I_NI_C_sc en la cromatografia de gel filtració. La línia blava fa referència a l'absorbància enregistrada a 280nm, i la fletxa vermella indica el pic el qual s'elueïa la proteïna d'interès. Cada un dels pics recollits es van analitzar per SDS-PAGE, el gel dels quals es mostra en cada un dels cromatogrames.

Mitjançant aquesta darrera etapa de purificació es van obtenir finalment les diferents variants de I_NI_C_sc pures, les quals es van liofilitzar i analitzar per espectrometria de masses MALDI-TOF. El valor de la massa molecular de les diferents variants de I_NI_C_sc obtingudes no difereix significativament de les esperades a partir del càlcul teòric (taula 4.8). Aquest fet confirmava que les diferents variants es trobaven sense la formació de ponts disulfur ni tampoc es van observar pics d'agregació o dimerització en l'espectre de MALDI-TOF obtingut.

Taula 4.8. Rendiments de purificació obtinguts per cada una de les variants de $I_N I_C$ sc i pesos moleculars obtinguts i teòrics per a cada una d'elles

Proteïna	Rendiment (mg/l de cultiu)	Mw teòric (Da)	Mw observat (Da)
$I_N I_C$ sc 114	6	13852	13854
$I_N I_C$ sc 123	10	14627	14636
$I_N I_C$ sc 131	15	15568	15571

4.3.2 Anàlisi bioinformàtica i predicció del grau de plegament

La bioinformàtica és un camp científic emergent i interdisciplinari que desenvolupa mètodes per a emmagatzemar, organitzar i analitzar dades biològiques. Un dels objectius principals d'aquesta disciplina és desenvolupar eines computacionals per tal de generar coneixements biològics útils.

Determinades eines permeten fer prediccions de si una proteïna es trobarà plegada o no. Aquestes eines es basen bàsicament en una anàlisi de la densitat de residus hidrofòbics que conté la proteïna i la càrrega neta mitjana. Proteïnes amb una elevada càrrega neta mitjana i baixa hidrofobicitat, tenen tendència a ser proteïnes desplegadas i proteïnes amb una baixa càrrega neta mitjana i alta hidrofobicitat, tenen tendència a ser proteïnes correctament plegades i estructurades (Uversky et al., 2000).

Per realitzar aquesta anàlisi bioinformàtica es va utilitzar el programa PONDR utilitzant el predictor VL-XT que, a partir de la seqüència de la proteïna, permet fer el càlcul de la hidrofobicitat de la proteïna i de la càrrega neta mitjana de la proteïna (Li et al., 1999; Romero et al., 1997; Romero et al., 2001). D'aquesta manera es poden construir gràfics de càrrega-hidrofobicitat com el que es mostra a la figura 4.26.

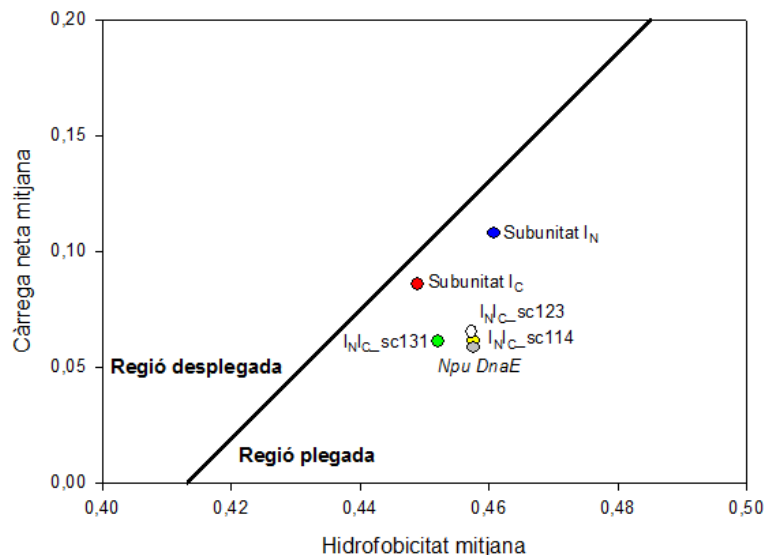


Fig. 4.26. Gràfic de càrrega-hidrofobicitat comparatiu de *Npu DnaE*, les diferents variants de $I_N I_C$ sc, la subunitat I_N i la subunitat I_C . La hidrofobicitat mitjana (H) es calcula a partir de l'escala normalitzada Kyte-Doolittle i la càrrega neta mitjana (R) es troba en valor absolut. La línia negra que separa la regió de proteïnes desplegadas de la regió de proteïnes plegades és definida com $R = 2,785 \cdot H - 1,151$ (Uversky et al., 2000).

El gràfic de càrrega-hidrofobicitat mostrava com totes les proteïnes estudiades es trobaven dins la regió plegada. Només les subunitats I_N i I_C es trobaven bastant al límit de la regió plegada. No obstant, cal remarcar que aquests gràfics de càrrega-hidrofobicitat presenten una fiabilitat del 75-83% amb una probabilitat de falsos negatius del 40% i de falsos positius del 22%. Per exemple, Zheng i col·laboradors van observar com la subunitat I_N de la *Ssp DnaE* es trobava dins la regió plegada i quan van fer els estudis de plegament mitjançant RMN, van observar com aquesta es trobava desplegada sense la presència de la subunitat I_C (Zheng et al., 2012). En un altre estudi de plegament de la *Npu DnaE* realitzat per Shah i col·laboradors observen també com les dues subunitats I_N i I_C es troben dins la regió plegada i en anàlisis de RMN aquestes es troben també desplegadas (Shah et al., 2013b).

Es va realitzar també l'anàlisi de funció de distribució acumulativa (CDF) de les tres variants $I_N I_C$ sc (Figura 4.27). En aquest tipus de anàlisi es fa una predicció del plegament de les proteïnes i es distribueixen en un gràfic el qual presenta un límit lineal que distingeix una proteïna ordenada d'una desordenada. En l'anàlisi CDF, les tres variants van aparèixer per sobre el límit lineal, la qual cosa indicava que es trobaven ordenades. Quan l'anàlisi CDF i el de càrrega hidrofobicitat coincideixen en que una proteïna es troba plegada o desplegada, cal pensar que és molt probable que la proteïna es trobi en l'estat predit (Uversky and Dunker, 2010). En ambdós anàlisis, les tres variants es trobaven a la regió plegada.

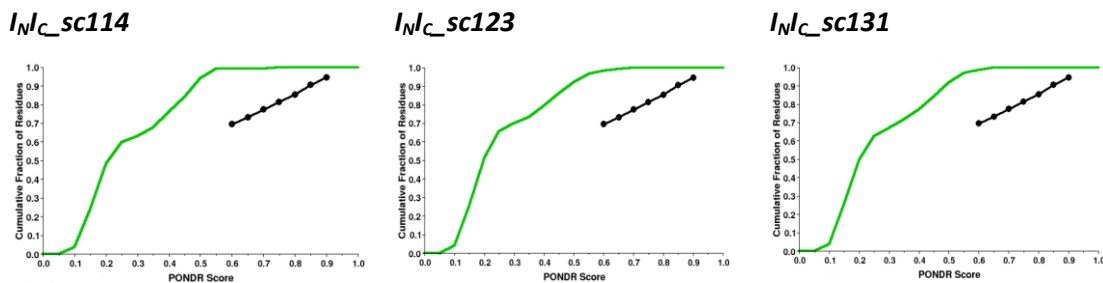


Fig. 4.27. Anàlisi CDF de les tres variants $I_N I_C$ sc. La línia negra representa el límit entre ordre (sobre) i desordre (sota). Les tres variants se situen en la regió ordenada (línia verda).

Per complementar aquests resultats es va realitzar també un altre tipus d'anàlisi que consisteix en la comparació dels segments ordenats i desordenats de cada una de les proteïnes. Aquest tipus d'anàlisi es basa en les freqüències relatives de cada aminoàcid i com aquests es troben en regions ordenades o desordenades en diferents bases de dades de proteïnes plegades i desplegadas.

Segons les bases de dades, s'ha descrit que les proteïnes intrínscament desordenades (IDP) tenen més residus promotors de desordre (A, R, G, Q, S, P, E i K) que residus promotors d'ordre (W, C, F, I, Y, V, L i N) (Dunker et al., 2001), la resta de residus es classifiquen com a neutres, atès que el seu impacte en el plegament global d'una proteïna és baix. Malgrat això, l'anàlisi de la composició

d'aminoàcids (figura 4.28) mostra que les diferents variants de $I_{N^I_C-sc}$ no pertanyen al grup de proteïnes IDP si es comparen amb la base de dades de proteïnes desplegadas Disprot 3.4 (Sickmeier et al., 2007). Aquests resultats suggereixen que les diferents variants serien, per tant, potencialment capaces d'adoptar estructura.

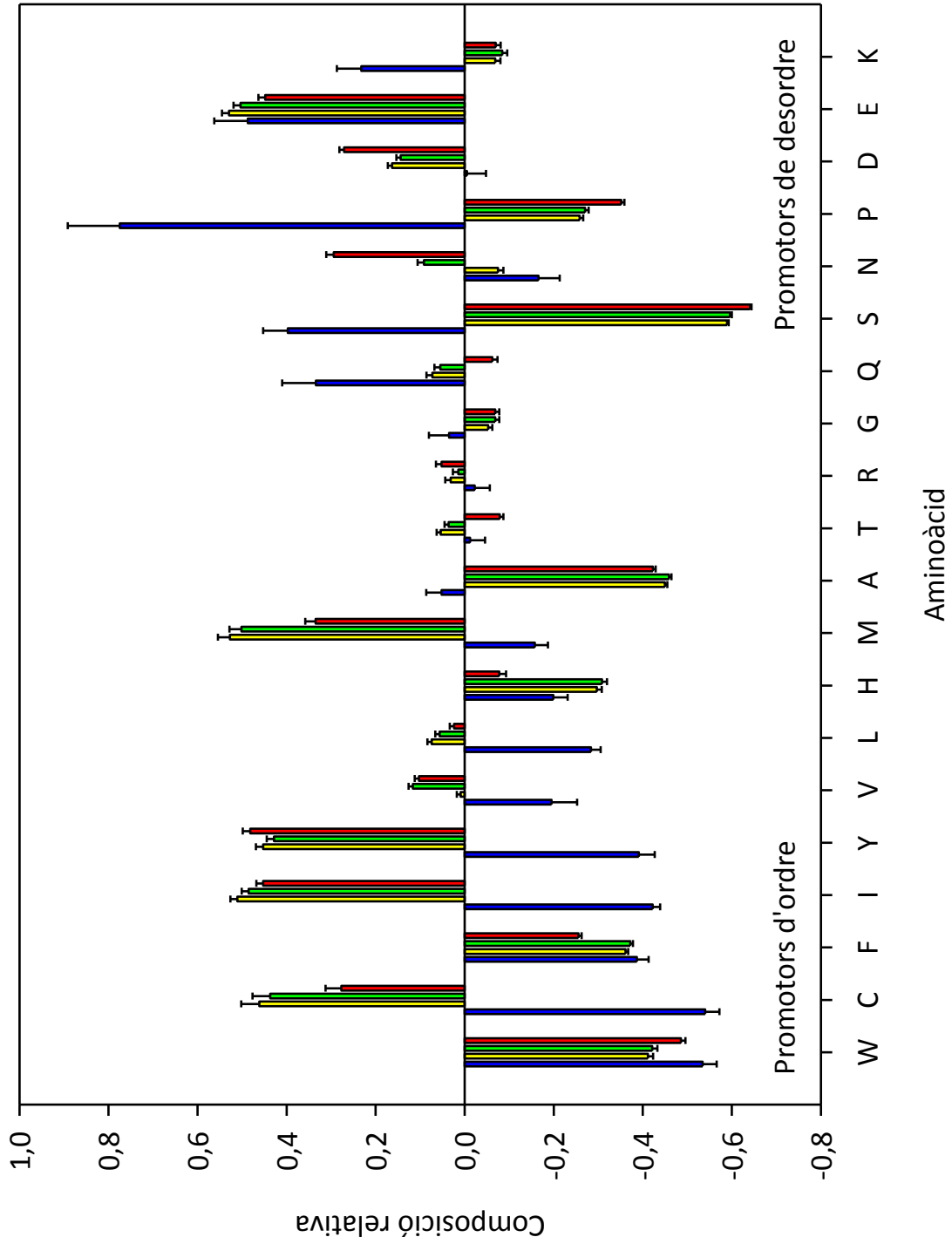


Fig. 4.28. Anàlisi on es compara la composició relativa d'aminoàcids de $I_{N^I_C-sc114}$ (barres grogues), $I_{N^I_C-sc123}$ (barres verdes), $I_{N^I_C-sc131}$ (barres vermelles) i la base de dades Disprot 3.4 amb PDB_Select_25 (barres blaves) utilitzant el programa informàtic Compositionn Profiler (Vacic et al., 2007). Els valors dels aminoàcids es troben en l'escala de flexibilitat segons Vihinen (Vihinen et al., 1994) i representen enriquiment (positiu) o empobriment (negatiu) relatiu a la base de dades PDB_select_25 . Totes les variants tenen menys composició relativa no significativa d'aminoàcids promotors de desordre que els que es troben a la base de dades PDB_select_25 i més composició relativa d'aminoàcids promotors d'ordre, utilitzant un valor significatiu de $p < 0.05$.

4.3.3 Anàlisi de l'estructura secundària per dicroisme circular

Una vegada havien estat purificades les diferents variants de $I_{N|C_sc}$ i analitzada prèviament la seva capacitat de formar estructura analitzant mitjançant gràfics de càrrega-hidrofobicitat, CDF, i la comparació dels segments ordenats i desordenats de cadascuna de les variants, el següent pas a seguir, per tal de veure si les diferents longituds de la subunitat I_C afectaven el plegament, era fer una anàlisi de l'estructura secundària. Per dur a terme aquesta anàlisi es va emprar l'espectroscòpia de dicroisme circular.

Es va obtenir per a totes les variants estudiades un perfil d'activitat òptica a l'ultraviolat llunyà (200 – 260nm). Aquests perfils d'activitat òptica a l'ultraviolat llunyà es van realitzar també per a la subunitat I_N i per la inteïna sencera *Npu DnaE*. Per il·lustrar aquests resultats, a la figura 4.29 es mostren els espectres de dicroisme circular enregistrats.

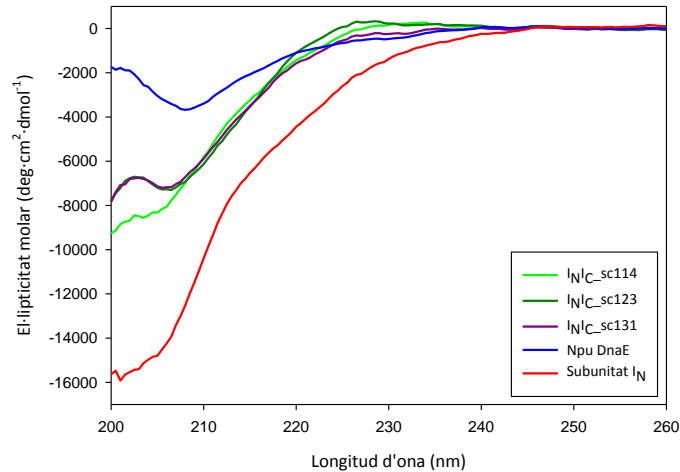


Fig. 4.29. Espectres de dicroisme circular normalitzats a l'ultraviolat llunyà (200 – 260nm) de les diferents variants de $I_{N|C_sc}$, de *Npu DnaE* i de la subunitat I_N . Condicions experimentals: tampó fosfat 10 mM, DTT 1mM pH 7 a una concentració de proteïna de 1mg/ml.

Com es pot observar a la figura 4.29, *Npu DnaE* presentava un mínim d'el·lipticitat als 208 nm mostrant que la proteïna es trobava composta principalment per fulls β . Si la proteïna fos composta només per fulls β aquest mínim hauria d'estar als 215 nm, no obstant, la presència d'hèlix α que presenta mínims d'el·lipticitat a 208 i 226 nm, va fer que el mínim es desplaçés dels 215 als 208 nm (Greenfield and Fasman, 1969; Holzwarth and Doty, 1965; Norma, 2006; Venyaminov et al., 1993). El mateix s'observava en el cas de les variants $I_{N|C_sc131}$ i 123. En aquests casos però, hi havia també mínims d'el·lipticitat als 200 nm indicant també que hi havia part de la proteïna que es trobava desestructurada. En el cas de la variant $I_{N|C_sc114}$ el mínim d'el·lipticitat s'observava principalment als 200 nm assenyalant que la proteïna tenia una part important desestructurada i per últim remarcar que la subunitat I_N presentava un mínim d'el·lipticitat molt evident als 200 nm la qual cosa indicava que aquesta es trobava bàsicament desestructurada.

4. Resultats i Discussió

A fi d'estimar el contingut d'hèlix α i full β de cada una de les proteïnes assajades, es van utilitzar els programes Dichroprot (Deleage and Geourjon, 1993) i K2D3 (Louis-Jeune et al., 2011). Els resultats obtinguts, juntament amb els valors d'estructura secundària de *Npu DnaE* observats per Oeemig i col·laboradors (Oeemig et al., 2009) es mostren a la taula 4.9.

Taula 4.9. Contingut en hèlix α i fulls β de les diferents proteïnes estudiades, amb predicció dels valors de *Npu DnaE* observats per Oeemig i col·laboradors (Oeemig et al., 2009)*

Proteïna	% Hèlix α	% Fulls β
<i>Npu DnaE</i> *	9	46
<i>Npu DnaE</i>	8.86	45.03
I _N I _C _sc131	8.75	37.86
I _N I _C _sc123	7.83	38.27
I _N I _C _sc114	7.51	36.14
Subunitat I _N	4.13	25.85

Els resultats que es presenten en la taula 4.10, per les diferents proteïnes estudiades mostren que l'element d'estructura secundària més abundant en aquestes variants eren els fulls β i que la *Npu DnaE* assajada presentava pràcticament els mateixos valors que els de la *Npu DnaE* descrita per Oeemig i col·laboradors. Utilitzant aquesta metodologia per a l'estimació d'estructura secundària, es consideren diferències significatives aquelles que són superiors al 5%. En el cas de l'hèlix α no hi havia unes diferències superiors al 5% exceptuant el cas de la subunitat I_N. En el cas dels fulls β si que hi havia diferències significatives entre la *Npu DnaE* i les variants I_NI_C_sc. Segons els valors obtinguts, aquestes variants haurien perdut part del seu contingut en fulls β i, per tant, hauria incrementat la quantitat de proteïna desplegada. Entre les tres variants de I_NI_C_sc no s'observaven diferències significatives en la quantitat de full β que presentaven. El contingut en fulls β de la subunitat I_N era molt inferior respecte a les altres proteïnes assajades. Per tant, segons aquests resultats, la subunitat I_N és la que presentava un grau de desestructuració més alt, aproximadament un 70% de la proteïna, i les variants I_NI_C_sc presentaven un grau de plegament similar entre elles i lleugerament inferior al de la *Npu DnaE*. Aquests resultats indiquen que la subunitat I_C contribueix de manera crítica al plegament global de la inteïna, i que, incorporant ja els primers residus d'aquesta (I_NI_C_sc114 (β 11)), el plegament ja és significatiu en comparació al plegament de la subunitat I_N.

4.3.4 Anàlisi de l'estructura per ressonància magnètica nuclear

Per tal de conèixer de manera més aproximada com la subunitat I_C influïa en el plegament global de la inteïna *Npu DnaE*, es van caracteritzar per RMN les diferents variants de I_NI_C_sc. Es va realitzar l'assignació dels espectres de ressonància de ¹H, ¹³C i ¹⁵N mitjançant una aproximació de les assignacions prèvies realitzades a la *Npu DnaE* (Heinamaki et al., 2009). En tots els casos es va dissoldre la proteïna en 90% H₂O / 10% D₂O per tal de poder fer correctament el lock de deuteri. Com a referència es va utilitzar DSS (àcid 2,2-dimetil-2-silapentan-5-sulfònic) a una concentració de

50 µM. La mostra es va tractar prèviament amb DTT 1 mM per tal d'eliminar així possibles interaccions de les cisteïnes. Es van enregistrar els espectres que es detallen a continuació:

^1H	Espectre monodimensional de protó. Permet avaluar de manera ràpida i senzilla la qualitat de la mostra de proteïna.
$^1\text{H}^{15}\text{N}$ HSQC	Espectre bidimensional que permet la visualització dels nitrògens i protons amida de l'esquelet de la cadena polipeptídica i de les cadenes laterals de les arginines, asparagines, glutamines i triptòfans.
$^1\text{H}^{13}\text{C}$ HSQC	Espectre bidimensional que permet la visualització dels carbonis i protons alifàtics i/o aromàtics.
TOCSY	<i>Total Correlation Spectroscopy</i> . Espectre bidimensional que permet la visualització dels protons d'un mateix sistema d'espí.
NOESY	<i>Nuclear Overhauser Effect Spectroscopy</i> . Espectre bidimensional que permet la visualització de les interaccions dipolars o efectes NOE entre protons de sistemes d'espí seqüencials ($n - n-1$) o ($n - n+1$) o entre protons que es troben a una distància inferior a 5 Å. Per tant, aquest espectre permet la visualització de NOE's, és a dir, la detecció de nuclis atòmics de diferents residus que es troben a una distància inferior a 5 Å en l'estructura tridimensional de la proteïna. A partir d'aquests espectres es pot obtenir informació important de la conformació d'una proteïna.

La informació proporcionada per cada un dels espectres bidimensionals va permetre realitzar les assignacions que es descriuen als següents apartats.

4.3.4.1 Espectres monodimensionals de ^1H

A fi d'avaluar de manera ràpida el plegament de les diferents proteïnes estudiades, es van realitzar espectres monodimensionals de ^1H . Els primers espectres que es van realitzar van ser els de la subunitat I_N i els de la inteïna *Npu DnaE* sencera. A la figura 4.30 es mostren els espectres obtinguts.

La dispersió dels desplaçaments químics en els espectres de RMN d'una proteïna són un bon indicador del seu estat natiu. Les proteïnes plegades presenten una dispersió de senyals molt superior a la d'una proteïna desplegada, en les quals, els senyals tendeixen a agrupar-se en regions (Zeeb and Balbach, 2004). En aquest cas, com es pot observar en la figura 4.31, la subunitat I_N presentava poca dispersió de senyals la qual cosa indicava que aquesta subunitat presentaria poca estructura i l'amplada dels senyals suggereix que adoptaria una conformació de glòbul fos. En el cas de la *Npu DnaE* presentava una elevada dispersió de senyals i a més, s'observaven molts senyals per sobre els 9 ppm i per sota els 0 ppm, la qual cosa indica que la proteïna es trobava plegada i estructurada.

Aquests resultats indiquen doncs, que la subunitat I_N sola no és capaç d'adoptar estructura per ella mateixa i que es necessita la subunitat I_C o part d'ella per tal d'adoptar l'estructura nativa. A partir d'aquí, es van realitzar els espectres monodimensionals de les diferents variants de $I_{N|C_sc}$ tal i com es mostra a la figura 4.30.

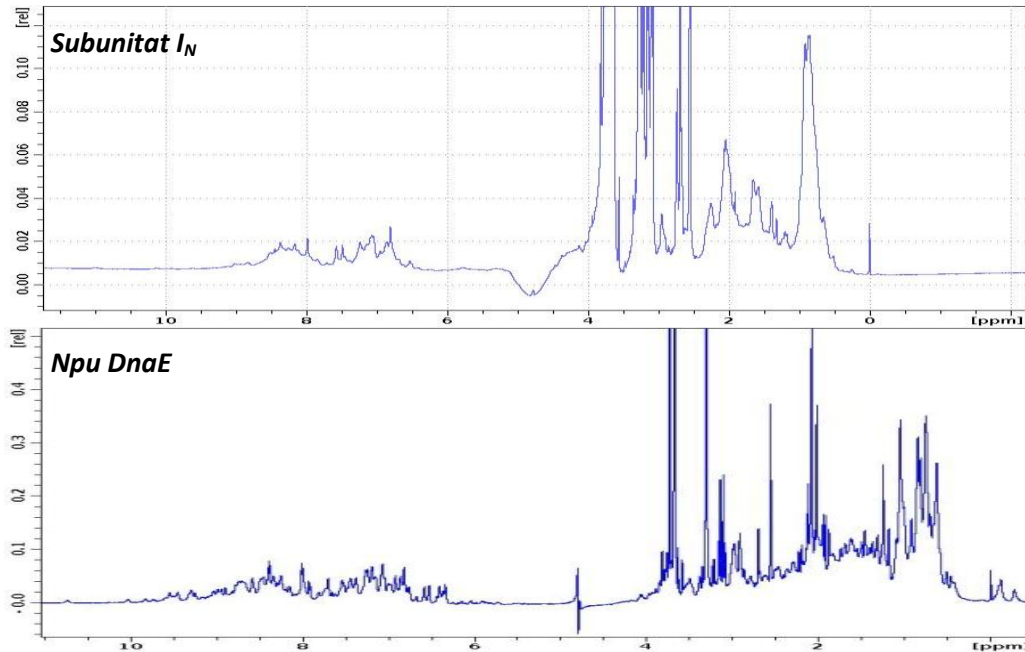


Fig. 4.30. Espectres monodimensionals de 1H de la subunitat I_N i de la inteïna *Npu DnaE*. Els espectres van ser obtinguts amb un espectròmetre Bruker de 800 MHz a 25°C en KH_2PO_4 10mM, DTT 1mM, 10% D_2O pH 7.

Els espectres monodimensionals de 1H de les diferents variants $I_{N|C_sc}$ mostraven com a mesura que el fragment de la subunitat I_C afegit era més gran, més dispersió hi havia, i per tant, més estructurada es trobava la proteïna. En el cas de la variant $I_{N|C_sc114}$, es va obtenir un espectre molt similar al de la subunitat I_N , poca dispersió i pics amples, la qual cosa indicava que aquesta variant presentava també una estructura de glòbul fos. Aquest resultat mostrava que incorporant només la seqüència pertanyent a la primera regió en conformació β ($\beta 11$) de la subunitat I_C , no és suficient per al plegament global de la proteïna. En les variants $I_{N|C_sc123}$ i $I_{N|C_sc131}$ s'obtenien uns espectres on els senyals eren més fins i dispersos. En aquestes dues variants doncs, els espectres indicaven que la proteïna es trobava ja, com a mínim, amb una regió plegada important. Per tant, incorporant la regió pertanyent a la subunitat I_C corresponent a les cadenes $\beta 11$, $\beta 12$ i $\beta 13$, s'aconseguia un plegament global important de la proteïna.

A partir d'aquestes observacions es van realitzar els espectres bidimensionals de les variants $I_{N|C_sc123}$ i $I_{N|C_sc131}$ per tal de provar de fer-ne les assignacions i obtenir l'estructura secundària de les dues variants. En el cas de la variant $I_{N|C_sc114}$, es va realitzar també algun espectre bidimensional (TOCSY, NOESY), els quals, presentaven unes senyals molt amples i poc disperses

indicant que la proteïna no presentava una estructura estable, d'acord amb els resultats de l'espectre monodimensional de ^1H .

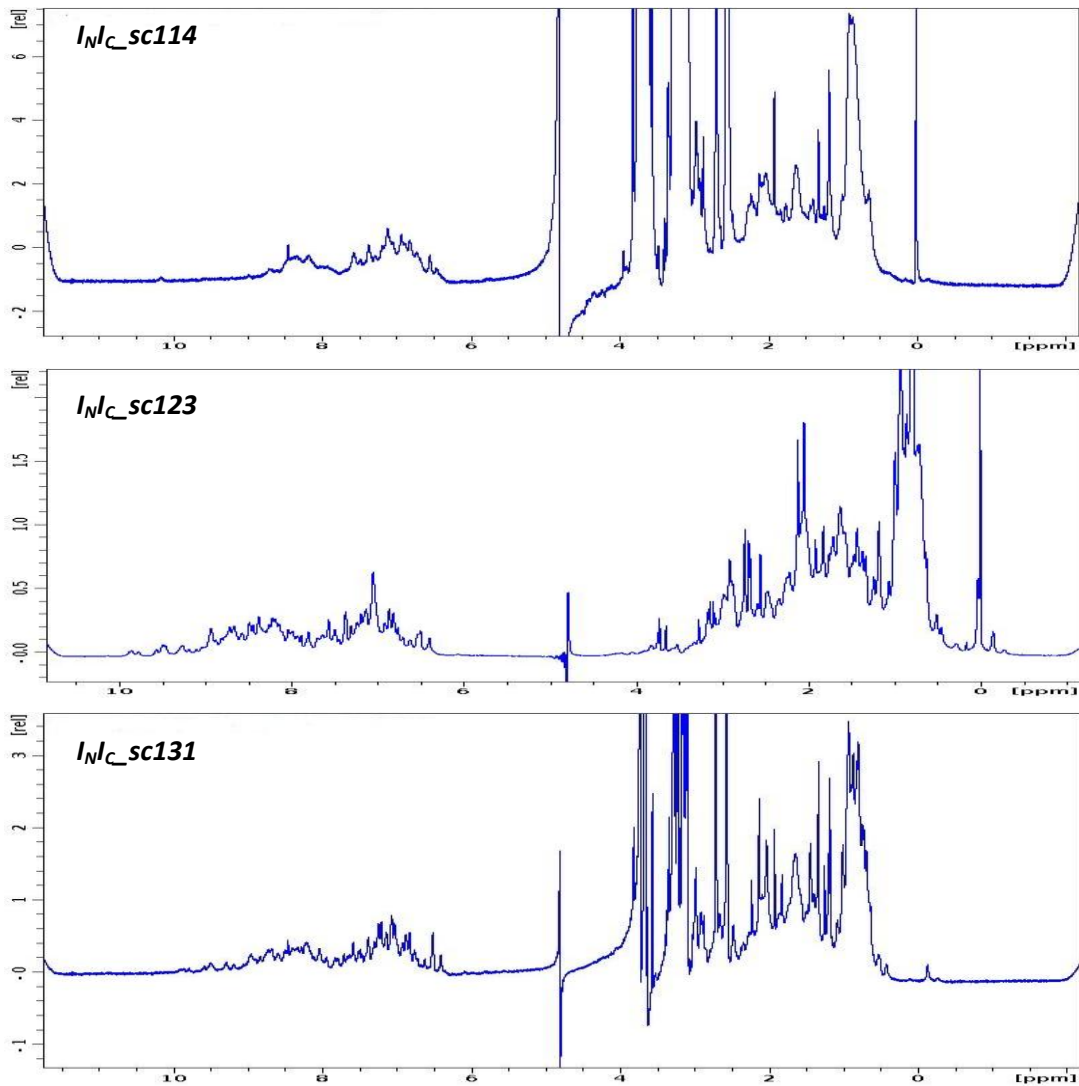


Fig. 4.31. Espectres monodimensionals de ^1H de les variants $I_{N|_C_sc114}$, $I_{N|_C_sc123}$ i $I_{N|_C_sc131}$. Els espectres van ser obtinguts amb un espectròmetre Bruker de 800 MHz a 25°C en el cas de $I_{N|_C_sc114}$ i $I_{N|_C_sc131}$ i a 5°C en el cas de $I_{N|_C_sc123}$. Les mostres es trobaven dissoltes en KH_2PO_4 10mM, DTT 1mM, 10% D_2O pH 7.

4.3.4.2 Determinació de l'estructura secundària de les variants $I_{N|_C_sc123}$ i $I_{N|_C_sc131}$

Els espectres bidimensionals enregistrats $^1\text{H}^{15}\text{N}$ HSQC, $^1\text{H}^{13}\text{C}$ HSQC, TOCSY, NOESY, es van dur a terme amb un espectròmetre Bruker de 800 MHz a 25°C en el cas de la variant $I_{N|_C_sc131}$ i a 5°C en el cas de la variant $I_{N|_C_sc123}$. Les assignacions es van fer seguint el mètode descrit per Wüthrich (Wüthrich, 1986) amb el programa Sparky (T. D. Goddard and D. G. Kneller, SPARKY 3, University of California, San Francisco) amb l'ajuda de les assignacions prèvies realitzades en la *Npu DnaE* (Heinamaki et al., 2009), dipositades al *Biological Magnetic Resonance Bank* amb el codi 16009 i

utilitzant la nomenclatura establerta per DYANA (“Dynamics algorithm for NMR applications”) (Gunter et al., 1997).

L’assignació dels espectres es va realitzar gràcies als desplaçaments químics descrits per Hideo i col·laboradors (Heinamaki et al., 2009) i es van assignar com a mínim tots els protons α i NH de tots els residus de les dues variants, exceptuant les regions dels extrems N i C terminals. No es va aconseguir assignar els residus 1 a 8 en el cas de la variant I_NI_C_sc123 i 1 a 4 en el cas de I_NI_C_sc131. En el cas de l’extrem C-terminal es va aconseguir fer les assignacions fins al residu 121 en ambdues variants. Una vegada feta l’assignació seqüencial, es van prendre els valors de desplaçament obtinguts dels protons α de cada residu (δ experimental) i es van restar els valors teòrics en pèptids no estructurals (δ random coil). Aquests valors teòrics es van calcular tenint en compte l’entorn del residu, el pH i temperatura de la mostra a l’hora de realitzar els espectres (Kjaergaard et al., 2011; Kjaergaard and Poulsen, 2011).

$$\Delta\delta = \delta_{\text{experimental}} - \delta_{\text{random coil}}$$

El valor resultant de la resta és el que s’anomena desplaçament conformacional ($\Delta\delta$) (figura 4.32). Un valor positiu de $\Delta\delta$ en els protons α indica que el residu es troba en una conformació característica de fulla β mentre que els valors negatius són indicatius d’una conformació tipus hèlix α . L’anàlisi d’aquests valors va indicar que el plegament de les variants s’ajustava a l’estructura secundària descrita anteriorment de la inteïna sencera *Npu DnaE* (Heinamaki et al., 2009), de manera que es pot concloure que les variants mantenen l’estructura secundària, almenys en les regions assignades.

4.3.4.3 Conservació de l’estructura terciària de les variants I_NI_C_sc123 i I_NI_C_sc131 respecte a la inteïna *Npu DnaE*

Malauradament els espectres enregistrats no van ser gaire resolutius tal i com es mostra a la figura 4.33 ja que els senyals eren molt amples i hi havia molta superposició entre ells la qual cosa dificultava el procés d’assignació. No obstant, amb l’ajuda de les assignacions prèvies de la inteïna sencera (Heinamaki et al., 2009), es van buscar en els espectres NOESY, tots aquells NOE’s de mitja (fins a i, i+4) i llarga distància (més de 4 residus) de les zones més desapantallades dels espectres més resolutius. Es van buscar sobretot NOE’s entre metils i es van assignar també NOE’s molt evidents entre NH-HA i NH-NH. L’assignació d’aquests NOE’s permetia observar regions de la proteïna pròximes entre sí, malgrat es trobessin allunyades en la seva seqüència aminoacídica. Si aquests NOE’s s’observaven també en la inteïna *Npu DnaE* sencera, era indicatiu que els aminoàcids els quals se’ls havia assignat el NOE conservaven l’estructura i posició respecte la inteïna sencera. Per tant, aquests NOE’s era un indicador que la conformació on es trobaven es mantenia en les variants estudiades. A la taula 4.10 es mostren tots els NOE’s de mitja i llarga distància assignats per cada una de les variants i que es troben conservats en la inteïna *Npu DnaE*.

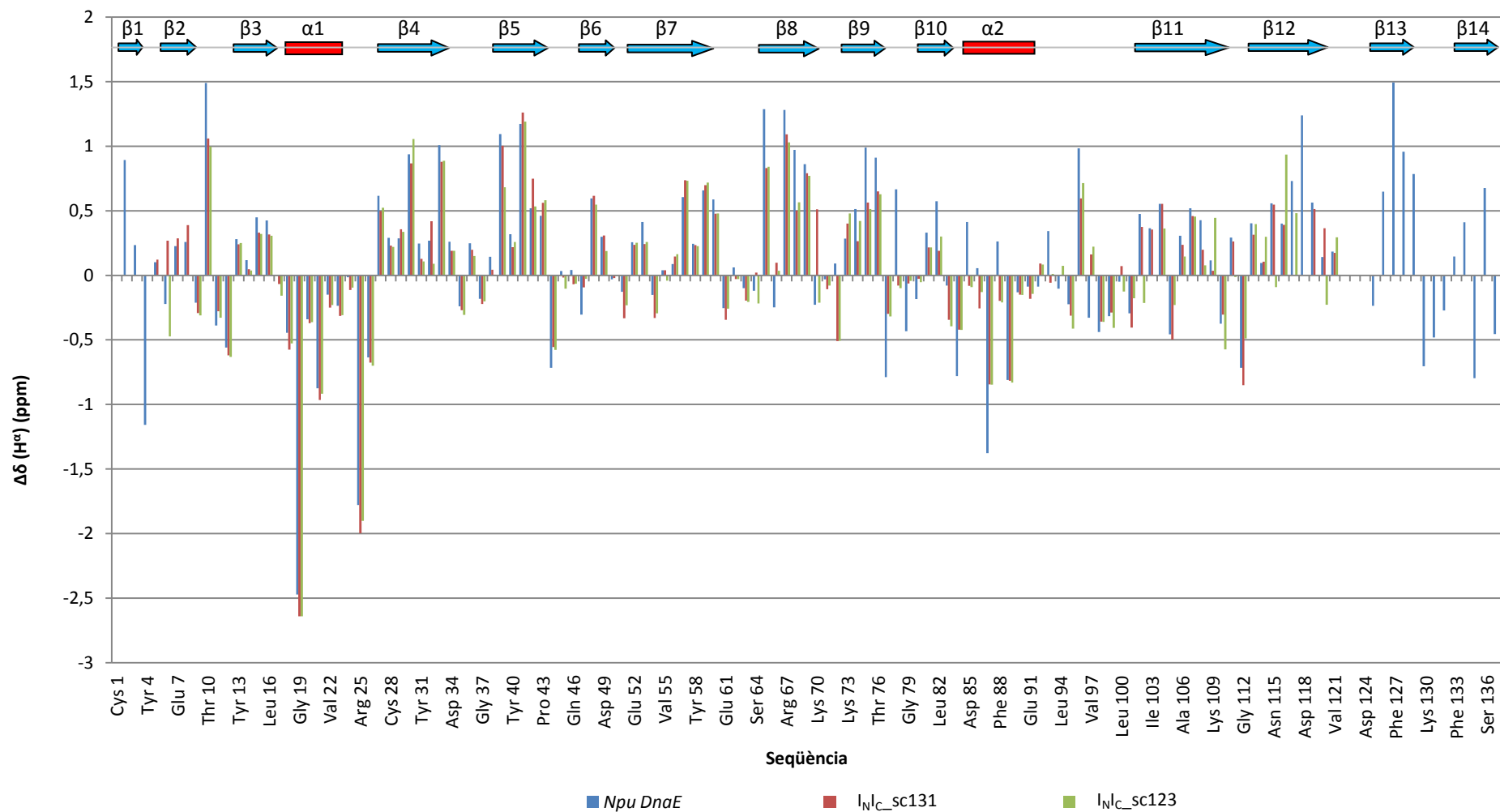


Fig. 4.32. Diferències entre els valors de desplaçament conformacional dels protons α de les variants *I_NC_sc131*, *I_NC_sc123* i la inteïna *Npu DnaE* (Biological Magnetic Resonance Bank codi: 16009) i els valors teòrics en pèptids no estructurats. Les fletxes blaves indiquen les regions amb conformació de cadena β i els rectangles vermells les regions amb conformació d'hèlix α .

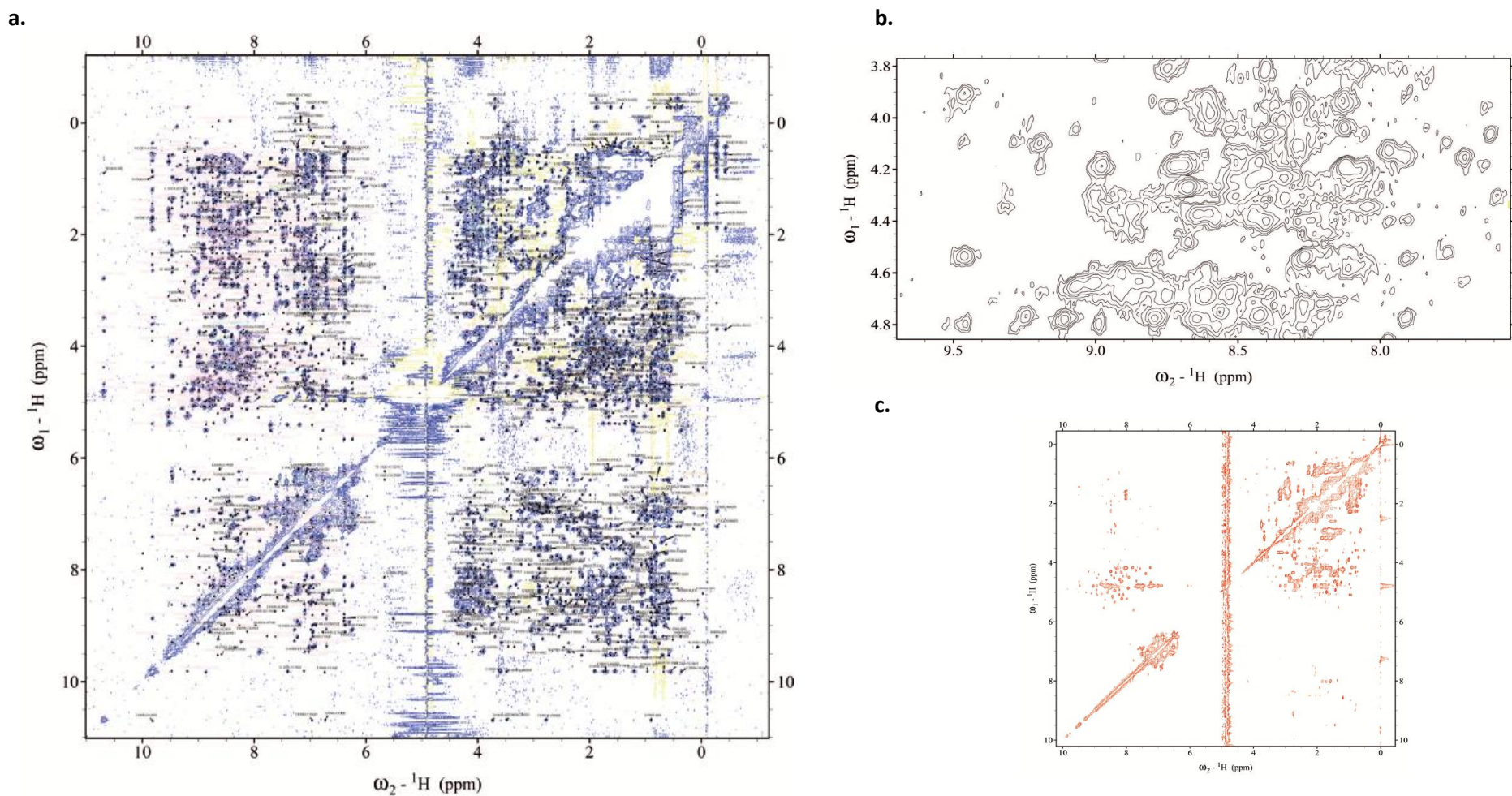


Fig. 4.33. Espectres bidimensionals de ^1H de $\text{INL}_{\text{sc123}}$. Els espectres van ser obtinguts amb un espectròmetre Bruker de 800 MHz a 5°C . (a) espectre NOESY, (b) ampliació de l'espectre NOESY on s'observa una regió amb senyals més desapantallades més fàcils d'assignar i una regió amb senyals més amples i solapats que van dificultar l'assignació, (c) espectre TOCSY poc resolutiu on hi mancaven molts senyals. Els espectres obtinguts de la variant $\text{INL}_{\text{sc131}}$ eren d'una qualitat similar als descrits.

4. Resultats i Discussió

Taula 4.10. NOE's de metils, HA-NH i NH-NH de mitja (fins a i, i+4) i llarga distància assignats a les variants I_{Nlc}_sc123 i I_{Nlc}_sc131 assignats també a la teïna *Npu DnaE* (codi PDB: 2KEQ)

NOE'S Metils					NOE'S Metils				
residu 1	residu 2	I _{Nlc} _sc123	I _{Nlc} _sc131	El. Estruc. 2ria	residu 1	residu 2	I _{Nlc} _sc123	I _{Nlc} _sc131	El. Estruc. 2ria
Y4HB2	Y117QE	X	X	L1-β12	I66 HG12	I84 HG12	X	X	β8-α2
Y4HB3	Y117QE	X	X	L1-β12	I66 HG12	I84 QD1	X	X	β8-α2
Y4 NH	Y117 QE	X	O	L1-β12	A68 QB	I84 QD1	X	X	β8-α2
L9 QD1	Y31 QD	X	X	β2-β4	F74 QD	I84 NH	X	O	β9-α2
L9 QD1	Y31 QE	X	X	β2-β4	F74 QD	I84 QD1	X	X	β9-α2
L9 HG	Y31 QD	X	X	β2-β4	F74 QD	I84 HG12	X	X	β9-α2
L9 HG	Y31 QE	X	X	β2-β4	F74 QE	I84 QD1	X	X	β9-α2
T10 HA	Y13 QD	X	X	L2	M75QB	M95NH	O	X	L9-L10
T10 QG2	G14 NH	X	O	L2	V77 QQG	D93 HB3	X	X	L9-L10
T10NH	G14HA	X	X	L2	V77 QQG	M95 NH	X	X	L9-L10
T10 HG12	Y31 QD	X	O	L2-β4	V77 QQG	N102 HD21	X	X	L9-L10
T10 HG12	Y31 QE	X	O	L2-β4	I84 QD1	I87 QG2	X	O	α2
V11 QG2	T29 HB	X	X	L2-β4	I84 QD1	I105 QG2	X	O	α2-β11
V11 QG2	Y31 QD	X	O	L2-β4	I84 HG13	I87 HB	X	O	α2
V11 QG2	Y31 QE	X	O	L2-β4	F88 QD	I105 HG13	X	X	α2-β11
E12 NH	C28 HB3	X	O	L2-L3	F88 QD	I105 HG2	X	X	α2-β11
Y13 QD	L16 HG	X	O	L2-β3	F88 QE	I105 HG2	X	X	α2-β11
Y13 QD	I26 QG2	X	X	L2-L3	F88 QE	T107 QG2	X	X	α2-β11
Y13 QD	C28 HB3	X	X	L2-L3	M95HA	N102HD21	X	X	L10
I21 NH	I26 QD1	X	O	α1-L3	I105HG2	R108HE	O	X	β11
I21 HG13	I26 QD1	X	X	α1-L3	NOE's HA NH				
I21 QG2	W47 HE1	X	X	α1-β6	residu 1	residu 2	I _{Nlc} _sc123	I _{Nlc} _sc131	El. Estruc. 2ria
I21 QG2	W47 HD1	X	X	α1-β6	L9 HA	L16 NH	X	O	β2-β3
I21 QD1	W47 HD1	X	O	α1-β6	T10 HA	Y13 NH	X	O	L2
V22HA	W47QB	X	X	α1-β6	T10 HA	G14 NH	X	X	L2
V22HA	W47HD1	X	X	α1-β6	T10NH	G14 HA	X	X	L2
V22HA	W47HE1	X	X	α1-β6	T10NH	L15 HA	X	O	L2-β3
V22 QG1	W47 HA	X	O	α1-β6	E12 NH	C28 HA	X	O	L2-L3
V22 QG1	H48 HA	X	O	α1-β6	T29 HA	V44 NH	X	X	β4-L5
R25 HB3	W47 HE1	X	X	L3-β6	V33 HA	Y40 NH	X	O	β4-β5
I26 NH	W47 HE1	X	X	L3-β6	F56 HA	K109 NH	O	X	β7-β11
I26 HA	W47 HE1	X	O	L3-β6	F56 HA	Y110 NH	X	O	β7-β11
E27 HA	W47 H22	X	O	L3-β6	F56 HA	L111 NH	X	X	β7-β11
T29 HB	T41 QG2	X	X	β4-β5	E57 HA	L111 NH	X	O	β7-β11
T29 QG2	Q42 HA	X	O	β4-β5	E57 NH	K109 HA	O	X	β7-β11
Y31 QD	T41 QG2	X	X	β4-β5	E57 NH	Y110 HA	X	O	β7-β11
Y31 QE	T41 QG2	X	X	β4-β5	Y58 NH	R67 HA	X	O	β7-β8
V33QQG	G37NH	O	X	β4-L4	C59 HA	I66 NH	X	O	β7-β8
D34 NH	Y40 HB3	X	O	L4-β5	L60 NH	L65 HA	X	O	β7-β8
D34 NH	Y40 QD	X	X	L4-β5	L60 HA	A106 NH	X	O	β7-β11
V44 QG2	W47 HE1	X	X	L5-β6	K73 HA	L82 NH	O	X	β9-β10
V44 QG1	W47 HE1	X	X	L5-β6	F74 NH	M81 HA	X	O	β9-β10
V44 QG1	W47 H22	X	X	L5-β6	I84 HA	I87 NH	X	X	α2
V55 QG1	A68 NH	X	X	β7-β8	E86 HA	E89 NH	X	X	α2
V55 NH	Y110 QD	X	O	β7-β11	I87 HA	L92 NH	X	X	α2-L10
V55 NH	Y110 QE	X	O	β7-β11	G112 HA	N115 NH	X	O	β11-L11
V55 HB	Q 114 HE22	X	X	β7-L11	NOE's NH NH				
F56 QD	T69 HA	X	X	β7-L8	residu 1	residu 2	I _{Nlc} _sc123	I _{Nlc} _sc131	El. Estruc. 2ria
F56 QE	Y110 QD	X	X	β7-β11	T10 NH	G14 NH	X	O	L2
F56 QE	Y110 QE	X	O	β7-β11	S32NH	Y40 NH	X	X	β4-β5
E57 NH	L111 HG	X	O	β7-β11	D34 NH	Y40 HB3	X	O	L4-β5
E57 NH	L111 QD2	X	O	β7-β11	D34 NH	G37 NH	X	X	L4
Y58 HB2	I84 QD1	X	O	β7-α2	D34 NH	N38 NH	X	X	L4
Y58 HB3	I84 QD1	X	O	β7-α2	V55 NH	G112 NH	X	O	β7-β11
C59 HA	L65 QD2	X	X	β7-β8	E57 NH	K109 NH	X	O	β7-β11
C59 NH	I105 QG2	X	X	β7-β11	Y58 NH	I66 NH	X	O	β7-β8
L60 QGD	I66 QG2	X	O	β7-β8	C59 NH	T107 NH	X	O	β7-β11
I66 HB	I84 QD1	X	O	β8-α2					

*El símbol x fa referència a NOE assignat i el símbol o fa referència a NOE no assignat per cada una de les variants. α fa referència a les diferents hèlix α, β a les diferents cadenes β i L als diferents loops que hi ha entre els diferents elements d'estructura secundària.

Una vegada assignats els NOE's de mitja i llarga distància, es van poder identificar quins eren els residus que conservaven la mateixa posició respecte a la inteïna sencera i plegada. D'aquesta manera, es va poder fer un mapejat per cada una de les variants per tal de veure en quines regions de la inteïna es trobaven aquests residus. Així, es va poder obtenir una aproximació de les regions d'estructura terciària conservada en cada una de les variants respecte la inteïna nativa i plegada. A la figura 4.34 es mostra el mapejat obtingut per cada una de les variants.

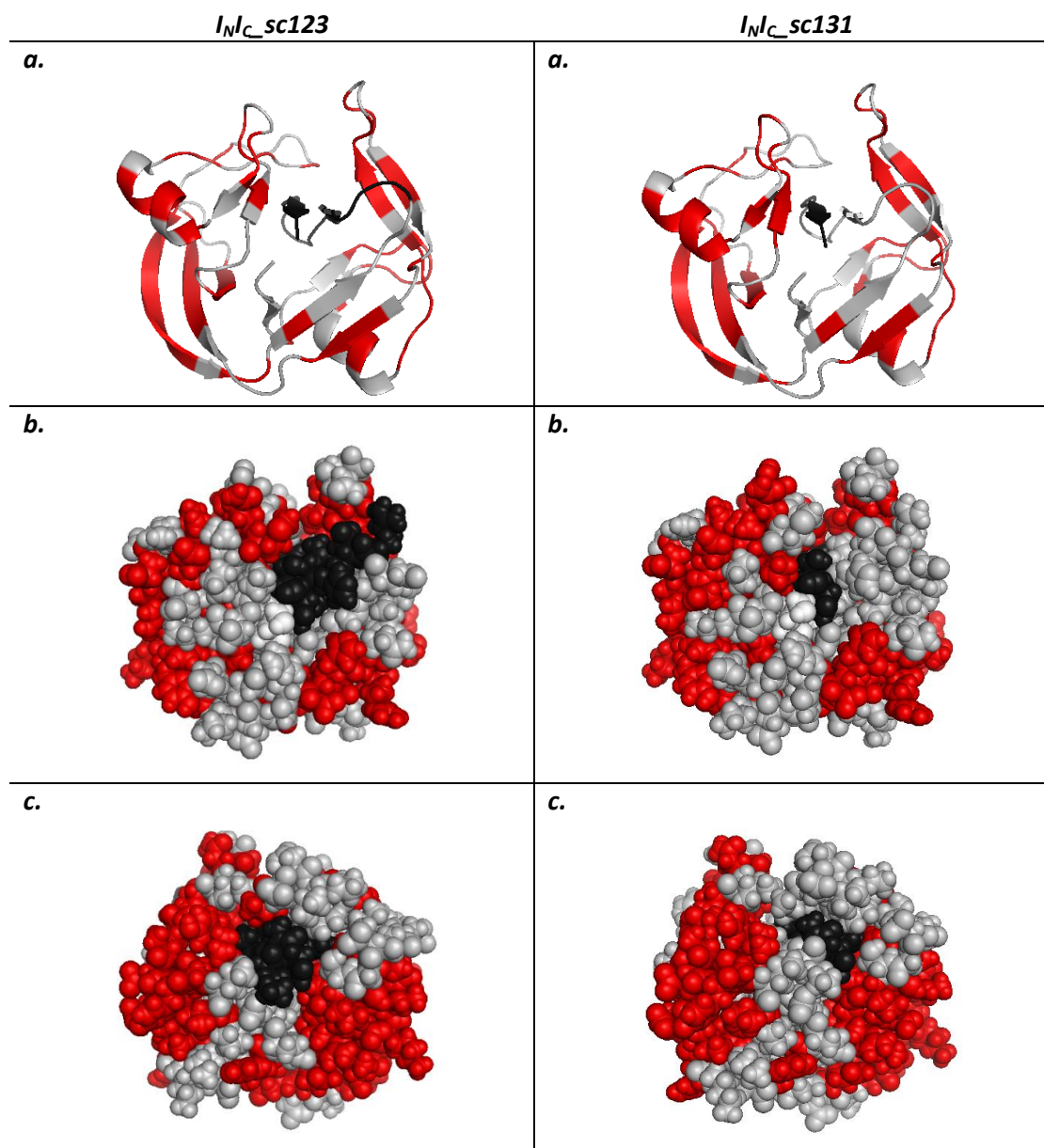


Fig. 4.34. Mapejat de les variants $I_{N/C}sc123$ i $I_{N/C}sc131$. En vermell s'observen els residus als quals se'ls ha assignat un NOE a mitja o llarga distància, en gris als residus els quals no se'ls ha assignat cap NOE a mitja o llarga distància i en negre els residus de la subunitat I_C que no són presents a cada una de les variants. Les figures (a) mostren una representació esquemàtica de l'estructura tridimensional de la inteïna en el qual s'indica gràficament els resultats de la taula 4.10. Això indica que aquests elements d'estructura secundària es mantindrien a cada una de les variants. Les figures (b i c) mostren una representació de la inteïna per la cara frontal i la cara dorsal respectivament utilitzant esferes per representar els àtoms basant-se en el radi de Van der Waals. S'observa que molts residus els quals no se'ls ha trobat un NOE a mitja o llarga distància són els que troben en contacte amb la regió de la subunitat I_C que manca a cada una de les variants.

4. Resultats i Discussió

Tal i com s'ha visualitzant en la figura 4.34, les dues variants $I_{Nl_c_sc123}$ i $I_{Nl_c_sc131}$ presentaven moltes regions amb la mateixa conformació que la inteïna *Npu DnaE*. Sobretot, aquesta conformació es mantenia en les zones que menys contacte tenien amb el fragment de la subunitat I_c que mancava en cada una d'elles. Tampoc s'observaven diferències significatives en les regions conservades en les dues variants $I_{Nl_c_sc123}$ i $I_{Nl_c_sc131}$. Per tal de confirmar que aquestes regions mantenien la mateixa estructura que la inteïna *Npu DnaE* sencera, a partir dels NOE's de mitja i llarga distància entre HA-NH i NH-NH, es va estudiar si l'orientació de diferents cadenes β de les diferents variants es mantenia igual a les de la inteïna sencera. En el cas de la *Npu DnaE*, la majoria de cadenes β es troben en orientació antiparal·lela. Per saber si dues cadenes β són antiparal·leles cal fixar-se amb uns NOE's concrets (NH-NH i NH-HA). És característic de cadenes β antiparal·les que el protó NH d'un residu formi NOE amb l'NH del residu $j-1$ que té enfrontat en l'altra cadena β i amb el HA del residu j que té enfrontat. A més, també es poden observar NOE's entre dos HA enfrontats en dues cadenes β i aquest HA pot formar a més, NOE amb l'NH del residu que té enfrontat. Observant almenys 2 NOE's d'aquest tipus, ja es pot deduir que l'estructura i orientació dels elements d'estructura secundària d'aquella regió es conserva. A la figura 4.35 es mostren els NOE's identificats a cada una de les cadenes β . La orientació que presenten aquestes cadenes β en les variants $I_{Nl_c_sc123}$ i $I_{Nl_c_sc131}$ i si aquesta orientació que presenten es conserva en la inteïna sencera *Npu DnaE* es troba recollida a la taula 4.11.

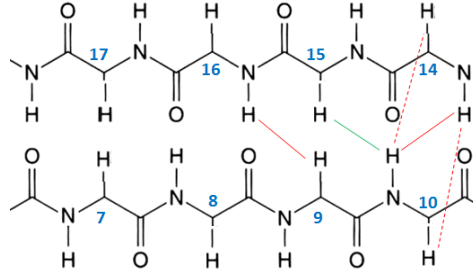
Taula 4.11. Relació entre les diferents cadenes β de la inteïna *Npu DnaE* i si aquesta relació es manté entre les variants $I_{Nl_c_sc123}$ i $I_{Nl_c_sc131}$

Cadena 1	Cadena 2	Sentit	Es manté la relació?
$\beta 1$	$\beta 12$	Paral·lel	No s'observa
$\beta 2$	$\beta 3$	Antiparal·lel	Sí
$\beta 4$	$\beta 5$	Antiparal·lel	Sí
$\beta 6$	$\beta 12$	Antiparal·lel	No s'observa
$\beta 7$	$\beta 11$	Antiparal·lel	Si
$\beta 7$	$\beta 8$	Antiparal·lel	Si
$\beta 9$	$\beta 10$	Antiparal·lel	Sí
$\beta 13$	$\beta 14$	Antiparal·lel	No es pot formar

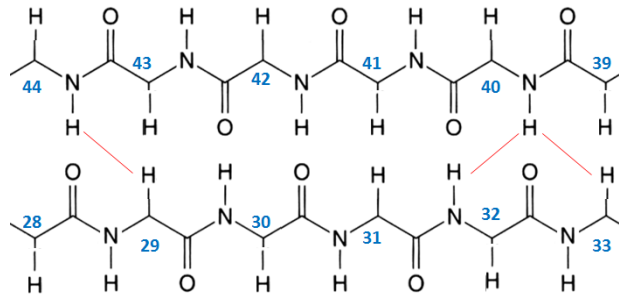
Malauradament no es van poder fer assignacions dels residus compresos entre l'1 i el 8 ni a partir del residu 121 en les dues variants $I_{Nl_c_sc123}$ i $I_{Nl_c_sc131}$. Per aquest motiu no s'han pogut identificar diferències significatives entre les dues variants, ja que, segons els resultats obtinguts, aquestes presenten un plegament molt semblant. Amb aquests resultats, es podria confirmar que introduint les dues primeres cadenes β de la subunitat I_c , s'aconsegueix ja una estructura global de la inteïna. Malauradament, a partir del residu 121 no s'ha pogut determinar quina estructura adopta aquesta regió en les diferents variants, sobretot en el cas de $I_{Nl_c_sc131}$. Això és probablement degut al fet que $\beta 13$ que no es trobaria en $I_{Nl_c_sc123}$ però si en $I_{Nl_c_sc131}$ no pot adoptar tampoc la conformació β a causa de l'absència en totes dues variants de $\beta 14$.

NOE's $\beta 2$ - $\beta 3$ **$\beta 2$ (7-9) $\beta 3$ (15-17)**

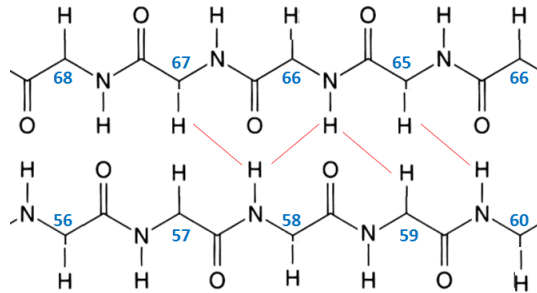
NH10-HA14
 NH10-HA15
 NH14-HA10
 NH16-HA9
 NH10-NH14

**NOE's $\beta 4$ - $\beta 5$** **$\beta 4$ (29-33) $\beta 5$ (39-43)**

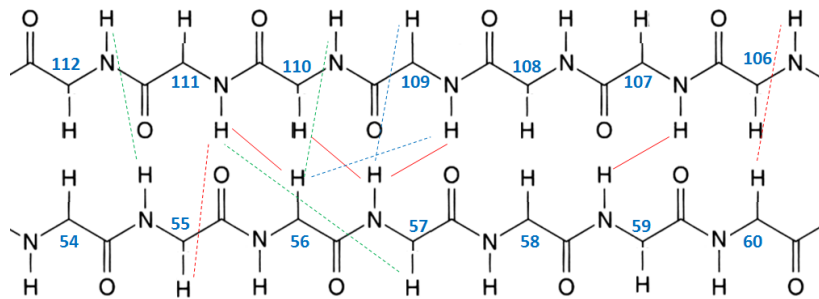
NH40-HA33
 NH44-HA29
 NH32-NH40

**NOE's $\beta 7$ - $\beta 8$** **$\beta 7$ (53-60) $\beta 8$ (65-68)**

NH58-HA67
 NH60-HA65
 NH66-HA59
 NH58-NH66

**NOE's $\beta 7$ - $\beta 11$** **$\beta 7$ (53-60) $\beta 11$ (105-113)**

NH57-HA109
 NH57-HA110
 NH106-HA60
 NH109-HA56
 NH110-HA56
 NH111-HA56
 NH111-HA57
 NH55-NH112
 NH57-NH109
 NH59-NH107

**NOE's $\beta 9$ - $\beta 10$** **$\beta 9$ (73-74) $\beta 10$ (81-83)**

NH74-HA81
 NH82-HA73

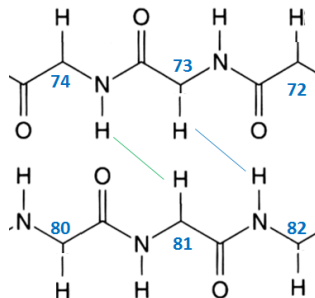


Fig. 4.35. NOE's a llarga distància identificats en les diferents cadenes β de les variants I_{Nlc_sc123} i I_{Nlc_sc131} i conservats en la intèina Npu DnaE. Per cada cadena β s'indica els residus els qual en formen part i el llistat de NOE's identificats. Les línies en vermell indiquen NOE's assignats a les dues variants, les línies en verd els NOE's assignats només a I_{Nlc_sc123} i les línies en blau als NOE's assignats només a I_{Nlc_sc131} . Les línies discontinues indiquen NOE's complementaris de cadena β assignats que són conseqüència de difusió de l'espí.

5. Discussió General

5. DISCUSSIÓ GENERAL

Les inteïnes estan esdevenint una interessant i potencial eina en el camp de la biotecnologia i la biologia molecular ja que s'utilitzen en diferents tècniques de purificació, marcatge isotòpic i control d'activitat de proteïnes (Elleuche and Poggeler, 2010). Per tal de dissenyar noves aplicacions i millorar les tècniques obsoletes es requereix una millor comprensió de la seva estructura i la seva funció com també un millor coneixement, en el cas de les inteïnes en trans, del procés associatiu de les subunitats I_N i I_C .

Una de les línies de recerca del nostre grup se centra en l'estudi estructural i la caracterització cinètica de determinades trans inteïnes. Dins de les trans inteïnes, la família *DnaE* és de les més estudiades per les excel·lents característiques que presenta. De totes elles però, la que ha rebut un major interès és la *Npu DnaE* de *Nostoc punctiforme* per la seva velocitat de reacció, la seva robustesa i l'elevada afinitat que presenten les subunitats I_N i I_C que fan d'ella una inteïna excel·lent per ser utilitzada com a eina en el camp de la biologia molecular (Shah et al., 2012; Shah et al., 2011; Zettler et al., 2009).

La majoria d'estudis realitzats en inteïnes incloent la *Npu DnaE*, se centren en l'estabilitat, cinètica i estudi del procés associatiu dels fragments de les trans inteïnes. Per aquest motiu, es va decidir progressar en aquests estudis per tal d'entendre millor com funcionen les inteïnes i dissenyar-ne i desenvolupar-ne de més eficients a partir de la trans inteïna *Npu DnaE*. El primer objectiu va ser estudiar la termoestabilitat que presenta aquesta inteïna quan forma part d'una única seqüència polipeptídica, quan es troben les seves subunitats I_N i I_C associades i quan es troba la subunitat I_N aïllada. Primerament, es va aconseguir expressar i purificar aquestes variants en la fracció soluble la qual cosa va permetre, simplificar molt el procés de purificació, atès que, es van evitar protocols de solubilització de cossos d'inclusió, plegament de la proteïna i concentració. Els estudis de termoestabilitat han demostrat que la variant més estable és aquella en que la inteïna es troba formant una única cadena polipeptídica i la menys estable la subunitat I_N aïllada. Per tant, això indicaria que les subunitats I_N i I_C han de romandre juntes per aconseguir una millor estabilitat a la inteïna. Aquest fet concordaria amb estudis previs els quals s'ha determinat que el plegament global de la inteïna, i per tant la seva estabilitat, és degut a una sinèrgia mútua entre la subunitat I_N i I_C (Shah et al., 2013b; Zheng et al., 2012). Aquesta sinèrgia de plegament vindria dominada per un component entàlpic que supera el cost entròpic desfavorable. El canvi d'entropia desfavorable es dona a causa de la pèrdua de llibertat conformacional causada per la transició d'estat desordenat a ordenat durant l'associació (He et al., 2001). Malgrat que l'entropia sigui desfavorable, el canvi d'entalpia és favorable i prové bàsicament del llarg nombre de contactes polars que s'estableixen entre les subunitats I_N i I_C (Zheng et al., 2012). Aquest fenomen és conegut com a compensació

entalpia-entropia (Calderone and Williams, 2001) i normalment va associat a un plegament proteic després de la unió de les subunitats. Tal i com s'ha demostrat en aquest treball, la subunitat I_N aïllada presenta una estructura de glòbul fos la qual cosa concorda amb que presenti una baixa estabilitat i també amb que no s'ha observat un replegament un cop desnaturalitzada per temperatura. Per altra banda, si que s'ha aconseguit un replegament parcial de les altres variants gràcies a la sinèrgia de plegament existent entre les subunitats I_N i I_C. És a dir, la presència de la subunitat I_C a través de les interaccions que es donen durant l'associació, permetria que la subunitat I_N es tornés a plegar, obtenint així, un replegament global de la inteïna. Aquest replegament va ser observable gràcies a la introducció de TCEP-HCl a les mostres. Aquest agent reductor va permetre evitar les possibles interaccions no desitjades de les diferents Cys que podrien dificultar l'estudi de la seva termoestabilitat. El fet que aquesta inteïna presenti una T_m d'entre 50 i 60°C tant quan es troba formant una única cadena polipeptídica com quan es troba associada en forma de I_N i I_C, permet el seu ús en aplicacions que necessitin temperatures fisiològiques de 37°C com és el cas del treball amb cèl·lules eucariotes de mamífer (Selgrade et al., 2013; Subramanyam et al., 2013). La termoestabilitat que presenta aquesta inteïna, també li confereix capacitat de ser activa a temperatures superiors a la fisiològica, de 40°C o més. Per altra banda, aquesta proteïna és també estable a baixes temperatures tal i com s'ha demostrat en altres treballs, fet que li confereix la capacitat de treballar amb exteïnes que són sensibles a temperatures altes (Zettler et al., 2009).

S'ha pogut comprovar que totes les variants tendeixen a ser més estables a pHs lleugerament bàsics. Aquest fet concordaria amb que diferents autors utilitzin valors de pH lleugerament bàsics per treballar amb aquesta inteïna (Cheriyana et al., 2013; Iwai et al., 2006; Shah et al., 2011). A aquests valors de pH, les Cys de la inteïna podrien trobar-se lleugerament ionitzades i actuar com a millors nucleòfils (Claus and Winyard, 2009; Paulsen and Carroll, 2013). En aquestes condicions, la inteïna es manté més soluble, ja que a partir de pH's lleugerament àcids (inferiors a 6) precipita segurament per l'aproximació al seu punt isoelèctric de 4.84. A més, la solubilitat que presenta és bona ja que en tots els assajos de termoestabilitat i de RMN, les seves diferents variants eren molt solubles a pHs lleugerament bàsics. S'ha arribat a obtenir concentracions superiors a 5mg/ml de proteïna sense observar la formació de precipitats.

Els resultats obtinguts a partir dels estudis d'estabilitat permeten concloure que la inteïna *Npu DnaE* tant formant una única cadena polipeptídica com en forma de subunitats I_N i I_C, presenta una bona termoestabilitat, per tant aconsegueix un dels requisits que permet que aquesta inteïna sigui interessant en el camp de la biotecnologia i biologia molecular. Un dels altres requisits que compleix la inteïna i que s'ha observat en aquest treball és l'elevada solubilitat que presenta tant la inteïna sencera com totes les variants estudiades en aquesta tesi doctoral. Aquesta elevada solubilitat de la inteïna ja ha estat descrita en estudis previs (Ramirez et al., 2013; Zettler et al., 2009).

5. Discussió General

Per altra banda, un dels esforços en l'estudi de la família d'intèines *DnaE* se centra en conèixer el paper de les Cys 1 i +1 i analitzar la seva substitució per altres residus que no siguin tant reactius, com Ser o Thr. La substitució de les Cys per aquests altres residus podria disminuir les possibles reaccions no desitjades que es poden produir a causa de la seva reactivitat. Els residus de Ser i Thr són usats en les posicions 1 i +1 en determinades intèines. Centrant-nos en la posició +1, de totes les intèines recollides a InBase, el 39.5% tenen Cys a la posició +1, el 34.6% presenten una Ser, el 22% tenen una Thr i només un 0.04% presenten altres residus (Shemella et al., 2011).

En aquest treball es va decidir substituir les Cys 1 i +1 per una Ser i veure com aquest canvi afectava l'activitat de la intèina. En aquest cas, es va decidir fer els estudis primerament només amb la Ser i no amb la Thr, ja que, la Ser és un aminoàcid molt més semblant a la Cys que no la Thr i per tant, això simplificaria la comprensió dels efectes que podria tenir aquesta mutació. També, en el conjunt global de totes les intèines de InBase, és més comú un residu de Ser a la posició +1 que no un de Thr. Així mateix, és factible pensar que de manera natural una Thr a la posició +1 és un residu menys eficient en el mecanisme de tall i unió comparant-lo amb els residus de Cys o Ser. Per exemple, en el nostre grup de recerca es treballa també amb la trans intèina *Neq Pol* de *Nostoc punctiforme* (Choi et al., 2006), la qual presenta una Thr a la posició +1. Fins a l'actualitat, no s'ha aconseguit veure activitat apreciable d'aquesta intèina utilitzant diferents extèines i diferents condicions de pH i temperatura. En un estudi fet per Iwaï i col·laboradors, es va avaluar i comparar l'activitat que presentaven 20 intèines diferents utilitzant GB1, MBP i el domini SH3 com a extèines. De totes elles, 8 presentaven una Thr a la posició +1, 8 presentaven una Cys i 4 presentaven una Ser a la posició +1. De les 8 intèines que presentaven una Thr en posició +1, 3 no presentaven activitat o era inferior al 5%. Concretament, aquestes 3 intèines eren *Pho CDC21-2*, *Mxe GyrA* i *Neq Pol*. Per altra banda, es van classificar aquestes 20 intèines estudiades en funció del seu rendiment. De les 20 intèines, 6 van ser classificades com a intèines amb rendiment baix (rendiment inferior al 5% o nul) i concretament, 3 d'elles presentaven una Thr a la posició +1, 2 d'elles presentaven una Cys i només una presentava una Ser a la posició +1 (Ellila et al., 2011). Aquests resultats, semblarien indicar que la Thr en la posició +1 és el residu menys eficient en el mecanisme de tall i unió comparat amb la Cys o la Ser.

A la base de dades InBase, es van analitzar totes les intèines de bacteris termòfils i es va observar que la meitat de les intèines d'aquests bacteris presenten una Ser a la posició +1. Aquest fet va fer pensar que, aprofitant l'elevada termoestabilitat que presenta *Npu DnaE*, s'aconseguiria recuperar l'activitat de la intèina amb la mutació C+1S incrementant la temperatura, ja que la Ser +1 era el residu predominant en les intèines d'aquests organismes termòfils. A la taula 5.1 es mostra un recull de totes les intèines de bacteris termòfils amb els residus que presenten a les posicions 1 i +1. Tenir un coneixement dels residus més abundants en aquestes dues posicions en intèines de bacteris termòfils, podria ajudar a entendre el perquè la *Npu DnaE* C+1S recupera l'activitat en incrementar la temperatura.

5. Discussió General

Primer de tot, centrant-nos en el residu 1, a la taula 5.1 s'observa que totes les inteïnes de bacteris termòfils tenen a la posició 1 una Cys. Cap d'elles presenta una Ser o Thr, aquest fet donaria suport als resultats obtinguts, on no s'ha observat activitat de la inteïna amb la mutació C1S. En aquest cas, es fa evident doncs, que la Cys en la posició 1 és essencial pel mecanisme de tall i unió de la inteïna *Npu DnaE*. Per completar aquest estudi, s'han buscat totes les trans inteïnes que es troben recollides a InBase i s'ha buscat quins residus tenen cada una d'elles en la posició 1 i +1. Aquest llistat es troba recollit a la taula 5.2.

Taula 5.1. Inteïnes de bacteris termòfils on es mostra el residu que tenen en la posició 1 i +1 recollides a InBase (Perler, 2002)

Nom	Residu 1	Residu +1	Organisme
<i>Aae RIR2</i>	C	C	<i>Aquifex aeolicus strain VF5</i>
<i>Chy RIR1</i>	C	C	<i>Carboxydotherrnus hydrogenoformans Z-2901</i>
<i>Dge DnaB</i>	C	T	<i>Deinococcus geothermalis DSM11300</i>
<i>Rma DnaB</i>	C	S	<i>Rhodothermus marinus</i>
<i>Rma DnaE</i>	C	S	<i>Rhodothermus marinus DSM 4252</i>
<i>Tfus RecA-1</i>	C	T	<i>Thermobifida fusca</i>
<i>Tfus RecA-2</i>	C	S	<i>Thermobifida fusca</i>
<i>Tfus Tfu2914</i>	C	S	<i>Thermobifida fusca</i>
<i>Tth-HB27 DnaE-1</i>	C	S	<i>Thermus thermophilus HB27</i>
<i>Tth-HB27 DnaE-2</i>	C	S	<i>Thermus thermophilus HB27</i>
<i>Tth-HB27 RIR1-1</i>	C	T	<i>Thermus thermophilus HB27</i>
<i>Tth-HB27 RIR1-2</i>	C	C	<i>Thermus thermophilus HB27</i>
<i>Tth-HB8 DnaE-1</i>	C	S	<i>Thermus thermophilus HB8</i>
<i>Tth-HB8 DnaE-2</i>	C	S	<i>Thermus thermophilus HB8</i>
<i>Tth-HB8 RIR1-1</i>	C	T	<i>Thermus thermophilus HB8</i>
<i>Tth-HB8 RIR1-2</i>	C	C	<i>Thermus thermophilus HB8</i>

Dels resultats obtinguts i recollits a la taula 5.2, s'observa que pràcticament totes les trans inteïnes d'aquesta base de dades, tenen una Cys tant en la posició 1 com en la +1. Només *Dra Snf2* i *Neq Pol* presenten altres residus que no són Cys, tot i que les dues tenen una Thr en la posició +1. Aquest fet indicaria doncs que les trans inteïnes en general necessiten d'una Cys en la posició 1 i +1 per dur a terme de manera eficient el mecanisme de tall i unió. Cal remarcar que totes aquestes inteïnes són d'eubacteris excepte *Neq Pol* que és d'un organisme arquea hipertermòfil (Choi et al., 2006).

El fet que la mutació de la Cys 1 inhibeixi l'activitat de la inteïna, fa que la inteïna doble mutant C1S – C+1S tampoc presenti activitat. A més, de totes les inteïnes de termòfils (aproximadament 120) que hi ha a InBase (Perler, 2002) siguin de bacteris, archea o virus, cap d'elles presenta una Ser a la posició 1 i +1. Aquest fet indicaria que una Ser a les posicions 1 i +1 no és viable per la correcte activitat de les inteïnes.

Taula 5.2. Trans inteïnes recollides a InBase (Perler, 2002) on es mostra el residu que tenen en la posició 1 i +1

Nom	Residu 1	Residu +1	Organisme
<i>Aha DnaE</i>	C	C	<i>Aphanothece halophytica</i>
<i>Aov DnaE</i>	C	C	<i>Aphanizomenon ovalisporum</i>
<i>Asp DnaE</i>	C	C	<i>Anabaena species PCC7120</i>
<i>Ava DnaE</i>	C	C	<i>Anabaena variabilis</i>
<i>Cra-CS505 DnaE</i>	C	C	<i>Cylindrospermopsis raciborskii CS-505</i>
<i>Csp-CCY0110 DnaE</i>	C	C	<i>Cyanothece sp. CCY0110</i>
<i>Csp-PCC7424 DnaE</i>	C	C	<i>Cyanothece sp. PCC 7424</i>
<i>Csp-PCC7822</i>	C	C	<i>Cyanothece sp. PCC 7822</i>
<i>Csp-PCC8801 DnaE</i>	C	C	<i>Cyanothece sp. PCC 8801</i>
<i>Cwa DnaE</i>	C	C	<i>Crocospaera watsonii WH 8501</i>
<i>Dra Snf2</i>	A	T	<i>Deinococcus radiodurans R1</i>
<i>Maer-NIES843</i>	C	C	<i>Microcystis aeruginosa NIES-843</i>
<i>Mcht-PCC7420 DnaE</i>	C	C	<i>Microcoleus chthonoplastes PCC7420</i>
<i>Mtu-T17 RecA</i>	C	C	<i>Mycobacterium tuberculosis T17</i>
<i>Nosp-CCY9414 DnaE</i>	C	C	<i>Nodularia spumigena CCY9414</i>
<i>Npu DnaE</i>	C	C	<i>Nostoc punctiforme PCC73102</i>
<i>Nsp-PCC7120</i>	C	C	<i>Nostoc species PCC7120</i>
<i>Oli DnaE</i>	C	C	<i>Oscillatoria limnetica</i>
<i>Sel-PC7942 DnaE</i>	C	C	<i>Synechococcus elongatus PC7942</i>
<i>Sel-PCC6301 DnaE</i>	C	C	<i>Synechococcus elongatus PCC 6301</i>
<i>Ssp DnaE</i>	C	C	<i>Synechocystis species, strain PCC6803</i>
<i>Ssp-PCC7002 DnaE</i>	C	C	<i>Synechocystis species, strain PCC 7002</i>
<i>Tel DnaE</i>	C	C	<i>Thermosynechococcus elongatus BP-1</i>
<i>Ter DnaE-3c</i>	C	C	<i>Trichodesmium erythraeum IMS101</i>
<i>Tvu DnaE</i>	C	C	<i>Thermosynechococcus vulcanus</i>
<i>Neq Pol</i>	S	T	<i>Nanoarchaeum equitans</i>

De manera general, la Cys 1 estableix ponts d'hidrogen amb l'Asp present en la regió F (taula 1.1 de la introducció) i s'ha vist que aquestes interaccions són essencials pel primer pas del mecanisme de tall i unió (Johnson et al., 2007; Sun et al., 2005; Van Roey et al., 2007). El residu d'Asp protonat estableix la cadena lateral de la Cys 1 facilitant-ne l'atac nucleofílic sobre el carboni carbonílic de l'extrem N-terminal (Volkman and Mootz, 2013). Aquest pont d'hidrogen que s'estableix entre l'Asp de la regió F i la Cys 1 s'observa també en la *Npu DnaE*. En aquest cas es tracta de l'Asp 118, el pont d'hidrogen es forma també quan la posició 1 és ocupada per una Ser però no quan és ocupada per una Thr (figura 5.1). El pont d'hidrogen que s'estableix entre l'Asp 118 i la Cys 1 té una longitud de 2.46 Å, mentre que quan és una Ser, la longitud del pont d'hidrogen és lleugerament més alta i passa a ser de 2.76 Å. Aquest lleuger increment podria ajudar, en part, a fer més inestable la interacció entre Ser 1 i Asp 118, contribuint a que una Ser en la posició 1 bloquejés el mecanisme de tall i unió malgrat s'incrementi la temperatura. També, el fet que el grup OH de la Ser sigui menys electronegatiu i per tant menys nucleofíl que el grup SH de la Cys (Barron et al., 2010), podria

explicar en part el bloqueig del mecanisme de tall i unió en aquest mutant. D'altra banda, es coneix també que tant la Thr com la His de la regió B és important per al primer pas del mecanisme de tall i unió (Du et al., 2011a; Pearl et al., 2007; Romanelli et al., 2004). Aquests residus s'ha vist en diferents inteïnes que estableixen ponts d'hidrogen amb el residu 1 i sembla que el paper d'aquestes interaccions serien l'estabilització de l'intermediari de la reacció d'aquest pas (Ding et al., 2003). *Npu DnaE* es caracteritza per no presentar la His de la regió B, però si que presenta una Thr en aquesta regió, concretament la Thr 69, que estableix també ponts d'hidrogen amb el residu 1 tant quan aquest és una Cys, Ser o Thr (fig. 5.1). Per tant, el fet de canviar el residu de Cys per Ser, sembla que no afectaria les interaccions que s'estableixen amb la Thr 69 malgrat que el procés de tall i unió es trobi inhibït o inapreciable amb les tècniques emprades en aquest estudi.

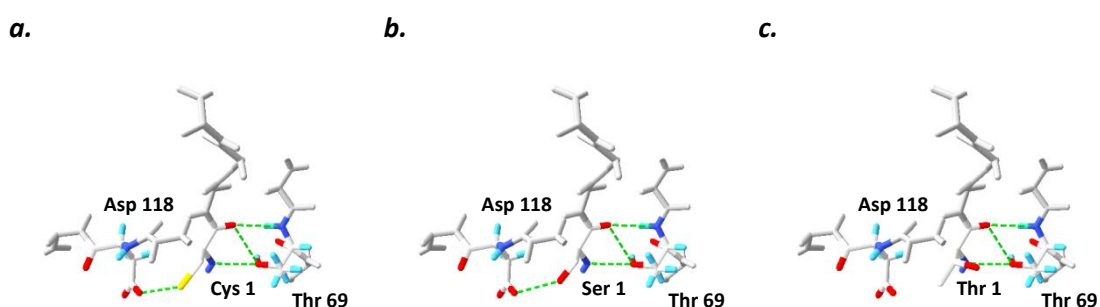


Fig. 5.1. Ponts d'hidrogen (línia discontinua verda) que s'estableixen entre el residu 1 i els residus Thr 69 i Asp 118 en la inteïna *Npu DnaE* i que estabilitzarien el pas 1 del mecanisme de tall i unió. (a) quan el residu 1 és una Cys, (b) quan el residu 1 és una Ser i (c) quan el residu 1 és una Thr. Imatges obtingudes a partir del programa *SwissPdb Viewer* (Guex and Peitsch, 1997).

Per altra banda, hi ha inteïnes que malgrat presentar un residu nucleofílic de Cys, Ser o Thr en la posició 1, requereixen també del triplet Trp, Cys i Thr, concretament es requereix d'un Trp a la regió B, una Cys a la regió F i una Thr a la regió G per tal que el mecanisme de tall i unió sigui actiu. El triplet Trp-Cys-Thr pot ser mutat per residus de característiques similars sense afectar el mecanisme de tall i unió menys la Cys. Aquesta Cys de la regió F resulta absolutament essencial per l'activitat de la inteïna i no pot ser canviat per cap altre residu malgrat sigui de característiques similars (Ser o Thr) (Volkman and Mootz, 2013). Un exemple d'aquest tipus de inteïnes és la MP-Catera Gp206 que presenta una Ser a la posició 1 juntament amb el triplet Trp-Cys-Thr. MP-Catera Gp206 sense la Cys de la regió F perd totalment la capacitat de dur a terme el mecanisme de tall i unió (Tori and Perler, 2011). En el cas de la *Npu DnaE*, aquesta no presenta cap dels tres residus en les regions esmentades i per tant, l'absència del triplet Trp-Cys-Thr podria contribuir també a explicar el perquè la mutació C1S inhibeix el mecanisme de tall i unió en la *Npu DnaE*. Cal remarcar que les inteïnes presenten una dinàmica en el seu esquelet polipeptídic essencial per a la seva activitat. En alguns casos, la mutació d'un dels residus essencials en el mecanisme de tall i unió provoca canvis significatius en la seva dinàmica i pot provocar canvis interns afectant la capacitat catalítica de determinats residus (Du et al., 2010; Lockless and Muir, 2009). Aquest fet es pot veure clarament en les diferències significatives d'activitat que presenten les trans inteïnes de la família *DnaE* malgrat la seva elevada homologia de

protó unit a l'oxigen ho és amb més força que l'unit al sofre. Així doncs, es necessita més energia per desprotegir l'enllaç OH que el SH, ja que l'energia d'enllaç depen en gran mesura i és proporcional a l'electronegativitat dels dos àtoms que formen la unió (Alcock, 1990). En el cas de la *Npu DnaE* amb la mutació C+1S, l'energia s'aconsegueix mitjançant un increment de temperatura. Aquest fet, explicaria que al incrementar la temperatura s'ha aconseguit desestabilitzar l'enllaç OH, fent que aquest grup perdés el protó, esdevingués més nucleòfil i incrementés l'activitat de la inteïna. Això també concordaria en què com més alta és la temperatura, més elevada és l'activitat d'aquesta ja que l'enllaç OH es troba més desestabilitzat i explicaria el perquè *Npu DnaE* amb la mutació C+1S és activa al incrementar la temperatura, en canvi, no seria suficient perquè fos funcional amb la mutació C1S. L'increment de temperatura també ha fet augmentar l'activitat que presenta la inteïna salvatge (Cys a la posició +1) però en menor mesura que en el cas del mutant C+1S (taula 4.6 resultats i discussió). El mateix que s'ha observat en el cas del mutant C+1S, l'augment de temperatura hauria ajudat a desestabilitzar l'enllaç SH de la cadena lateral de la Cys incrementant lleugerament l'activitat de la inteïna salvatge. En aquest cas, l'activitat de la inteïna no ha augmentat de manera tant significativa ja que, a temperatures baixes la inteïna ja presenta una molt bona activitat.

Cal esmentar que un increment de temperatura augmenta la flexibilitat i disminueix la compactació en enzims (Feller, 2010) i, a la vegada, una flexibilitat interna d'una proteïna és essencial per a la seva activitat ja que es redueixen les barreres energètiques i es permet un accés més fàcil al centre actiu (Daniel et al., 1998; Kokkinidis et al., 2012). El fet que un augment de temperatura augmenti la flexibilitat i accessibilitat als residus que duen a terme el procés de tall i unió podria contribuir en l'augment d'activitat en els mutants C+1S.

Globalment, en aquests estudis cinètics s'ha pogut comprovar que una Ser a la posició 1 o +1 inhibeix l'activitat de la inteïna a temperatura ambient. No obstant, un increment de la temperatura superior a 40°C fa que la inteïna recuperi la seva activitat quan la Ser es troba en posició +1 però no quan es troba en posició 1. Aquest fet ha estat observat tant quan la inteïna *Npu DnaE* es troba formant part d'una única cadena polipeptídica com quan es troba separada en forma de subunitat I_N i I_C. Els resultats, porten a la conclusió que la Cys 1 és essencial per l'activitat d'aquesta inteïna tal i com s'ha descrit recentment en altres treballs (Shah et al., 2013a). Per altra banda, la Cys +1 malgrat que també és essencial per el procés de tall i unió, es pot substituir per una Ser si es treballa a temperatures superiors als 40°C encara que amb una velocitat de reacció lleugerament inferior a la descrita per a la inteïna salvatge.

El fet que un increment de temperatura activi la inteïna *Npu DnaE* amb la mutació C+1S, fa pensar que aquesta inteïna podria ser utilitzada en aplicacions biotecnològiques basades en el mecanisme de tall i unió condicionat per temperatura. A més d'obtenir un mecanisme de tall i unió regulable per

5. Discussió General

temperatura, s'elimina, una de les 3 Cys que conté la inteïna, disminuint en part, la probabilitat de que es donin reaccions secundàries no desitjades. Concretament s'eliminarà la Cys+1 que en realitat es troba a l'exteïna (figura 5.3). Això suposaria una avantatge, ja que l'exteïna obtinguda a partir d'aquest mutant de *Npu DnaE*, no presentaria una Cys sinó una Ser, la qual cosa, disminuiria la reactivitat que podria presentar l'exteïna evitant-se reaccions no desitjades o alteracions en el seu plegament. L'únic requisit que s'hauria de tenir en compte per utilitzar aquesta inteïna en el tall i unió condicionat per temperatura, és que l'exteïna a utilitzar també tingués una bona termoestabilitat.

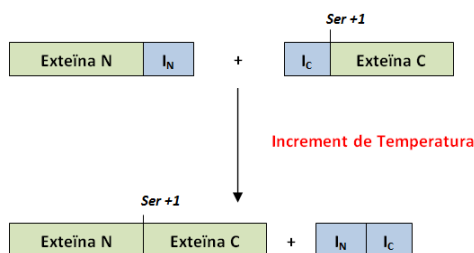


Fig. 5.3. Esquema del procés de tall i unió de la variant C+1S on s'observa que la Ser +1 formaria el primer residu de l'extrem C-terminal de l'exteïna al final del procés de tall i unió.

A part d'avaluar la importància dels residus de la posició 1 i +1 en el procés de tall i unió de *Npu DnaE*, també s'ha realitzat, a la vegada, estudis cinètics dels diferents mutants els quals poden ser comparats amb altres treballs cinètics. En el nostre cas, s'ha observat un $t_{1/2}$ de formació del producte de tall i unió que varia bastant en funció del mutant i de la temperatura. La millor condició en el cas del mutant C+1S en cis utilitzant la GFP com a exteïna ha sigut una temperatura de 60°C, mentre que, en el cas del mutant C+1S en trans utilitzant GB1 com a exteïna, les millors condicions es troben a 50 i 60°C. Els valors de $t_{1/2}$ han diferit molt segons si la inteïna es trobava en trans o en cis, en el cas de la inteïna en cis el valor de $t_{1/2}$ a 60°C és de 198.4 minuts, mentre que en el cas de la inteïna en trans, el valor de $t_{1/2}$ a 50 i 60°C ronda els 25 minuts. Per tant, el mecanisme de tall i unió en trans utilitzant la GB1 és 8 vegades més ràpid que el mecanisme de tall i unió en cis utilitzant la GFP com a exteïna. Això evidencia la capacitat que té l'exteïna de modificar l'activitat d'una inteïna tal i com s'ha demostrat en altres treballs (Ellila et al., 2011; Shah et al., 2013a). Per altra banda, cal esmentar que aquestes diferències de $t_{1/2}$ poden venir condicionades també a l'estabilitat que presenta la inteïna, ja que quan forma part d'una única cadena polipeptídica presenta una major termoestabilitat que quan es troba en forma de I_N i I_C associades. Per tant, aquesta major termoestabilitat que presenta la inteïna formant una única cadena polipeptídica ha permès que presentés una millor activitat a 60°C respecte als 50°C, mentre que aquest augment no s'observa quan la inteïna es troba partida ja que a 60°C es troba per sobre del seu valor de T_m i dificultaria la seva velocitat de reacció. A més, la termoestabilitat que presenten les exteïnes unides poden influenciar de manera significativa en l'activitat de la inteïna. Una exteïna poc termoestable pot disminuir l'activitat de la inteïna si es troba en condicions de temperatures superiors a la seva T_m. És

important remarcar que els valors de $t_{1/2}$ obtinguts són bastant més elevats als descrits per Muir i col·laboradors, els quals calculen un $t_{1/2}$ de la *Npu DnaE* en trans utilitzant SUMO com a exteïna de només 30.9 s a 30°C i 18.9 s a 37°C (Shah et al., 2012). Aquestes diferències en els valors de $t_{1/2}$ correspondrien bàsicament a que les condicions no eren les mateixes ja que ells treballen amb una altra exteïna, a pH 7,2 i els residus +1+2+3 de la exteïna C són Cys-Phe-Asn en comptes de Ser-Phe-Asn. En aquest mateix treball, estudien l'efecte de la mutació F+2G en la velocitat del mecanisme de tall i unió i observen un $t_{1/2}$ que incrementa fins a 70 min a 37°C, uns valors ja més aproximats als descrits en aquest treball per la mutació C+1S. Cal esmentar que en els assajos cinètics realitzats en aquest treball amb la inteïna en trans amb la Cys en posició +1, ja s'obtenen uns valors de $t_{1/2}$ d'uns 5-6 minuts, els quals són superiors als descrits al treball de Muir i col·laboradors. Es torna a fer evident doncs, la importància que té l'exteïna i les condicions de treball en el procés de tall i unió. A la vegada, aquests valors són també superiors al $t_{1/2}$ descrit de només 1 minut per l'equip de Mootz i col·laboradors utilitzant la *Npu DnaE* en trans a 37°C amb *head protein D* (gdP) com a exteïna N i tioredoxina com a exteïna C (Zettler et al., 2009). Una altra vegada, aquestes diferències, potser no tant significatives en aquest treball de Mootz i col·laboradors, es deuen a l'exteïna unida i també a les condicions de treball emprades.

Els estudis estructurals duts a terme utilitzant RMN han permès comprovar primerament com les subunitats I_N i I_C aïllades de la inteïna *Npu DnaE* es troben desestructurades i en forma de glòbul fos en el cas de la subunitat I_N . Un cas semblant a aquest, es descriu per la inteïna *Ssp DnaE* (Zheng et al., 2012), una trans inteïna de la família *DnaE* que presenta una homologia molt elevada amb la *Npu DnaE*. En global presenten un 54.7% d'homologia respecte la seqüència de la *Npu DnaE*, tot i que aquesta homologia ascendeix fins a un 67% si es compara només la subunitat I_N (figura 5.2) (Iwai et al., 2006). A més, els aminoàcids responsables de la funció catalítica son idèntics en ambdues inteïnes. En el cas de la subunitat I_N , la *Ssp DnaE* presenta un fragment de 21 residus que no es troba en la *Npu DnaE*. Aquests residus però, no formen part de la zona estructurada de la inteïna (Sun et al., 2005) i poden ser eliminats sense que perdi l'activitat (Ghosh et al., 2001). Per tant, coincidint amb aquest treball i el realitzat paral·lament al nostre treball sobre la *Npu DnaE* (Shah et al., 2013b), es pot concloure que el plegament global de la inteïna requereix l'associació de les subunitats degut primerament per la complementarietat de càrregues i posteriorment d'una sinèrgia en el plegament que indueix de manera sinèrgica el co-plegament d'ambdues subunitats per donar lloc a la inteïna sencera.

Una de les qüestions a resoldre és si la subunitat I_N requereix de tota la subunitat I_C per tal de plegar-se, o bé, només amb un fragment d'aquesta ja és suficient pel correcte plegament. En el mecanisme de captura i col·lapse proposat per Muir i col·laboradors, només amb els 13 primers residus de la subunitat I_C , ja és suficient perquè es doni un cert plegament global de la inteïna (Shah et al., 2013b). Els 13 primers residus correspondrien a la variant $I_N|_C_sc114$ descrita en aquesta tesi doctoral. En el

nostre cas, aquesta variant continuava presentant una estructura de glòbul fos tal i com es va observar també en la subunitat I_N . Malauradament, no es van aconseguir espectres bidimensionals bons d'aquesta variant. L'obtenció d'espectres bidimensionals de major qualitat, hauria permès observar segurament, zones estructurades de la variant que no s'haurien observat en el cas de la subunitat I_N aïllada. Això vindria suportat pel fet que en l'espectre monodimensional de $I_{N|C_sc114}$ malgrat que hi hagi poca dispersió de senyals, sí que es veuen alguns senyals que indiquen zones estructurades. A la vegada, l'espectre de dicroisme circular també indica que hi ha certa estructuració encara que menor que per les variants $I_{N|C_sc123}$ i $I_{N|C_sc131}$. En les diferents anàlisis informàtiques realitzades en aquest treball suggereixen que, la variant $I_{N|C_sc114}$ podria presentar cert grau d'estructuració. Per tant, aquesta variant necessitaria ser estudiada en més detall per acabar de classificar-la com a proteïna estructurada o desestructurada i poder veure les diferències que presenta respecte la subunitat I_N aïllada. Shah i col·laboradors van descriure un conjunt de parells iònics en determinats residus de les subunitats I_N i I_C per tal de comprovar que l'associació de les subunitats I_N i I_C es dona en primer lloc per una complementarietat de càrregues en la inteïna *Npu DnaE*. En global, van acabar identificant 4 grups de parells iònics: (E7, K130), (E52, E54, K113), (D85, E89, R108) i (E61, E91 i K104) (Shah et al., 2011). De tots aquests parells iònics, hi ha involucrats 3 residus que formen part dels primers 13 residus de la subunitat I_C , concretament són els residus 104, 108 i 113, per tant, tots els residus que formen part d'aquests parells electrònics excepte el 130 i que es troben en la subunitat I_C , es troben dins dels 13 primers residus. Cada un d'aquests residus identificats en la subunitat I_C es troba en un dels parells iònics identificats, per tant, això concordaria amb el fet que només amb els 13 primers residus de la subunitat I_C , ja és suficient perquè es doni una correcta complementarietat de càrregues per tal de desencadenar un plegament global de la inteïna. Per tant, aquests resultats estarien d'acord amb els que posteriorment el mateix grup de Shah i col·laboradors proposen amb el model de captura i col·lapse (apartat 1.4.5 de la introducció) (Shah et al., 2013b) (figura 5.4).

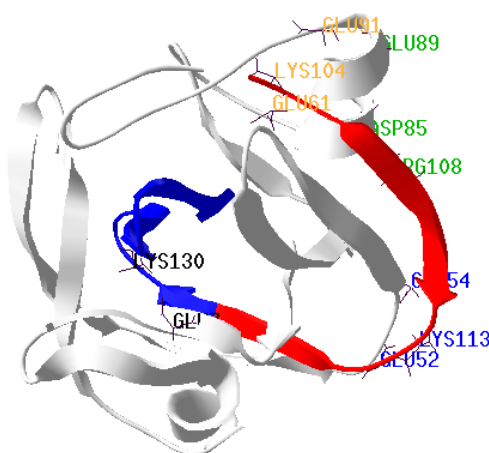


Fig. 5.4. Estructura tridimensional de *Npu DnaE* on s'indiquen els parells iònics identificats per Shah i col·laboradors. En gris s'identifica la part pertanyent a la subunitat I_N , en vermell els 13 primers residus de la subunitat I_C i en blau la resta de residus de la subunitat I_C . Cada grup de parells iònics es troba etiquetat en un color diferent. Dels 4 residus que formen parells iònics identificats en la subunitat I_C , 3 d'ells es troben en els 13 primers residus d'aquesta subunitat. Concretament són els residus 104, 108 i 113.

Els estudis preliminars de les variants $I_{N|C_sc123}$ i $I_{N|C_sc131}$ per RMN, espectres monodimensionals de 1H i NOESY posteriorment, concordaven tots en el fet que les dues variants $I_{N|C_sc123}$ i $I_{N|C_sc131}$ presentaven un bon grau d'estructuració i que, a la vegada, la variant $I_{N|C_sc131}$ presentava un grau d'estructuració més elevat que la variant $I_{N|C_sc123}$. D'aquesta observació, s'ha pogut concloure que la subunitat I_C és essencial pel plegament global de la inteïna i que com més part de subunitat I_C hi ha, més elevat és en principi el seu grau d'estructuració global.

Amb l'assignació dels espectres bidimensionals de 1H s'ha pogut observar, en primer lloc, que les variants $I_{N|C_sc123}$ i $I_{N|C_sc131}$ presenten una estructura secundària principalment de cadenes β , típic de la família d'inteiïnes *DnaE* (Duan et al., 1997) i a la vegada, conserven l'estructura secundària respecte la inteïna sencera *Npu DnaE* (Oeemig et al., 2009). Malauradament, no s'han pogut assignar els residus dels extrems N i C terminals i per tant, no s'ha pogut determinar quina estructura presenten les regions terminals. Aquests resultats juntament amb l'observació de NOE's a mitjana i llarga distància assignats en les dues variants i la conservació de l'orientació de la majoria de les cadenes β , demostra que les dues variants conserven l'estructura de la inteïna sencera i que, una vegada més, la subunitat I_C es troba involucrada en el plegament global de la inteïna. El fet que la primera part de la subunitat I_C es trobés involucrada en el plegament global de la inteïna, estaria d'acord amb el model de captura i col·lapse proposat per Muir i col·laboradors en la inteïna *Npu DnaE* (Shah et al., 2013b). En aquest model, es donaria primer una captura de la subunitat I_N i el fragment de la subunitat I_C a causa d'una interacció electrostàtica entre les dues regions desordenades seguit d'unes interaccions hidrofòbiques (Schreiber and Fersht, 1996; Shah et al., 2011) que conduirien a un col·lapse global del complex el qual permetria el plegament global de cada una de les variants (figura 1.19 de la introducció).

Per altra banda, cal esmentar que la subunitat I_C , composta per 36 residus i formada per les cadenes β_{11} , β_{12} , β_{13} i β_{14} , s'ha vist en aquest treball que estaria organitzada en dues regions independents. Per una banda els residus corresponents a les cadenes β_{11} i β_{12} i per l'altra els residus corresponents a β_{13} i β_{14} . En el cas de β_{11} ja ha adquirit en les variants $I_{N|C_sc123}$ i $I_{N|C_sc131}$ la mateixa conformació en quant a estructura secundària que la de la inteïna plegada, malgrat l'absència de la resta de residus que formen la subunitat I_C (β_{13} i β_{14} en el cas de la variant $I_{N|C_sc123}$ i β_{14} en el cas de la variant $I_{N|C_sc131}$). Segurament, els residus de β_{12} (fins el residu 120) també formen part d'aquesta primera part de la subunitat I_C , malgrat que en cap d'ambdues variants s'ha aconseguit assignar cap residu d'aquesta regió. Per altre banda, els residus que formen part de β_{13} i β_{14} sí que s'ha observat que tenen una forta dependència entre sí, ja que en el cas de la variant $I_{N|C_sc131}$ que presenta els residus corresponents a β_{13} però no de β_{14} , aquesta cadena β no ha estat observada, molt probablement, a causa de l'absència de β_{14} , indicant que impedeix la formació del β -gir β (β_{13} -gir β_{14}) i fent que aquesta regió es trobi completament desestructurada. Per tant, sembla com si els residus de β_{11} i β_{12} serien els que es trobarien més involucrats en el

plegament global de la inteïna, ja que són els que interaccionen més amb elements d'estructura secundària de la subunitat I_N . Això concordaria amb l'estructura tridimensional de *Npu DnaE* (codi PDB: 2KEQ) on s'observa que β_{11} i β_{12} interaccionen amb β_7 i β_6 , respectivament, que formen part de la subunitat I_N , mentre que, β_{13} i β_{14} interaccionen entre sí per adoptar l'esmentada estructura de β -gir- β fent que sigui una regió més independent de la subunitat I_N . L'estructura de β -gir- β s'observa també en les dues últimes cadenes β de *Ssp DnaE* (figura 5.5), les quals conformen una regió més independent de la subunitat I_N (Sun et al., 2005).

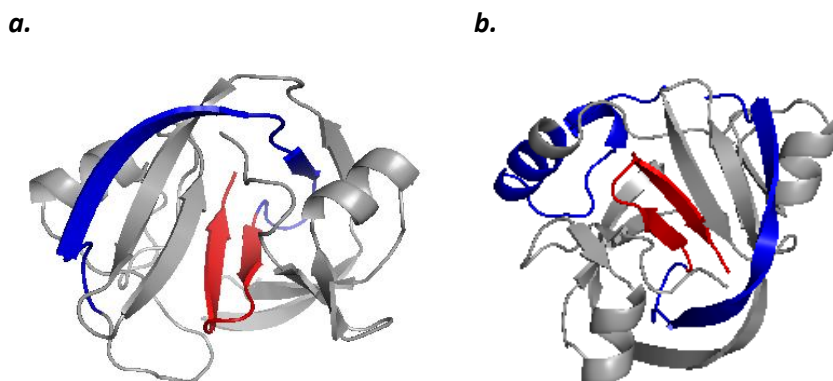


Fig. 5.5. Representació esquemàtica de l'estructura tridimensional de *Npu DnaE* (a) i *Ssp DnaE* (b). En gris es mostra la subunitat I_N i en blau la subunitat I_C exceptuant la regió final de cada una d'elles que és en color vermell per ressaltar l'estructura de les dues últimes cadenes β de la regió C-terminal en forma de β -gir- β .

Per tant, en els treballs d'estudi del plegament mitjançant RMN, hauria estat interessant haver obtingut una millor resolució en els espectres de les tres variants per haver pogut observar més diferències entre elles. També hauria estat interessant observar les diferències de plegament de cada una de les variants respecte la subunitat I_N i veure, a més, si el plegament és progressiu a mesura que s'afegeixen més residus de la subunitat I_C o no. Segons els resultats de dicromisme circular, el grau de plegament de les diferents variants seria així, però no s'ha pogut arribar a aquest nivell de resolució en els espectres de RMN. Malgrat tot, sí que s'ha aconseguit demostrar però que hi ha una sinèrgia de plegament entre la subunitat I_N i I_C pel plegament global de la inteïna, tal i com s'ha descrit també en la inteïna *Ssp DnaE* (Zheng et al., 2012). S'ha observat com els primers residus de la subunitat I_C ja contribueixen en gran manera al plegament de la inteïna. Això concordaria també amb els estudis de Muir i col·laboradors que amb els primers residus de la subunitat I_C també observen ja un cert grau de plegament en la inteïna (Shah et al., 2013b).

Com a perspectives de futur a aquest treball, seria interessant d'entrada estudiar més a fons els mutants C1S i C+1S. En primer lloc caldria veure quin efecte podria tenir la introducció del triplet Trp-Cys-Thr en l'activitat del mutant C1S. Mitjançant mutagènesi dirigida es podrien introduir aquestes mutacions i veure si això afecta a l'activitat del mutant. També seria interessant mitjançant RMN veure si la mutació C1S pot estar afectant significativament l'estructura de la inteïna i que per aquest motiu es perdi l'activitat en aquest mutant. Així es podrien obtenir més coneixements del paper de

la Cys 1 en el mecanisme de tall i unió de la inteïna *Npu DnaE*. Pel que fa al mutant C+1S seria interessant per una banda determinar quins efectes tindrien diferents exteïnes sobre aquesta inteïna mutant en el valor de $t_{1/2}$. A partir d'aquest mutant, es podria posar a punt alguna tècnica de tall i unió condicionat per temperatura i amb diferents exteïnes per tal de determinar si el mutant C+1S de *Npu DnaE* és un bon candidat a ser utilitzat en aplicacions biotecnològiques. Seria convenient també avaluar el paper que tindria el mutant C+1T i en cas que la mutació provoqués uns efectes semblants als de la mutació C+1S, es podria avaluar quina de les dues mutacions és més eficient per tal d'utilitzar els mutants en el camp de la biologia molecular. Totes aquestes mutacions seria interessant i convenient també estudiar-les mitjançant RMN i veure quins canvis conformationals provoquen en la inteïna *Npu DnaE*. Detectar i analitzar els canvis permetria obtenir informació essencial per tal d'entendre millor el paper que tenen cada un dels residus en cada una de les etapes del mecanisme de tall i unió.

Pel que fa a l'estudi del procés associatiu seria interessant, a partir dels coneixements que es tenen i que s'han obtingut en el treball realitzat, veure amb major detall les diferències conformationals de cada una de les variants estudiades per tal de tenir un coneixement més exhaustiu de com el fragment de I_C es troba involucrat en el plegament global de la inteïna. Es podrien realitzar una anàlisi de l'estabilitat tèrmica per tal de veure quins efectes té el fragment I_C en l'estabilitat de la inteïna. Seria convenient veure amb més detall quins són els residus del fragment de la I_C que són més essencials en el procés sinèrgic de plegament i de quina manera aquests residus interaccionen amb els residus del fragment I_N . La millora del coneixement a nivell molecular de com es produeixen aquestes interaccions i quins són els residus responsables, no només permetria una millor comprensió de tot el procés, sinó, una bona base pel disseny i millores de les tècniques de biotecnologia i biologia molecular que utilitzen inteïnes. aquesta classe inusual i cada vegada més útil de proteïnes.

6. Conclusions

6. Conclusions

1. S'ha posat a punt un protocol d'expressió i purificació de *Npu DnaE* activa formant una única cadena polipeptídica (*cis*), per les subunitats I_N i I_C per separat (*trans*) formant un complex i de la subunitat I_N aïllada.
2. L'estabilitat tèrmica de *Npu DnaE* activa formant una única cadena polipeptídica, en forma de complex I_N-I_C i la subunitat I_N aïllada, presenta un màxim a pHs lleugerament bàsics (al voltant de 8). La millor estabilitat es troba quan la inteïna es troba formant una única cadena polipeptídica. Per altra banda, la menor estabilitat tèrmica s'observa per la subunitat I_N aïllada. La correlació entre estabilitat i pH té també un sentit funcional: els valors de pH lleugerament bàsics afavoririen que les Cys 1 i +1 catalítiques es trobessin parcialment ionitzades i actuessin com a nucleòfils, afavorint així l'activitat de la inteïna.
3. El replegament de la inteïna es veu condicionat per com es troben associades les seves subunitats. Quan formen una única cadena polipeptídica s'observa un millor percentatge de replegament que no pas quan es troben associades formant un complex. El replegament de la subunitat I_N aïllada és nul. Aquestes observacions estarien d'acord amb un procés cooperatiu en el plegament d'ambdues subunitats de la inteïna per assolir l'estat natiu.
4. La mutació C+1S disminueix de manera molt significativa l'activitat de la inteïna a temperatures baixes tant quan es troba en *cis* com en *trans*. L'activitat en aquests mutants es recupera al incrementar la temperatura i millora considerablement a partir dels 40°C. En el cas del mutant C+1S en *trans*, malgrat presentar un $t_{1/2}$ inferior al que s'observa en el cas de la inteïna salvatge, la formació del producte de tall i unió format a les 24h és el mateix a 50 i 60°C.
5. El mutant C1S i el doble mutant C1S-C+1S disminueixen de manera molt significativa l'activitat de la inteïna en *trans* a temperatures baixes. Aquesta activitat no millora malgrat s'incrementi la temperatura. Aquests resultats indiquen que la C1 és més crítica que no pas la C+1 en el mecanisme del procés de tall i unió d'aquesta inteïna. La mutació C+1S podria ser utilitzada en aplicacions basades en el tall i unió condicionat per temperatura, sempre i quan l'exteïna presentés una bona estabilitat tèrmica i amb l'avantatge d'eliminar una Cys en l'exteïna C i per tant en el producte final del tall i unió.
6. Els resultats del dicromisme circular han mostrat que el plegament global de la inteïna està relacionat amb la longitud de la I_C unida a la I_N . S'ha observat que la subunitat I_N presenta un nivell molt baix d'estructura secundària i que aquesta va augmentant a mesura que s'incrementa la longitud de subunitat I_C . El dicromisme circular ha permès comprovar també

6. Conclusions

que les variants I_NI_C_sc114, I_NI_C_sc123 i I_NI_C_sc131 presenten una estructura secundària principalment formada per cadenes β , fet que concorda amb el plegament de la inteïna completa *Npu DnaE*.

7. Els espectres monodimensionals de ¹H han permès fer una primera observació del grau d'estructuració de cada una de les variants. En els espectres obtinguts s'ha observat com la subunitat I_N i la variant I_NI_C_sc114 presenten una estructura de glòbul fos i que, s'adopta una estructura més estable i ordenada en les variants I_NI_C_sc123 i I_NI_C_sc131.
8. A partir de l'assignació de NOE's dels espectres bidimensionals NOESY i el càlcul dels desplaçaments químics del protons s'ha observat que la majoria de residus de les variants I_NI_C_sc123 i I_NI_C_sc131 es troben en elements d'estructura secundària seguint un patró molt similar al que presenten els residus en la inteïna completa.
9. La visualització, de tots els NOE's de mitja i llarga distància assignats, en models tridimensionals ha estat essencial per veure quines regions d'aquestes variants eren les més desestructurades i diferents respecte la inteïna completa. Concretament les regions més desestructurades han estat les que es trobaven més pròximes a la regió de la part de la subunitat I_C absent a cada una de les variants.
10. Els NOE's NH-NH i NH-HA han permès determinar l'orientació de les cadenes β de cada una de les variants i observar que l'orientació de les cadenes β de cada una de les variants es conserva respecte a la inteïna completa. A més, s'ha pogut veure de manera molt clara, a cada una de les variants, com es forma la cadena β 11 formada per residus de la subunitat I_C. Aquesta observació ha permès concloure que els residus que constitueixen β 11 es pleguen de manera independent a la resta de residus de la subunitat I_C.
11. El grau d'estructura observat a I_NI_C_sc123 i I_NI_C_sc131 és pràcticament el mateix. Els residus que té de més I_NI_C_sc131 són els que constitueixen β 13 i aquests estableixen de manera significativa moltes interaccions amb els residus que constitueixen β 14 absents en aquesta variant. Per aquest motiu, els residus de β 13 no han pogut formar les interaccions necessàries i es troben desordenats en I_NI_C_sc131.
12. Com a conclusió global a aquest treball, s'ha obtingut un millor coneixement de la inteïna *Npu DnaE*, s'ha comprovat que és una inteïna robusta, que presenta una bona termoestabilitat a pHs lleugerament bàsics i que les mutacions C+1S i C1S disminueixen de manera significativa la seva activitat. Aquesta disminució però, és reversible mitjançant temperatura en el cas del mutant senzill C+1S. Per altra banda, s'ha pogut corroborar la

6. Conclusions

sinèrgia en el plegament que presenten les subunitats I_N i I_C i com els primers residus de la subunitat I_C contribueixen de forma crítica en el plegament global de *Npu DnaE*.

7. Bibliografia

7. Bibliografia

- Abel-Santos, E., Scott, C.P., and Benkovic, S.J. (2003). Use of inteins for the in vivo production of stable cyclic peptide libraries in *E. coli*. *Methods in molecular biology* 205, 281-294.
- Albertsen, L., Shaw, A.C., Norrild, J.C., and Stromgaard, K. (2013). Recombinant production of peptide C-terminal alpha-amides using an engineered intein. *Bioconjugate chemistry* 24, 1883-1894.
- Alcock, N.W. (1990). *Bonding and Structure: structural principles in inorganic and organic chemistry*. Ellis Horwood Ltd, 40-42.
- Amitai, G., Belenkiy, O., Dassa, B., Shainskaya, A., and Pietrokovski, S. (2003). Distribution and function of new bacterial intein-like protein domains. *Molecular microbiology* 47, 61-73.
- Amitai, G., Callahan, B.P., Stanger, M.J., Belfort, G., and Belfort, M. (2009). Modulation of intein activity by its neighboring extein substrates. *Proceedings of the National Academy of Sciences of the United States of America* 106, 11005-11010.
- Amitai, G., Dassa, B., and Pietrokovski, S. (2004). Protein splicing of inteins with atypical glutamine and aspartate C-terminal residues. *The Journal of biological chemistry* 279, 3121-3131.
- Anraku, Y., Mizutani, R., and Satow, Y. (2005). Protein splicing: its discovery and structural insight into novel chemical mechanisms. *IUBMB life* 57, 563-574.
- Appleby-Tagoe, J.H., Thiel, I.V., Wang, Y., Wang, Y., Mootz, H.D., and Liu, X.Q. (2011). Highly efficient and more general cis- and trans-splicing inteins through sequential directed evolution. *The Journal of biological chemistry* 286, 34440-34447.
- Appleby, J.H., Zhou, K., Volkmann, G., and Liu, X.Q. (2009). Novel split intein for trans-splicing synthetic peptide onto C terminus of protein. *The Journal of biological chemistry* 284, 6194-6199.
- Aranko, A.S., Oeemig, J.S., Kajander, T., and Iwai, H. (2013). Intermolecular domain swapping induces intein-mediated protein alternative splicing. *Nature chemical biology* 9, 616-622.
- Aranko, A.S., Oeemig, J.S., Zhou, D., Kajander, T., Wlodawer, A., and Iwai, H. (2014). Structure-based engineering and comparison of novel split inteins for protein ligation. *Molecular bioSystems*.
- Aranko, A.S., Zuger, S., Buchinger, E., and Iwai, H. (2009). In vivo and in vitro protein ligation by naturally occurring and engineered split DnaE inteins. *PLoS one* 4, e5185.
- Artimo, P., Jonnalagedda, M., Arnold, K., Baratin, D., Csardi, G., de Castro, E., Duvaud, S., Flegel, V., Fortier, A., Gasteiger, E., et al. (2012). ExpASY: SIB bioinformatics resource portal. *Nucleic acids research* 40, W597-603.
- Bakhrat, A., Jurica, M.S., Stoddard, B.L., and Raveh, D. (2004). Homology modeling and mutational analysis of Ho endonuclease of yeast. *Genetics* 166, 721-728.
- Barron, A., Smith, C., and Hamilton, C. (2010). *Chemistry of the Main Group Elements* Open Educational Resource.
- Bellmann-Sickert, K., and Beck-Sickinger, A.G. (2010). Peptide drugs to target G protein-coupled receptors. *Trends in pharmacological sciences* 31, 434-441.
- Berrade, L., and Camarero, J.A. (2009). Expressed protein ligation: a resourceful tool to study protein structure and function. *Cellular and molecular life sciences : CMLS* 66, 3909-3922.
- Berrade, L., Kwon, Y., and Camarero, J.A. (2010). Photomodulation of protein trans-splicing through backbone photocaging of the DnaE split intein. *Chembiochem : a European journal of chemical biology* 11, 1368-1372.

7. Bibliografía

Better, M., Chang, C.P., Robinson, R.R., and Horwitz, A.H. (1988). Escherichia coli secretion of an active chimeric antibody fragment. *Science* 240, 1041-1043.

Binschik, J., and Mootz, H.D. (2013). Chemical bypass of intein-catalyzed N-S acyl shift in protein splicing. *Angewandte Chemie* 52, 4260-4264.

Binschik, J., Zettler, J., and Mootz, H.D. (2011). Photocontrol of protein activity mediated by the cleavage reaction of a split intein. *Angewandte Chemie* 50, 3249-3252.

Brace, L.E., Southworth, M.W., Tori, K., Cushing, M.L., and Perler, F. (2010). The *Deinococcus radiodurans* Snf2 intein caught in the act: detection of the Class 3 intein signature Block F branched intermediate. *Protein science : a publication of the Protein Society* 19, 1525-1533.

Brenzel, S., Kurpiers, T., and Mootz, H.D. (2006). Engineering artificially split inteins for applications in protein chemistry: biochemical characterization of the split Ssp DnaB intein and comparison to the split Sce VMA intein. *Biochemistry* 45, 1571-1578.

Busche, A.E., Aranko, A.S., Talebzadeh-Farooji, M., Bernhard, F., Dotsch, V., and Iwai, H. (2009). Segmental isotopic labeling of a central domain in a multidomain protein by protein trans-splicing using only one robust DnaE intein. *Angewandte Chemie* 48, 6128-6131.

Calderone, C.T., and Williams, D.H. (2001). An enthalpic component in cooperativity: the relationship between enthalpy, entropy, and noncovalent structure in weak associations. *Journal of the American Chemical Society* 123, 6262-6267.

Callahan, B.P., Stanger, M., and Belfort, M. (2013). A redox trap to augment the intein toolbox. *Biotechnology and bioengineering* 110, 1565-1573.

Callahan, B.P., Topilina, N.I., Stanger, M.J., Van Roey, P., and Belfort, M. (2011). Structure of catalytically competent intein caught in a redox trap with functional and evolutionary implications. *Nature structural & molecular biology* 18, 630-633.

Cambon-Bonavita, M.A., Schmitt, P., Zieger, M., Flaman, J.M., Lesongeur, F., Raguene, G., Bindel, D., Frisch, N., Lakkis, Z., Dupret, D., et al. (2000). Cloning, expression, and characterization of DNA polymerase I from the hyperthermophilic archaea *Thermococcus fumicolans*. *Extremophiles : life under extreme conditions* 4, 215-225.

Carrington, J.C., and Dougherty, W.G. (1988). A viral cleavage site cassette: identification of amino acid sequences required for tobacco etch virus polyprotein processing. *Proceedings of the National Academy of Sciences of the United States of America* 85, 3391-3395.

Carvajal-Vallejos, P., Pallisse, R., Mootz, H.D., and Schmidt, S.R. (2012). Unprecedented rates and efficiencies revealed for new natural split inteins from metagenomic sources. *The Journal of biological chemistry* 287, 28686-28696.

Caspi, J., Amitai, G., Belenkiy, O., and Pietrokovski, S. (2003). Distribution of split DnaE inteins in cyanobacteria. *Molecular microbiology* 50, 1569-1577.

Claus, J., and Winyard, P. (2009). Redox Signaling and Regulation in Biology and Medicine. *Curr Opin Chem Biol* 1, 1-12.

Contreras Martinez, L.M., Borrero Quintana, E.E., Escobedo, F.A., and DeLisa, M.P. (2008). In silico protein fragmentation reveals the importance of critical nuclei on domain reassembly. *Biophysical journal* 94, 1575-1588.

7. Bibliografia

- Chen, J., Zheng, X.F., Brown, E.J., and Schreiber, S.L. (1995). Identification of an 11-kDa FKBP12-rapamycin-binding domain within the 289-kDa FKBP12-rapamycin-associated protein and characterization of a critical serine residue. *Proceedings of the National Academy of Sciences of the United States of America* 92, 4947-4951.
- Chen, L., Benner, J., and Perler, F.B. (2000). Protein splicing in the absence of an intein penultimate histidine. *The Journal of biological chemistry* 275, 20431-20435.
- Chen, W., Li, L., Du, Z., Liu, J., Reitter, J.N., Mills, K.V., Linhardt, R.J., and Wang, C. (2012). Intramolecular disulfide bond between catalytic cysteines in an intein precursor. *Journal of the American Chemical Society* 134, 2500-2503.
- Cheriyian, M., Pedamallu, C.S., Tori, K., and Perler, F. (2013). Faster protein splicing with the Nostoc punctiforme DnaE intein using non-native extein residues. *The Journal of biological chemistry* 288, 6202-6211.
- Choi, J.J., Nam, K.H., Min, B., Kim, S.J., Soll, D., and Kwon, S.T. (2006). Protein trans-splicing and characterization of a split family B-type DNA polymerase from the hyperthermophilic archaeal parasite Nanoarchaeum equitans. *Journal of molecular biology* 356, 1093-1106.
- Daniel, R.M., Smith, J.C., Ferrand, M., Hery, S., Dunn, R., and Finney, J.L. (1998). Enzyme activity below the dynamical transition at 220 K. *Biophysical journal* 75, 2504-2507.
- Dassa, B., Amitai, G., Caspi, J., Schueler-Furman, O., and Pietrokovski, S. (2007). Trans protein splicing of cyanobacterial split inteins in endogenous and exogenous combinations. *Biochemistry* 46, 322-330.
- Dassa, B., London, N., Stoddard, B.L., Schueler-Furman, O., and Pietrokovski, S. (2009). Fractured genes: a novel genomic arrangement involving new split inteins and a new homing endonuclease family. *Nucleic acids research* 37, 2560-2573.
- Dassa, B., and Pietrokovski, S. (2005). Origin and Evolution of Inteins and Other Hint Domains. *Nucleic Acids and Molecular Biology* 16, 211-231.
- Deleage, G., and Geourjon, C. (1993). An interactive graphic program for calculating the secondary structure content of proteins from circular dichroism spectrum. *Computer applications in the biosciences : CABIOS* 9, 197-199.
- Dhar, T., and Mootz, H.D. (2011). Modification of transmembrane and GPI-anchored proteins on living cells by efficient protein trans-splicing using the Npu DnaE intein. *Chemical communications* 47, 3063-3065.
- Ding, Y., Xu, M., Ghosh, I., Chen, X., Ferrandon, S., Lesage, G., and Rao, Z. (2003). Crystal structure of a mini-intein reveals a conserved catalytic module involved in side chain cyclization of asparagine during protein splicing. *J Biol Chem* 278, 39133-39142.
- Du, Z., Liu, J., Albracht, C.D., Hsu, A., Chen, W., Marieni, M.D., Colelli, K.M., Williams, J.E., Reitter, J.N., Mills, K.V., et al. (2011a). Structural and mutational studies of a hyperthermophilic intein from DNA polymerase II of *Pyrococcus abyssi*. *The Journal of biological chemistry* 286, 38638-38648.
- Du, Z., Liu, Y., Ban, D., Lopez, M.M., Belfort, M., and Wang, C. (2010). Backbone dynamics and global effects of an activating mutation in minimized Mtu RecA inteins. *Journal of molecular biology* 400, 755-767.
- Du, Z., Shemella, P., Liu, Y., McCallum, S., Pereira, B., Nayak, S., Belfort, G., Belfort, M., and Wang, C. (2009). Highly conserved histidine plays a dual catalytic role in protein splicing: a pKa shift mechanism. *Journal of the American Chemical Society* 131, 11581-11589.

7. Bibliografía

- Du, Z., Zheng, Y., Patterson, M., Liu, Y., and Wang, C. (2011b). pK(a) coupling at the intein active site: implications for the coordination mechanism of protein splicing with a conserved aspartate. *Journal of the American Chemical Society* 133, 10275-10282.
- Duan, X., Gimble, F.S., and Quioco, F.A. (1997). Crystal structure of PI-Scel, a homing endonuclease with protein splicing activity. *Cell* 89, 555-564.
- Dunker, A.K., Lawson, J.D., Brown, C.J., Williams, R.M., Romero, P., Oh, J.S., Oldfield, C.J., Campen, A.M., Ratliff, C.M., Hipps, K.W., et al. (2001). Intrinsically disordered protein. *Journal of molecular graphics & modelling* 19, 26-59.
- Duplay, P., Bedouelle, H., Fowler, A., Zabin, I., Saurin, W., and Hofnung, M. (1984). Sequences of the malE gene and of its product, the maltose-binding protein of Escherichia coli K12. *The Journal of biological chemistry* 259, 10606-10613.
- Elleuche, S., and Poggeler, S. (2007). Trans-splicing of an artificially split fungal mini-intein. *Biochemical and biophysical research communications* 355, 830-834.
- Elleuche, S., and Poggeler, S. (2010). Inteins, valuable genetic elements in molecular biology and biotechnology. *Applied microbiology and biotechnology* 87, 479-489.
- Ellila, S., Jurvansuu, J.M., and Iwai, H. (2011). Evaluation and comparison of protein splicing by exogenous inteins with foreign exteins in Escherichia coli. *FEBS letters* 585, 3471-3477.
- Evans, T.J.T., and Xu, M.Q. (2002). Mechanistic and kinetic considerations of protein splicing. *Chemical reviews* 102, 4869-4884.
- Farfan-Arribas, D.J., Stern, L.J., and Rock, K.L. (2012). Using intein catalysis to probe the origin of major histocompatibility complex class I-presented peptides. *Proceedings of the National Academy of Sciences of the United States of America* 109, 16998-17003.
- Feller, G. (2010). Protein stability and enzyme activity at extreme biological temperatures. *Journal of physics Condensed matter : an Institute of Physics journal* 22, 323101.
- Flavell, R.R., and Muir, T.W. (2009). Expressed protein ligation (EPL) in the study of signal transduction, ion conduction, and chromatin biology. *Accounts of chemical research* 42, 107-116.
- Font, J., Benito, A., Torrent, J., Lange, R., Ribo, M., and Vilanova, M. (2006). Pressure- and temperature-induced unfolding studies: thermodynamics of core hydrophobicity and packing of ribonuclease A. *Biological chemistry* 387, 285-296.
- Frokjaer, S., and Otzen, D.E. (2005). Protein drug stability: a formulation challenge. *Nature reviews Drug discovery* 4, 298-306.
- Frutos, S., Goger, M., Giovani, B., Cowburn, D., and Muir, T. (2010). Branched intermediate formation stimulates peptide bond cleavage in protein splicing. *Nat Chem Biol* 6, 527-533.
- Ghosh, I., Sun, L., and Xu, M. (2001). Zinc inhibition of protein trans-splicing and identification of regions essential for splicing and association of a split intein. *J Biol Chem* 276, 24051-24058.
- Gierach, I., Li, J., Wu, W.Y., Grover, G.J., and Wood, D.W. (2012). Bacterial biosensors for screening isoform-selective ligands for human thyroid receptors alpha-1 and beta-1. *FEBS open bio* 2, 247-253.
- Gogarten, J.P., Senejani, A.G., Zhaxybayeva, O., Olendzenski, L., and Hilario, E. (2002). Inteins: structure, function, and evolution. *Annual review of microbiology* 56, 263-287.

7. Bibliografia

- Greenfield, N., and Fasman, G.D. (1969). Computed circular dichroism spectra for the evaluation of protein conformation. *Biochemistry* 8, 4108-4116.
- Greenwood, H.C., Bloom, S.R., and Murphy, K.G. (2011). Peptides and their potential role in the treatment of diabetes and obesity. *The review of diabetic studies : RDS* 8, 355-368.
- Guan, D., Ramirez, M., and Chen, Z. (2013). Split intein mediated ultra-rapid purification of tagless protein (SIRP). *Biotechnology and bioengineering* 110, 2471-2481.
- Guex, N., and Peitsch, M.C. (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* 18, 2714-2723.
- Guhan, N., and Muniyappa, K. (2003). Mycobacterium tuberculosis RecA intein, a LAGLIDADG homing endonuclease, displays Mn(2+) and DNA-dependent ATPase activity. *Nucleic acids research* 31, 4184-4191.
- Guntert, P., Mumenthaler, C., and Wuthrich, K. (1997). Torsion angle dynamics for NMR structure calculation with the new program DYANA. *Journal of molecular biology* 273, 283-298.
- Gururaja, T.L., Narasimhamurthy, S., Payan, D.G., and Anderson, D.C. (2000). A novel artificial loop scaffold for the noncovalent constraint of peptides. *Chemistry & biology* 7, 515-527.
- Hall, T.M., Porter, J.A., Young, K.E., Koonin, E.V., Beachy, P.A., and Leahy, D.J. (1997). Crystal structure of a Hedgehog autoprocessing domain: homology between Hedgehog and self-splicing proteins. *Cell* 91, 85-97.
- Hanahan, D. (1983). Studies on transformation of Escherichia coli with plasmids. *Journal of molecular biology* 166, 557-580.
- He, X., Chow, D., Martick, M.M., and Garcia, K.C. (2001). Allosteric activation of a spring-loaded natriuretic peptide receptor dimer by hormone. *Science* 293, 1657-1662.
- Heinamaki, K., Oemig, J.S., Paakkonen, K., Djupsjobacka, J., and Iwai, H. (2009). NMR resonance assignment of DnaE intein from Nostoc punctiforme. *Biomolecular NMR assignments* 3, 41-43.
- Hirata, R., Ohsumi, Y., Nakano, A., Kawasaki, H., Suzuki, K., and Anraku, Y. (1990). Molecular structure of a gene, VMA1, encoding the catalytic subunit of H(+)-translocating adenosine triphosphatase from vacuolar membranes of Saccharomyces cerevisiae. *The Journal of biological chemistry* 265, 6726-6733.
- Holzwarth, G., and Doty, P. (1965). The Ultraviolet Circular Dichroism of Polypeptides. *Journal of the American Chemical Society* 87, 218-228.
- Huang, X., Narayanaswamy, R., Fenn, K., Szpakowski, S., Sasaki, C., Costa, J., Blancafort, P., and Lizardi, P.M. (2012). Sequence-specific biosensors report drug-induced changes in epigenetic silencing in living cells. *DNA and cell biology* 31 Suppl 1, S2-10.
- Iwai, H., Zuger, S., Jin, J., and Tam, P.H. (2006). Highly efficient protein trans-splicing by a naturally split DnaE intein from Nostoc punctiforme. *FEBS letters* 580, 1853-1858.
- Jeffries, C.M., Graham, S.C., Stokes, P.H., Collyer, C.A., Guss, J.M., and Matthews, J.M. (2006). Stabilization of a binary protein complex by intein-mediated cyclization. *Protein science : a publication of the Protein Society* 15, 2612-2618.
- Jermutus, L., Honegger, A., Schwesinger, F., Hanes, J., and Pluckthun, A. (2001). Tailoring in vitro evolution for protein affinity or stability. *Proceedings of the National Academy of Sciences of the United States of America* 98, 75-80.

7. Bibliografia

- Johnson, M.A., Southworth, M.W., Herrmann, T., Brace, L., Perler, F.B., and Wuthrich, K. (2007). NMR structure of a KlbA intein precursor from *Methanococcus jannaschii*. *Protein science : a publication of the Protein Society* 16, 1316-1328.
- Jung, D., Min, K., Jung, J., Jang, W., and Kwon, Y. (2013). Chemical biology-based approaches on fluorescent labeling of proteins in live cells. *Molecular bioSystems* 9, 862-872.
- Kane, P.M., Yamashiro, C.T., Wolczyk, D.F., Neff, N., Goebel, M., and Stevens, T.H. (1990). Protein splicing converts the yeast TFP1 gene product to the 69-kD subunit of the vacuolar H(+)-adenosine triphosphatase. *Science* 250, 651-657.
- Kanno, A., Ozawa, T., and Umezawa, Y. (2009). Bioluminescent imaging of MAPK function with intein-mediated reporter gene assay. *Methods in molecular biology* 574, 185-192.
- Kanno, A., Yamanaka, Y., Hirano, H., Umezawa, Y., and Ozawa, T. (2007). Cyclic luciferase for real-time sensing of caspase-3 activities in living mammals. *Angewandte Chemie* 46, 7595-7599.
- Kapust, R.B., and Waugh, D.S. (1999). *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein science : a publication of the Protein Society* 8, 1668-1674.
- Karra, E., Chandarana, K., and Batterham, R.L. (2009). The role of peptide YY in appetite regulation and obesity. *The Journal of physiology* 587, 19-25.
- Kawasaki, M., Nogami, S., Satow, Y., Ohya, Y., and Anraku, Y. (1997). Identification of three core regions essential for protein splicing of the yeast Vma1 protozyme. A random mutagenesis study of the entire Vma1-derived endonuclease sequence. *The Journal of biological chemistry* 272, 15668-15674.
- Kay, L.E., Torchia, D.A., and Bax, A. (1989). Backbone dynamics of proteins as studied by 15N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry* 28, 8972-8979.
- Kenneth, V., and Paulus, H. (2005). Biochemical Mechanisms of Intein-Mediated Protein Splicing. *Nucleic Acids and Molecular Biology* 16, 233-255.
- Kim, S.B., Ozawa, T., Watanabe, S., and Umezawa, Y. (2004). High-throughput sensing and noninvasive imaging of protein nuclear transport by using reconstitution of split Renilla luciferase. *Proceedings of the National Academy of Sciences of the United States of America* 101, 11542-11547.
- Kjaergaard, M., Brander, S., and Poulsen, F.M. (2011). Random coil chemical shift for intrinsically disordered proteins: effects of temperature and pH. *Journal of biomolecular NMR* 49, 139-149.
- Kjaergaard, M., and Poulsen, F.M. (2011). Sequence correction of random coil chemical shifts: correlation between neighbor correction factors and changes in the Ramachandran distribution. *Journal of biomolecular NMR* 50, 157-165.
- Kokkinidis, M., Glykos, N.M., and Fadoulglou, V.E. (2012). Protein flexibility and enzymatic catalysis. *Advances in protein chemistry and structural biology* 87, 181-218.
- Kritzer, J.A., Hamamichi, S., McCaffery, J.M., Santagata, S., Naumann, T.A., Caldwell, K.A., Caldwell, G.A., and Lindquist, S. (2009). Rapid selection of cyclic peptides that reduce alpha-synuclein toxicity in yeast and animal models. *Nature chemical biology* 5, 655-663.
- Kurpiers, T., and Mootz, H.D. (2008). Site-specific chemical modification of proteins with a prelabelled cysteine tag using the artificially split Mxe GyrA intein. *ChemBiochem : a European journal of chemical biology* 9, 2317-2325.

7. Bibliografia

Lane, D., and Harlow, E. (1982). Two different viral transforming proteins bind the same host tumour antigen. *Nature* 298, 517.

Lange, R., J. Frank, J. -L. Saldana, and Balny, C. (1996). Fourth derivative UV-spectroscopy of proteins under high pressure I. Factors affecting the fourth derivative spectrum of the aromatic amino acids. *European Biophysics Journal* 24, 277-283.

Lee, Y.T., Su, T.H., Lo, W.C., Lyu, P.C., and Sue, S.C. (2012). Circular permutation prediction reveals a viable backbone disconnection for split proteins: an approach in identifying a new functional split intein. *PloS one* 7, e43820.

Li, J., Gierach, I., Gillies, A.R., Warden, C.D., and Wood, D.W. (2011). Engineering and optimization of an allosteric biosensor protein for peroxisome proliferator-activated receptor gamma ligands. *Biosensors & bioelectronics* 29, 132-139.

Li, X., Romero, P., Rani, M., Dunker, A.K., and Obradovic, Z. (1999). Predicting Protein Disorder for N-, C-, and Internal Regions. *Genome informatics Workshop on Genome Informatics* 10, 30-40.

Lindman, S., Johansson, I., Thulin, E., and Linse, S. (2009). Green fluorescence induced by EF-hand assembly in a split GFP system. *Protein science : a publication of the Protein Society* 18, 1221-1229.

Liu, X.Q. (2000). Protein-splicing intein: Genetic mobility, origin, and evolution. *Annual review of genetics* 34, 61-76.

Liu, X.Q., and Yang, J. (2003). Split dnaE genes encoding multiple novel inteins in *Trichodesmium erythraeum*. *The Journal of biological chemistry* 278, 26315-26318.

Lockless, S.W., and Muir, T.W. (2009). Traceless protein splicing utilizing evolved split inteins. *Proceedings of the National Academy of Sciences of the United States of America* 106, 10999-11004.

Louis-Jeune, C., Andrade-Navarro, M.A., and Perez-Iratxeta, C. (2011). Prediction of protein secondary structure from circular dichroism using theoretically derived spectra. *Proteins*.

Ludwig, C., Pfeiff, M., Linne, U., and Mootz, H.D. (2006). Ligation of a synthetic peptide to the N terminus of a recombinant protein using semisynthetic protein trans-splicing. *Angewandte Chemie* 45, 5218-5221.

Luginbuhl, P., Pervushin, K.V., Iwai, H., and Wuthrich, K. (1997). Anisotropic molecular rotational diffusion in ¹⁵N spin relaxation studies of protein mobility. *Biochemistry* 36, 7305-7312.

Martin, D.D., Xu, M.Q., and Evans, T.C., Jr. (2001). Characterization of a naturally occurring trans-splicing intein from *Synechocystis* sp. PCC6803. *Biochemistry* 40, 1393-1402.

Massoud, T.F., Paulmurugan, R., and Gambhir, S.S. (2010). A molecularly engineered split reporter for imaging protein-protein interactions with positron emission tomography. *Nature medicine* 16, 921-926.

Matsumura, H., Takahashi, H., Inoue, T., Yamamoto, T., Hashimoto, H., Nishioka, M., Fujiwara, S., Takagi, M., Imanaka, T., and Kai, Y. (2006). Crystal structure of intein homing endonuclease II encoded in DNA polymerase gene from hyperthermophilic archaeon *Thermococcus kodakaraensis* strain KOD1. *Proteins* 63, 711-715.

Merkler, D.J. (1994). C-terminal amidated peptides: production by the in vitro enzymatic amidation of glycine-extended peptides and the importance of the amide to bioactivity. *Enzyme and microbial technology* 16, 450-456.

7. Bibliografia

- Miao, J., Wu, W., Spielmann, T., Belfort, M., Derbyshire, V., and Belfort, G. (2005). Single-step affinity purification of toxic and non-toxic proteins on a fluidics platform. *Lab on a chip* 5, 248-253.
- Mills, K.V., Dorval, D.M., and Lewandowski, K.T. (2005). Kinetic analysis of the individual steps of protein splicing for the *Pyrococcus abyssi* PolII intein. *The Journal of biological chemistry* 280, 2714-2720.
- Mootz, H.D. (2009). Split inteins as versatile tools for protein semisynthesis. *Chembiochem : a European journal of chemical biology* 10, 2579-2589.
- Mozhaev, V.V., Heremans, K., Frank, J., Masson, P., and Balny, C. (1996). High pressure effects on protein structure and function. *Proteins* 24, 81-91.
- Muir, T.W., Sondhi, D., and Cole, P.A. (1998). Expressed protein ligation: a general method for protein engineering. *Proceedings of the National Academy of Sciences of the United States of America* 95, 6705-6710.
- Mujika, J., Lopez, X., and Mulholland, A. (2009). Modeling protein splicing: reaction pathway for C-terminal splice and intein scission. *J Phys Chem B* 113, 5607-5616.
- Mujika, J.I., Lopez, X., and Mulholland, A.J. (2012). Mechanism of C-terminal intein cleavage in protein splicing from QM/MM molecular dynamics simulations. *Organic & biomolecular chemistry* 10, 1207-1218.
- Muona, M., Aranko, A.S., Raulinaitis, V., and Iwai, H. (2010). Segmental isotopic labeling of multi-domain and fusion proteins by protein trans-splicing in vivo and in vitro. *Nature protocols* 5, 574-587.
- Nicastri, M.C., Xega, K., Li, L., Xie, J., Wang, C., Linhardt, R.J., Reitter, J.N., and Mills, K.V. (2013). Internal disulfide bond acts as a switch for intein activity. *Biochemistry* 52, 5920-5927.
- Nichols, N.M., Benner, J.S., Martin, D.D., and Evans, T.C., Jr. (2003). Zinc ion effects on individual Ssp DnaE intein splicing steps: regulating pathway progression. *Biochemistry* 42, 5301-5311.
- Nichols, N.M., and Evans, T.C., Jr. (2004). Mutational analysis of protein splicing, cleavage, and self-association reactions mediated by the naturally split Ssp DnaE intein. *Biochemistry* 43, 10265-10276.
- Norma, J.G. (2006). Using circular dichroism spectra to estimate protein secondary structure. *Nat Protoc* 1, 2876-2890.
- Oeemig, J.S., Aranko, A.S., Djupsjobacka, J., Heinamaki, K., and Iwai, H. (2009). Solution structure of DnaE intein from *Nostoc punctiforme*: structural basis for the design of a new split intein suitable for site-specific chemical modification. *FEBS letters* 583, 1451-1456.
- Ozawa, T., Nogami, S., Sato, M., Ohya, Y., and Umezawa, Y. (2000). A fluorescent indicator for detecting protein-protein interactions in vivo based on protein splicing. *Analytical chemistry* 72, 5151-5157.
- Pace, C.N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995). How to measure and predict the molar absorption coefficient of a protein. *Protein science : a publication of the Protein Society* 4, 2411-2423.
- Paramban, R.I., Bugos, R.C., and Su, W.W. (2004). Engineering green fluorescent protein as a dual functional tag. *Biotechnology and bioengineering* 86, 687-697.
- Paulsen, C.E., and Carroll, K.S. (2013). Cysteine-mediated redox signaling: chemistry, biology, and tools for discovery. *Chemical reviews* 113, 4633-4679.

7. Bibliografia

- Paulus, H. (1998). The chemical basis of protein splicing. *Chem Soc Rev* 27, 375 - 386.
- Pearl, E.J., Tyndall, J.D., Poulter, R.T., and Wilbanks, S.M. (2007). Sequence requirements for splicing by the Cne PRP8 intein. *FEBS letters* 581, 3000-3004.
- Pereira, B., Shemella, P.T., Amitai, G., Belfort, G., Nayak, S.K., and Belfort, M. (2011). Spontaneous proton transfer to a conserved intein residue determines on-pathway protein splicing. *Journal of molecular biology* 406, 430-442.
- Perler, F.B. (1998). Protein splicing of inteins and hedgehog autoproteolysis: structure, function, and evolution. *Cell* 92, 1-4.
- Perler, F.B. (2002). InBase: the InteIn Database. *Nucleic acids research* 30, 383-384.
- Perler, F.B. (2005). Protein splicing mechanisms and applications. *IUBMB life* 57, 469-476.
- Perler, F.B., Davis, E.O., Dean, G.E., Gimble, F.S., Jack, W.E., Neff, N., Noren, C.J., Thorner, J., and Belfort, M. (1994). Protein splicing elements: inteins and exteins--a definition of terms and recommended nomenclature. *Nucleic acids research* 22, 1125-1127.
- Perler, F.B., Olsen, G.J., and Adam, E. (1997). Compilation and analysis of intein sequences. *Nucleic acids research* 25, 1087-1093.
- Pietrokovski, S. (1998). Modular organization of inteins and C-terminal autocatalytic domains. *Protein science : a publication of the Protein Society* 7, 64-71.
- Pietrokovski, S. (2001). InteIn spread and extinction in evolution. *Trends in genetics : TIG* 17, 465-472.
- Qi, X., Meng, Q., and Liu, X.Q. (2011). Spontaneous C-cleavage of a mini-intein without its conserved N-terminal motif A. *FEBS letters* 585, 2513-2518.
- Ramirez, M., Valdes, N., Guan, D., and Chen, Z. (2013). Engineering split intein DnaE from *Nostoc punctiforme* for rapid protein purification. *Protein engineering, design & selection : PEDS* 26, 215-223.
- Reitter, J.N., and Mills, K.V. (2011). Canonical protein splicing of a class 1 intein that has a class 3 noncanonical sequence motif. *Journal of bacteriology* 193, 994-997.
- Romanelli, A., Shekhtman, A., Cowburn, D., and Muir, T.W. (2004). Semisynthesis of a segmental isotopically labeled protein splicing precursor: NMR evidence for an unusual peptide bond at the N-extein-intein junction. *Proceedings of the National Academy of Sciences of the United States of America* 101, 6397-6402.
- Romero, Obradovic, and Dunker, K. (1997). Sequence Data Analysis for Long Disordered Regions Prediction in the Calcineurin Family. *Genome informatics Workshop on Genome Informatics* 8, 110-124.
- Romero, P., Obradovic, Z., Li, X., Garner, E.C., Brown, C.J., and Dunker, A.K. (2001). Sequence complexity of disordered protein. *Proteins* 42, 38-48.
- Sakamoto, S., Terauchi, M., Hugo, A., Kim, T., Araki, Y., and Wada, T. (2013). Creation of a caspase-3 sensing system using a combination of split-GFP and split-intein. *Chemical communications* 49, 10323-10325.
- Saleh, L., and Perler, F.B. (2006). Protein splicing in cis and in trans. *Chemical record* 6, 183-193.

7. Bibliografia

Saleh, L., Southworth, M.W., Considine, N., O'Neill, C., Benner, J., Bollinger, J.M., Jr., and Perler, F.B. (2011). Branched intermediate formation is the slowest step in the protein splicing reaction of the Ala1 KlbA intein from *Methanococcus jannaschii*. *Biochemistry* 50, 10576-10589.

Sanger, F., and Coulson, A.R. (1978). The use of thin acrylamide gels for DNA sequencing. *FEBS letters* 87, 107-110.

Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America* 74, 5463-5467.

Schoepfer, R. (1993). The pRSET family of T7 promoter expression vectors for *Escherichia coli*. *Gene* 124, 83-85.

Schreiber, G., and Fersht, A.R. (1996). Rapid, electrostatically assisted association of proteins. *Nature structural biology* 3, 427-431.

Selgrade, D.F., Lohmueller, J.J., Lienert, F., and Silver, P.A. (2013). Protein scaffold-activated protein trans-splicing in mammalian cells. *Journal of the American Chemical Society* 135, 7713-7719.

Shah, N.H., Dann, G.P., Vila-Perello, M., Liu, Z., and Muir, T.W. (2012). Ultrafast protein splicing is common among cyanobacterial split inteins: implications for protein engineering. *Journal of the American Chemical Society* 134, 11338-11341.

Shah, N.H., Eryilmaz, E., Cowburn, D., and Muir, T.W. (2013a). Extein residues play an intimate role in the rate-limiting step of protein trans-splicing. *Journal of the American Chemical Society* 135, 5839-5847.

Shah, N.H., Eryilmaz, E., Cowburn, D., and Muir, T.W. (2013b). Naturally Split Inteins Assemble through a "Capture and Collapse" Mechanism. *Journal of the American Chemical Society* 135, 18673-18681.

Shah, N.H., Vila-Perello, M., and Muir, T.W. (2011). Kinetic control of one-pot trans-splicing reactions by using a wild-type and designed split intein. *Angewandte Chemie* 50, 6511-6515.

Shao, Y., and Paulus, H. (1997). Protein splicing: estimation of the rate of O-N and S-N acyl rearrangements, the last step of the splicing process. *The journal of peptide research : official journal of the American Peptide Society* 50, 193-198.

Shao, Y., Xu, M.Q., and Paulus, H. (1995). Protein splicing: characterization of the aminosuccinimide residue at the carboxyl terminus of the excised intervening sequence. *Biochemistry* 34, 10844-10850.

Shao, Y., Xu, M.Q., and Paulus, H. (1996). Protein splicing: evidence for an N-O acyl rearrangement as the initial step in the splicing process. *Biochemistry* 35, 3810-3815.

Shemella, P.T., Topilina, N.I., Soga, I., Pereira, B., Belfort, G., Belfort, M., and Nayak, S.K. (2011). Electronic structure of neighboring extein residue modulates intein C-terminal cleavage activity. *Biophysical journal* 100, 2217-2225.

Shi, J., and Muir, T.W. (2005). Development of a tandem protein trans-splicing system based on native and engineered split inteins. *Journal of the American Chemical Society* 127, 6198-6206.

Sickmeier, M., Hamilton, J.A., LeGall, T., Vacic, V., Cortese, M.S., Tantos, A., Szabo, B., Tompa, P., Chen, J., Uversky, V.N., et al. (2007). DisProt: the Database of Disordered Proteins. *Nucleic acids research* 35, D786-793.

7. Bibliografia

- Singla, N., Erdjument-Bromage, H., Himanen, J.P., Muir, T.W., and Nikolov, D.B. (2011). A semisynthetic Eph receptor tyrosine kinase provides insight into ligand-induced kinase activation. *Chemistry & biology* 18, 361-371.
- Sorci, M., Dassa, B., Liu, H., Anand, G., Dutta, A.K., Pietrokovski, S., Belfort, M., and Belfort, G. (2013). Oriented covalent immobilization of antibodies for measurement of intermolecular binding forces between zipper-like contact surfaces of split inteins. *Analytical chemistry* 85, 6080-6088.
- Southworth, M.W., Benner, J., and Perler, F.B. (2000). An alternative protein splicing mechanism for inteins lacking an N-terminal nucleophile. *The EMBO journal* 19, 5019-5026.
- Starokadomskyy, P.L. (2007). Protein Splicing. *Molecular Biology* 41, 278-293.
- Studier, F.W., and Moffatt, B.A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *Journal of molecular biology* 189, 113-130.
- Subramanyam, P., Chang, D.D., Fang, K., Xie, W., Marks, A.R., and Colecraft, H.M. (2013). Manipulating L-type calcium channels in cardiomyocytes using split-intein protein transsplicing. *Proceedings of the National Academy of Sciences of the United States of America*.
- Sun, P., Ye, S., Ferrandon, S., Evans, T.C., Xu, M.Q., and Rao, Z. (2005). Crystal structures of an intein from the split dnaE gene of *Synechocystis* sp. PCC6803 reveal the catalytic model without the penultimate histidine and the mechanism of zinc ion inhibition of protein splicing. *Journal of molecular biology* 353, 1093-1105.
- Tavassoli, A., and Benkovic, S.J. (2007). Split-intein mediated circular ligation used in the synthesis of cyclic peptide libraries in *E. coli*. *Nature protocols* 2, 1126-1133.
- Thiel, I.V., Volkmann, G., Pietrokovski, S., and Mootz, H.D. (2014). An Atypical Naturally Split Intein Engineered for Highly Efficient Protein Labeling. *Angewandte Chemie*.
- Topilina, N.I., and Mills, K.V. (2014). Recent advances in in vivo applications of intein-mediated protein splicing. *Mobile DNA* 5, 5.
- Tori, K., Cheriyan, M., Pedomallu, C.S., Contreras, M.A., and Perler, F.B. (2012). The *Thermococcus kodakaraensis* Tko CDC21-1 intein activates its N-terminal splice junction in the absence of a conserved histidine by a compensatory mechanism. *Biochemistry* 51, 2496-2505.
- Tori, K., Dassa, B., Johnson, M.A., Southworth, M.W., Brace, L.E., Ishino, Y., Pietrokovski, S., and Perler, F.B. (2010). Splicing of the mycobacteriophage Bethlehem DnaB intein: identification of a new mechanistic class of inteins that contain an obligate block F nucleophile. *The Journal of biological chemistry* 285, 2515-2526.
- Tori, K., and Perler, F.B. (2011). Expanding the definition of class 3 inteins and their proposed phage origin. *Journal of bacteriology* 193, 2035-2041.
- Torrent, J., Connelly, J.P., Coll, M.G., Ribo, M., Lange, R., and Vilanova, M. (1999). Pressure versus heat-induced unfolding of ribonuclease A: the case of hydrophobic interactions within a chain-folding initiation site. *Biochemistry* 38, 15952-15961.
- Uversky, V.N., and Dunker, A.K. (2010). Understanding protein non-folding. *Biochimica et biophysica acta* 1804, 1231-1264.
- Uversky, V.N., Gillespie, J.R., and Fink, A.L. (2000). Why are "natively unfolded" proteins unstructured under physiologic conditions? *Proteins* 41, 415-427.

7. Bibliografia

- Vacic, V., Uversky, V.N., Dunker, A.K., and Lonardi, S. (2007). Composition Profiler: a tool for discovery and visualization of amino acid composition differences. *BMC bioinformatics* 8, 211.
- Van Roey, P., Pereira, B., Li, Z., Hiraga, K., Belfort, M., and Derbyshire, V. (2007). Crystallographic and mutational studies of *Mycobacterium tuberculosis* recA mini-inteins suggest a pivotal role for a highly conserved aspartate residue. *Journal of molecular biology* 367, 162-173.
- Venyaminov, S., Baikalov, I.A., Shen, Z.M., Wu, C.S., and Yang, J.T. (1993). Circular dichroic analysis of denatured proteins: inclusion of denatured proteins in the reference set. *Analytical biochemistry* 214, 17-24.
- Vihinen, M., Torkkila, E., and Riikonen, P. (1994). Accuracy of protein flexibility predictions. *Proteins* 19, 141-149.
- Vila-Perello, M., Liu, Z., Shah, N.H., Willis, J.A., Idoyaga, J., and Muir, T.W. (2013). Streamlined expressed protein ligation using split inteins. *Journal of the American Chemical Society* 135, 286-292.
- Vila-Perello, M., and Muir, T.W. (2010). Biological applications of protein splicing. *Cell* 143, 191-200.
- Vila-Perello, M., Pratt, M.R., Tulin, F., and Muir, T.W. (2007). Covalent capture of phospho-dependent protein oligomerization by site-specific incorporation of a diazirine photo-cross-linker. *Journal of the American Chemical Society* 129, 8068-8069.
- Vlieghe, P., Lisowski, V., Martinez, J., and Khrestchatskiy, M. (2010). Synthetic therapeutic peptides: science and market. *Drug discovery today* 15, 40-56.
- Volkman, G., and Iwai, H. (2010). Protein trans-splicing and its use in structural biology: opportunities and limitations. *Molecular bioSystems* 6, 2110-2121.
- Volkman, G., and Liu, X.Q. (2009). Protein C-terminal labeling and biotinylation using synthetic peptide and split-intein. *PLoS one* 4, e8381.
- Volkman, G., and Liu, X.Q. (2011). Intein lacking conserved C-terminal motif G retains controllable N-cleavage activity. *The FEBS journal* 278, 3431-3446.
- Volkman, G., and Mootz, H.D. (2013). Recent progress in intein research: from mechanism to directed evolution and applications. *Cellular and molecular life sciences : CMLS* 70, 1185-1206.
- Wang, H., Meng, X.L., Xu, J.P., Wang, J., Wang, H., and Ma, C.W. (2012). Production, purification, and characterization of the cecropin from *Plutella xylostella*, pxCECA1, using an intein-induced self-cleavable system in *Escherichia coli*. *Applied microbiology and biotechnology* 94, 1031-1039.
- Wang, S., and Liu, X.Q. (1997). Identification of an Unusual Intein in Chloroplast ClpP Protease of *Chlamydomonas eugametos*. *J Biol Chem* 272, 11869-11873.
- Wasmuth, A., Ludwig, C., and Mootz, H.D. (2013). Structure-activity studies on the upstream splice junction of a semisynthetic intein. *Bioorganic & medicinal chemistry* 21, 3495-3503.
- Williams, N.K., Liepinsh, E., Watt, S.J., Prosselkov, P., Matthews, J.M., Attard, P., Beck, J.L., Dixon, N.E., and Otting, G. (2005). Stabilization of native protein fold by intein-mediated covalent cyclization. *Journal of molecular biology* 346, 1095-1108.
- Wood, D.W., Wu, W., Belfort, G., Derbyshire, V., and Belfort, M. (1999). A genetic system yields self-cleaving inteins for bioseparations. *Nature biotechnology* 17, 889-892.

7. Bibliografia

Wu, W., Wood, D.W., Belfort, G., Derbyshire, V., and Belfort, M. (2002). Intein-mediated purification of cytotoxic endonuclease I-TevI by insertional inactivation and pH-controllable splicing. *Nucleic acids research* 30, 4864-4871.

Wüthrich, K. (1986). *NMR of Proteins and Nucleic Acids*.

Xu, M.Q., Comb, D.G., Paulus, H., Noren, C.J., Shao, Y., and Perler, F.B. (1994). Protein splicing: an analysis of the branched intermediate and its resolution by succinimide formation. *The EMBO journal* 13, 5517-5522.

Xu, M.Q., and Evans, T.C., Jr. (2005). Recent advances in protein splicing: manipulating proteins in vitro and in vivo. *Current opinion in biotechnology* 16, 440-446.

Xu, M.Q., and Perler, F.B. (1996). The mechanism of protein splicing and its modulation by mutation. *The EMBO journal* 15, 5146-5153.

Xu, M.Q., Southworth, M.W., Mersha, F.B., Hornstra, L.J., and Perler, F.B. (1993). In vitro protein splicing of purified precursor and the identification of a branched intermediate. *Cell* 75, 1371-1377.

Yang, J., Meng, Q., and Liu, X.Q. (2004). Intein harbouring large tandem repeats in replicative DNA helicase of *Trichodesmium erythraeum*. *Molecular microbiology* 51, 1185-1192.

Zeeb, M., and Balbach, J. (2004). Protein folding studied by real-time NMR spectroscopy. *Methods* 34, 65-74.

Zeidler, M.P., Tan, C., Bellaiche, Y., Cherry, S., Hader, S., Gayko, U., and Perrimon, N. (2004). Temperature-sensitive control of protein activity by conditionally splicing inteins. *Nature biotechnology* 22, 871-876.

Zettler, J., Schutz, V., and Mootz, H.D. (2009). The naturally split Npu DnaE intein exhibits an extraordinarily high rate in the protein trans-splicing reaction. *FEBS letters* 583, 909-914.

Zhang, Y., Yang, W., Chen, L., Shi, Y., Li, G., and Zhou, N. (2011). Development of a novel DnaE intein-based assay for quantitative analysis of G-protein-coupled receptor internalization. *Analytical biochemistry* 417, 65-72.

Zheng, Y., Wu, Q., Wang, C., Xu, M.Q., and Liu, Y. (2012). Mutual synergistic protein folding in split intein. *Bioscience reports* 32, 433-442.

Zuger, S., and Iwai, H. (2005). Intein-based biosynthetic incorporation of unlabeled protein tags into isotopically labeled proteins for NMR studies. *Nature biotechnology* 23, 736-740.