

Tesi doctoral

Agregació i desagregació proteica
en *Escherichia coli*: Biologia dels cossos d'inclusió

Maria del Mar Carrió Llach
2002

1. Introducció	1
1.1. Presentació	1
1.2. Paper dels cossos d'inclusió en l'enginyeria de proteïnes	2
1.2.1. Utilització d' <i>E.coli</i> com a fàbrica cel.lular	3
1.2.1.1. Morfologia, localització i composició dels cossos d'inclusió	4
1.2.2. Estratègies per millorar l'expressió de la proteïna soluble i evitar la formació de cossos d'inclusió	5
1.2.3. Aprofitament dels cossos d'inclusió com a magatzem de proteïna replegable ...	6
1.2.3.1. Protocols utilitzats per a recuperar proteïna nativa dels cossos d'inclusió ...	6
1.2.3.1.1. Preparació dels cossos d'inclusió	7
1.2.3.1.2. Aïllament dels cossos d'inclusió	8
1.2.3.1.3. Solubilització dels cossos d'inclusió	8
1.2.3.1.4. Replegament de la proteïna solubilitzada	9
1.3. Els cossos d'inclusió com a model per a estudiar les bases moleculars de l'agregació de proteïnes	11
1.3.1. Agregats proteics responsables de patologies	13
1.3.2. L'agregació de proteïnes no sempre provoca disfuncions	15
1.3.3. Propietats estructurals de la proteïna agregada en forma de cossos d'inclusió. Comparació amb altres agregats descrits.	16
1.4. Paper de les xaperones i les proteases cel·lulars en el plegament i l'agregació de proteïnes	17
1.4.1. El "trigger factor" i el sistema Hsp70	19
1.4.2. El sistema Hsp70 - Xaperones GroEL/ES	20
1.4.3. La família de les Hsp100 o Clp ATPases	21
1.4.4. La ClpB en cooperació amb la DnaK funcionen com un mecanisme de desagregació	21
1.4.5. Les Ibp	22

ÍNDEX

1.4.6. Paper de la proteolisi en l'agregació de proteïnes	22
1.4.7. Inducció de la resposta al xoc tèrmic durant la sobreproducció de proteïnes recombinants i formació de cossos d'inclusió	23
2. Antecedents del treball de recerca	25
2.1. Proteïnes recombinants utilitzades	25
2.2. Utilització d'un sistema d'expressió termoinduïble per a la producció de les proteïnes de fusió VP1LAC i LACVP1	26
2.3. Soques productores	26
2.4. Agregació i estabilitat de les proteïnes VP1LAC i LACVP1	27
2.5. La proteïna de fusió TSPA com a model per estudiar la formació de cossos d'inclusió ...	27
3. Objectius de la tesi	31
4. Resultats	33
4.1. Estudi cinètic de la formació dels cossos d'inclusió en el citoplasma d' <i>E.coli</i>	33
4.1.1. Treball 1: Dynamics of <i>in vivo</i> protein aggregation: building inclusion bodies in recombinant bacteria	35
4.2. Els cossos d'inclusió són estructures dinàmiques: transició entre la forma soluble i insoluble	43
4.2.1. Treball 2: Proteolytic digestion of bacterial inclusion body proteins during dynamic transition between soluble and insoluble forms	45
4.3. Reversibilitat de la formació dels cossos d'inclusió	53
4.3.1. Treball 3: Protein aggregation as bacterial inclusion bodies is reversible	55
4.3.2. Treball 4: Desintegració dels cossos d'inclusió quan s'atura la síntesi proteica ...	60
4.4. Organització interna dels cossos d'inclusió bacterians	61
4.4.1. Treball 5: Fine architecture of bacterial inclusion bodies	63
4.4.2. Treball 6: <i>In situ</i> proteolytic digestion of inclusion body polypeptides occurs as a cascade process	69
4.4.3. Treball 7: Aproximació estructural de la proteïna agregada en els cossos d'inclusió ...	75
4.5. Paper dels cossos d'inclusió en la biologia cel·lular de l'estrés	85

4.5.1. Treball 8: Construction and deconstruction of bacterial inclusion bodies	87
4.5.2. Treball 9: Role of molecular chaperones in inclusion body formation	97
4.6. Noves estratègies per a recuperar proteïna nativa dels cossos d'inclusió	119
4.6.1. Treball 10: Utilització d'extractes cel·lulars per a la solubilització de cossos d'inclusió ..	121
4.6.2. Treball 11: Utilització de la pressió hidrostàtica com a una eina per a solubilitzar proteïna dels cossos d'inclusió	123
5. Discussió	129
5.1. Formació i creixement dels cossos d'inclusió en el citoplasma d' <i>Escherichia coli</i>	129
5.1.1. Morfologia, mida i distribució dels cossos d'inclusió	129
5.1.2. Anàlisi de la composició dels cossos d'inclusió durant el seu creixement	130
5.1.3. Els cossos d'inclusió són estructures dinàmiques	131
5.1.4. Desintegració dels cossos d'inclusió en el citoplasma cel·lular	133
5.1.5. La formació dels cossos d'inclusió podria representar un mecanisme per reduir la toxicitat provocada pels polipèptids mal plegats	134
5.2. Estructura i organització interna dels cossos d'inclusió	134
5.2.1. Els cossos de VP1LAC i LACVP1 són proteolitzats ràpidament amb tripsina	134
5.2.2. La proteïna dels cossos d'inclusió presenta una gran heterogeneïtat conformacional	135
5.2.3. La proteïna dels cossos d'inclusió és proteolitzada <i>in vitro</i> a través d'un procés en cascada	136
5.2.4. Dinamisme conformacional dels polipèptids que formen els cossos d'inclusió ...	136
5.3. Biologia dels cossos d'inclusió en la resposta d'estrés cel·lular	137
5.3.1. La DnaK és la xaperona més important en prevenir la formació de cossos d'inclusió	137
5.3.2. El sistema DnaK/ClpB sembla que no exerceix la seva funció desagregadora amb els cossos d'inclusió	138
5.3.3. La xaperonina GroEL afavoreix la formació de cossos d'inclusió	139
5.3.3.1. Possibles mecanismes dels paper de la GroEL en la construcció dels cossos	

ÍNDIX

d'inclusió	139
5.3.4. Les IbpAB coagreguen amb els cossos d'inclusió	140
5.3.5. ClpA i ClpB no intervenen individualment en el procés de formació dels cossos d'inclusió	141
5.4. Generalitats dels cossos d'inclusió bacterians i eucariotes - Formació biològica d'agregats proteics	142
5.5. Noves estratègies per millorar el rendiment dels processos de producció de proteïnes recombinants que formen els cossos d'inclusió	143
5.5.1. Com es pot minimitzar la formació dels cossos d'inclusió	144
5.5.2. Nous mètodes per a recuperar proteïna dels cossos d'inclusió	144
5.5.3. Procés de recuperació de proteïna activa dels cossos d'inclusió utilitzant extractes cel.lulars	145
5.5.4. Utilització de la pressió hidrostàtica com a mètode per a recuperar proteïna activa dels cossos d'inclusió	146
6. Conclusions	147
7. Bibliografia	152

1.1. PRESENTACIÓ

L'estructura tridimensional d'una proteïna constitueix la forma funcional de la informació genètica que transfereix el genoma dels diferents organismes vius. Recentment, la seqüenciació del genoma de centenars d'organismes, inclòs el genoma humà, està revolucionant el món de la biologia. D'aquí a poc coneixerem la seqüència de nucleòtids de tots i cadascun dels gens d'organismes tan diversos com procariotes, llevats, nemàtodes, plantes, peixos, mamífers, etc., i per tant, coneixerem la seqüència d'aminoàcids de totes les proteïnes per les que aquests gens codifiquen. Tanmateix, malgrat la importància d'aquesta informació, encara no sabem predir l'estructura tridimensional de les proteïnes ni la via de transformació que la seqüència d'aminoàcids ha de seguir per plegar-se correctament i adquirir l'estructura funcional. Per això els avenços en la determinació de la seqüència d'aminoàcids, resten incomplets i són difícils d'explotar sense entendre el procés de plegament, que desxifra les vies de transformació cap a una unitat funcional.

Les funcions d'una proteïna venen determinades per la seva estructura, que delimita les interaccions específiques amb altres molècules. Aquesta estructura s'assoleix mitjançant el mecanisme de plegament, propi de cada proteïna. Durant l'extensió de les cadenes d'aminoàcids naixents sintetitzades pels ribosomes, aquestes s'empaqueten en estats més

globulars, a través de la unió de les regions més hidrofòbiques que s'interioritzen per evitar el contacte amb el medi aquós, i d'aquesta manera adquireixen una estructura soluble. Les conformacions proteïques que es troben en la natura són solubles o associades a membranes cel·lulars. Però el procés de plegament de proteïnes és susceptible a errors i això pot generar agregats insolubles de proteïnes que resulten en la pèrdua de la seva funció biològica.

L'estudi del plegament de proteïnes és una àrea de recerca molt activa, que inclou aspectes de biologia, química, bioquímica, bioinformàtica i física. El gran interès en entendre les vies de plegament que dirigeixen una cadena polipeptídica a la seva estructura tridimensional i activa es deu, principalment, a dues raons: 1) últimament s'estan identificant un número creixent de patologies associades al mal plegament de proteïnes i 2) l'agregació de proteïnes ocasiona entrebancs pel desenvolupament del disseny i producció de noves proteïnes amb funcions especials.

Darrerament s'han descrit moltes malalties humanes i animals causades per l'agregació de proteïnes, entre les quals es coneixen les encefelopaties espongiformes, que inclouen les malalties priòniques, com l'escrapie en ovelles, la malaltia de les vaques boges, o la malaltia de Creutzfeld-Jacob en humans; i les patologies provocades per la formació de plaques amiloides, com l'Alzheimer o el Parkinson.

El desenvolupament de la biotecnologia i, sobretot, la utilització d'organismes vius

INTRODUCCIÓ

com a fàbriques cel·lulars per a la producció de proteïnes d'interès, també s'ha vist perjudicada per problemes de plegament de proteïnes, errors en aquest procés porten a la formació de grans agregats de proteïna, coneguts com a cossos d'inclusió, que representen un dels principals obstacles per obtenir un bon rendiment dels processos de producció.

Els cossos d'inclusió van ser descrits per primera vegada per Prouty i col. el 1975. Durant el creixement d'*Escherichia coli* en presència de l'anàleg d'aminoàcids canavinina, s'observaren uns grànuls citoplasmàtics amorfes que en el microscopi òptic de contrast de fases presentaven un aspecte de partícules denses i retràctils, i en el microscopi electrònic es mostraven com estructures desorganitzades (Fig.1) (Prouty i col., 1975). L'estudi molecular d'aquests grànuls va mostrar que estaven compostos bàsicament per cadenes polipeptídiques i no presentaven cap membrana que els aïllés del medi exterior. Aquests agregats de proteïna intracel·lulars es van anomenar cossos d'inclusió i, durant molts anys, es va creure que la seva formació es donava com a una resposta cel·lular aberrant induïda pel creixement del cultiu en condicions no fisiològiques.

A partir dels anys 80, amb l'aparició i el desenvolupament de l'enginyeria de proteïnes, la formació de cossos d'inclusió s'ha vist associada a la sobreproducció de proteïnes recombinants. Les condicions que s'utilitzen per a produir aquestes proteïnes en una cèl·lula hoste afavoreixen la seva formació, els polipèptids sintetitzats

a taxes elevades no són capaços de plegar-se correctament per adquirir la seva estructura nativa i precipiten formant nuclis d'agregació.

En els processos de producció de proteïnes recombinants, la formació de cossos d'inclusió presenta avantatges i inconvenients. Aquests agregats són una font rica de proteïna recombinant concentrada, que es pot purificar més fàcilment que la proteïna soluble. De totes maneres, la proteïna purificada que s'obté no és activa, perquè no presenta la seva estructura nativa i sovint, els processos de replegament i recuperació de la proteïna soluble són tan complexes que redueixen molt el rendiment. Per això, la formació de cossos d'inclusió representa un dels principals obstacles en la indústria biotecnològica i fa que hi hagi un gran interès en estudiar les condicions que eviten la seva formació i els possibles mètodes per a recuperar la proteïna nativa d'aquests agregats. Tot i que s'han descrit diferents mètodes per minimitzar la seva formació i millorar els rendiments en la producció de proteïna soluble i activa (Strandberg i Enfors, 1991), els resultats de l'aplicació d'aquests protocols per a una proteïna en particular són imprevisibles i, en general, poc satisfactoris.

Aquest treball pretén estudiar els mecanismes involucrats en la formació dels cossos d'inclusió per buscar solucions de la problemàtica que representa i per entendre millor les bases moleculars que porten a l'agregació de proteïnes.

1.2. PAPER DELS COSSOS D'INCLUSIÓ EN L'ENGINYERIA DE PROTEÏNES

A finals dels anys 70 es van produir descobriments importants que van revolucionar la biologia molecular i en van permetre un ràpid desenvolupament. Amb el descobriment dels enzims de restricció i la implantació de protocols pel seu ús va sorgir la tecnologia de DNA recombinant, que permet l'obtenció de proteïnes d'interès a partir de la clonació dels gens que les codifiquen i la seva expressió en una cèl·lula hoste. Quan aquests gens s'introdueixen a la cèl·lula hoste reemplaçant el seu promotor natural per un promotor regulable, és possible controlar la seva expressió i, per tant, es poden produir grans quantitats de proteïnes desitjades a nivells econòmicament rentables.

La majoria de les proteïnes d'interès biomèdic com hormones, enzims, factors de creixement o antígens s'obtenen en proporcions molt petites de les seves fonts naturals i fa que la seva comercialització no sigui viable. Per això la producció de proteïnes heteròlogues en bacteris ha estat un gran avenç i actualment és un procés molt utilitzat per la indústria farmacològica, ja que són processos ràpids i de baix cost.

L'enginyeria de proteïnes ha esdevingut una poderosa eina per a la manipulació de proteïnes naturals i la creació de noves proteïnes recombinants. Un dels principals objectius d'aquesta nova tecnologia és poder dissenyar i construir proteïnes amb una estructura i funció desitjada i optimitzar la seva producció i purificació. Des que va

començar aquesta era s'han ideat moltes proteïnes recombinants amb funcions biomèdiques d'interès, així com vacunes, anticossos, etc. La possibilitat de produir qualsevol proteïna en quantitats desitjades, també ha estat molt important per a la recerca bàsica (crystallografia, bioquímica, etc.), ja que ha permès realitzar estudis estructurals, funcionals i/o bioquímics.

1.2.1. Utilització d'*Escherichia coli* com a fàbrica cel·lular

La utilització de bacteris per a la producció de proteïnes recombinants és la manera més simple i barata d'aconseguir grans quantitats de proteïna desitjada per a recerca o per la indústria. *Escherichia coli* ha estat l'organisme més usat com a fàbrica cel·lular degut a l'exhaustiu coneixement que es té d'aquest microorganisme, la seva fàcil i econòmica manipulació i la gran varietat de sistemes i vectors d'expressió que s'han desenvolupat per a la seva utilització. Actualment hi ha tants vectors comercials per a l'expressió de proteïnes recombinants en *E. coli*, que ens permeten tenir-la quasi a la carta. Aquests vectors consten d'un promotor, que pot ser induïble o constitutiu; un origen de replicació; un "polilinker" (lloc múltiple de clonatge); i un gen de resistència a un antibiòtic, per seleccionar els clons. Per altra banda, també s'han desenvolupat soques d'*E. coli* més adients per a la producció de proteïnes recombinants, com són la BL21 (Studier i Moffatt, 1986), deficient en proteases, la BL21-Star, deficient en RNAasa (Grunberg-Manago *i col.*, 1999), la BL21-Codon plus, que té incorporats els tRNA que s'uneixen a codons poc comuns (Jerpseth *i*

col., 1998) o la Orgami, en la que hi ha una millora en l'eficàcia de l'enllaç disofre (buscar ref.).

Tot i els avantatges esmentats, *E.coli* no és la solució definitiva per a produir qualsevol tipus de proteïna recombinant, ja que apareixen diversos problemes quan es pretenen obtenir grans quantitats del producte en la seva forma biològica activa. Els inconvenients més comuns amb els que ens trobem quan utilitzem *E.coli* com a organisme productor són: la toxicitat de la proteïna sobreproduïda, la incapacitat de realitzar modificacions post-traduccionals en proteïnes d'origen eucariota, la pèrdua de plasmidi durant els processos de producció, i, sobretot, la incapacitat de plegar la proteïna correctament, el que implica la seva degradació o l'acumulació en forma insoluble. Darrerament, s'ha dedicat molt esforç a l'estudi i comprensió d'aquests mecanismes. La formació de cossos d'inclusió és un dels majors problemes d'aquests processos i per això es té un gran interès en evitar que es formin i buscar estratègies per a millorar l'expressió de la proteïna soluble, com s'explica més endavant.

D'un temps ençà, els cossos d'inclusió també s'han utilitzat com a una font de proteïna pràcticament pura, mal plegada, però amb potencial per a ésser replegada a la seva forma biològicament activa. De totes maneres, el procés de recuperació de la proteïna soluble dels agregats és difícil i costós, però degut al seu interès industrial, el seu estudi està en continu desenvolupament.

Actualment, hi ha varis protocols a seguir per a intentar recuperar la proteïna atrapada en cossos d'inclusió, però les condicions òptimes són diferents per a cada proteïna individual i el procés d'optimització resulta lent i treballós.

1.2.1.1. Morfologia, localització i composició dels cossos d'inclusió

Els cossos d'inclusió han estat descrits com a agregats proteics densos i refractils, resistents a proteases i a alguns detergents, compostos majoritàriament per la proteïna recombinant mal plegada i inactiva, que es formen en el citoplasma o en el periplasma d'*E.coli* durant la seva sobreproducció (Marston, 1986). La seva formació presenta avantatges i inconvenients, la proteïna que els compona no és funcional, però sembla que es manté protegida de la proteolisis, que és el factor que fa reduir més els nivells de producció. Així doncs es pot buscar obtenir grans quantitats de proteïna agregada i prosseguir amb un procés de replegament o limitar els nivells de producció, i obtenir la proteïna en el seu estat natiu.

Tot i que l'aparició d'aquests agregats representa un problema força extens, estudis sobre la formació i les propietats d'aquests agregats són molt escassos. Els treballs que han analitzat les propietats morfològiques de cossos d'inclusió d'algunes proteïnes (Bowden *i col.*, 1991; Taylor *i col.*, 1986) mostren que la seva mida pot variar entre un rang de 0,4 μm fins a tenir el mateix volum cel·lular (2 μm). La forma dels agregats pot ser cilíndrica, esfèrica, ovoide o amorfa, en funció de la proteïna que s'aglutini i de la

localització cel·lular. Normalment, els cossos d'inclusió es localitzen en el citoplasma, si bé les proteïnes de secreció poden formar cossos d'inclusió en l'espai periplasmàtic. Estudis de cossos d'inclusió de β -lactamasa formats al periplasma i citoplasma revelen notables diferències morfològiques. Els cossos d'inclusió citoplasmàtics són cilíndrics, regulars i grans, encanvi els periplasmàtics són molt més petits, irregulars i de forma semi-esfèrica. Aquests últims, es dissocien més fàcilment en presència d'agents desnaturalitzants, que no pas els citoplasmàtics, per això es creu que les interaccions proteïna-proteïna que es produeixen a l'espai periplasmàtic són més

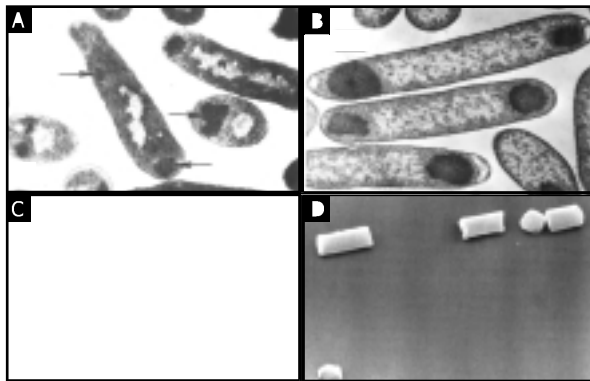


Figura 1. Imatges de microscopia electrònica de diferents cossos d'inclusió. En el panel A es mostren cèl·lules d'*E. coli* 27873 crescudes amb canavinina, les fletxes indiquen les inclusions proteiques que es formen en aquestes condicions (Prouty *i col.*, 1975). En el panel B es poden observar els cossos d'inclusió que es formen en *E. coli* recombinants, que produeixen TSP. En el panel C es mostren les micrografies electròniques de transmissió de cèl·lules d'*E. coli* RB791 que contenen cossos de β -lactamasa i en el D les imatges dels mateixos cossos purificats per centrifugació en gradient de cesi (Bowden *i col.*, 1991).

dèbils (Bowden *i col.*, 1991).

El component principal dels cossos d'inclusió és la proteïna recombinant, però també coagreguen altres components cel·lulars, com xaperones, lípids i petites quantitats d'àcids nucleics. De totes maneres, la

puresa dels agregats varia en funció de la proteïna, de l'estratègia d'expressió gènica, i de les condicions de creixement. Per exemple, hi ha diferències rellevants en els cossos d'inclusió de β -lactamasa, quan es produeixen al citoplasma o al periplasma. Els que es troben en el citoplasma són molt més homogenis (95% de proteïna recombinant), en comparació amb els que es formen en el periplasma (sols un 35%) (Valax i Georgiou, 1993).

1.2.2. Estratègies per millorar l'expressió de la proteïna soluble i evitar la formació de cossos d'inclusió

La majoria de treballs realitzats sobre cossos d'inclusió han estat dirigits a resoldre la problemàtica que comporta la seva aparició, ja sigui per evitar que es formin, o per recuperar la proteïna d'interès agregada.

Donat que l'agregació proteica representa un dels principals obstacles per obtenir bons rendiments en els processos de producció de proteïnes heteròlogues funcionals, hi ha un gran interès en correlacionar la tendència de formar cossos d'inclusió amb les propietats de la proteïna recombinant, el sistema d'expressió i la cèl·lula hoste. Tot i que no s'ha aconseguit establir unes condicions generals per evitar l'agregació de la proteïna sobreproduïda, s'han identificat molts dels factors que afavoreixen la seva formació, com les temperatures elevades, el pH àcid del citosol, la seqüència d'aminoàcids, etc. (Schein, 1991; Weicker, 1996; Chalmers, 1990; Strandberg i Enfors, 1991), els quals es poden manipular per

mirar d'esquivar-los.

Modificant les condicions d'inducció que perjudiquen el plegament de la proteïna recombinant, com són les temperatures elevades, l'acidificació del medi i l'elevada taxa de síntesi proteica, es beneficia la producció de la proteïna soluble. Per a minimitzar l'elevada concentració de polipèptids naixents, s'usen promotors més dèbils o condicions d'inducció atenuades (Baneyx, 1999; Strandberg i Enfors, 1991).

L'ambient reductor del citoplasma d'*E. coli*, no és un entorn propici pel plegament de proteïnes que presenten ponts disofre en la seva estructura nativa. Recentment, s'ha descrit una soca d'*E. coli* mutant, que té un citosol més oxidatiu, escaient pel plegament d'aquestes proteïnes (Bessette *i col.*, 1999). Produir la proteïna recombinant en el periplasma, en comptes del citoplasma, també afavoreix el plegament de proteïnes eucariotes, ja que permet la formació i isomerització de ponts disofre (Georgiou, 1995; Hockney, 1994).

S'ha observat que petits canvis en la seqüència d'aminoàcids també poden alterar les vies de plegament *in vivo* i com a conseqüència modular la tendència a agregar. S'han descrit algunes fórmules que prediuen l'estructura i la solubilitat d'una proteïna, a partir de la seva seqüència (Arakawa i Timasheff, 1985; Hoop i Woods, 1981; Wilkinson i Harrison, 1991; Carugo, 2000; Takano i Yutani, 2001), que són interessants de tenir en compte a l'hora de dissenyar i produir una nova proteïna recombinant.

Una altra maniobra que s'utilitza per a augmentar la solubilitat d'una proteïna

recombinant és fusionar-li una proteïna petita altament soluble, com la tioredoxina o el pèptid d'unió a maltosa, que facilita l'expressió de la proteïna de forma soluble (Sachdev i Chirgwin, 1998).

Tot i que la sobreproducció d'una proteïna heteròloga indueix l'expressió de gens de xaperones (veure apartat 5), que ajuden a plegar la proteïna recombinant, sovint no poden abordar l'assistència de la gran quantitat de proteïna produïda per unitat de temps. Per a restaurar aquest problema de saturació del sistema de xaperones, s'han construït soques que les coexpressen (Wall i Pluckthun, 1995; Thomas i Baneyx, 1996). S'ha vist que la sobreexpressió del sistema de xaperones DnaK/DnaJ/GrpE millora la solubilitat de la proteïna a totes les temperatures, mentre que l'expressió del sistema GroEL/GroES la fa augmentar a 30°C, però no a temperatures més elevades de 37°C (Thomas i Baneyx, 1996). L'addició d'un 3% d'etanol indueix la resposta cel·lular a l'estrés i per tant la producció de les xaperones que auxiliïen el plegament, aquest efecte també es pot aprofitar per augmentar el subministrament de xaperones (Thomas i Baneyx, 1996).

1.2.3. Aprofitament dels cossos d'inclusió com a magatzem de proteïna replegable

La producció de la proteïna recombinant en forma de cossos d'inclusió també pot resultar avantatjosa, ja que s'obté molt concentrada, en un estat que és relativament fàcil aïllar-la. Tanmateix, recuperar-la en la seva forma activa és més complex i requereix

una sèrie de passos que es descriuen a continuació.

1.2.3.1. Protocols utilitzats per a recuperar proteïna nativa de cossos d'inclusió

L'estratègia general que s'utilitza per a renaturalitzar la proteïna agregada en cossos d'inclusió inclou tres passos: primer l'aïllament i rentat dels cossos d'inclusió; segon, la solubilització de proteïna agregada, que porta a la seva desnaturalització; i finalment, el replegament de la proteïna solubilitzada. Normalment, els dos primers passos tenen una eficiència força bona, però el procés de replegament de la proteïna nativa és difícil i limita molt el rendiment del procés. Un altre factor important, que afavoreix que el procés sigui profitós, és partir d'una preparació de cossos d'inclusió pura, que depèn de les substàncies que coagreguen, durant la seva formació.

Seguint aquesta estratègia hi ha un ventall de possibles protocols a usar, que es mostren

esquemàticament a la taula 1, i s'expliquen, més detalladament, en el text. El procediment a seguir en cada cas depèn de la naturalesa de la proteïna produïda, de l'escala de producció, la qualitat i el rendiment que se'n vulgui treure.

1. PREPARACIÓ DELS COSSOS D'INCLUSIÓ

Un dels factors més importants per a obtenir un bon rendiment en la recuperació de proteïna activa dels cossos d'inclusió és la qualitat dels agregats, per la qual s'entén el grau de puresa de proteïna recombinant agregada. Aquesta depèn de dos paràmetres: 1) la quantitat de polipèptids i altres macromolècules que coagreguen durant la formació dels cossos d'inclusió i 2) l'habilitat de separar els cossos d'inclusió d'altres partícules cel·lulars que tenen el mateix coeficient de sedimentació i de material que s'absorbeix a la superfície dels agregats durant la lisi cel·lular, bàsicament proteïnes de membrana. Sovint és difícil identificar si les proteïnes alienes pertanyen a partícules

Taula 1. Sistemes més usats per al replegament de la proteïna de cossos d'inclusió

				químiques irreversibles.
--	--	--	--	--------------------------

coagregades o representen contaminants de la purificació.

La coagregació de diferents espècies proteiques depèn, principalment, del grau d'especificitat de les interaccions que porten a l'agregació de proteïnes, que varia en funció de la proteïna. Com a conseqüència, la puresa dels cossos d'inclusió depèn, en part, de la proteïna recombinant (Fink, 1998). No obstant, les condicions d'expressió, com la temperatura i la taxa de síntesi proteica també hi influeixen. Si aconseguim taxes d'agregació més baixes, es pot reduir la coagregació.

Les impureses no proteiques que coagreguen amb els cossos d'inclusió, com DNA, RNA ribosòmic i lípids, també poden interferir en el replegament posterior de la proteïna. De totes maneres, sembla que la seva interferència en el procés de replegament varia força en agregats de diferent naturalesa proteica. Mentre que en alguns casos aquestes impureses no afecten significativament la renaturalització de la proteïna (Maachupali-Reddy *i col.*, 1997), en altres la perjudiquen dramàticament (Darby i Creighton, 1990).

Minimitzar la presència de substàncies contaminants és important, ja que provoca complicacions durant els passos posteriors de solubilització i replegament. Per això es poden variar les condicions d'expressió de la proteïna recombinant, com utilitzar temperatures baixes o sistemes d'expressió suaus, per a reduir la taxa d'agregació i així augmentar la puresa del cossos d'inclusió.

2. AÏLLAMENT DELS COSSOS D'INCLUSIÓ

Les cèl·lules que contenen cossos d'inclusió, es poden rompre mitjançant una pressió elevada, per sonicació, o per l'acció

d'un enzim lític, com la lisozima. Un cop s'obté la suspensió de l'extracte cel·lular lliurat, els cossos d'inclusió poden aïllar-se fàcilment per centrifugació, ja que tenen una densitat força elevada (De Bernandez Clark *i col.*, 1998).

Hi ha proteases que es poden copurificar durant la centrifugació i es poden activar amb la presència de desnaturalitzants que s'utilitzen per a la solubilització, degradant ràpidament les proteïnes mal plegades, i reduint molt el rendiment del procés (Puri *i col.*, 1993; Babbit *i col.*, 1990). Per evitar-ho és important rentar el sediment resultant amb tampons que contenen petites concentracions d'agents caotròpics, com l'urea i l'hidroclorur de guanidina (Gdn-HCl) i detergents, com el tritón X-100 (Cardamone *i col.*, 1995; Cowley i Mackin, 1997; Rudolph, *i col.*, 1997) i sodium-dodecil-sulfat (SDS) (Burgess, 1996; Cardamone *i col.*, 1995; Kurucz *i col.*, 1995). Aquests rentats també serveixen per a eliminar els lípids i les altres proteïnes contaminants que han coagregat amb els cossos d'inclusió. S'ha d'anar en compte amb les concentracions d'urea i Gdn-HCl usades, ja que si són excessives poden portar a la solubilització dels cossos d'inclusió. Com a mètode alternatiu, es pot sotmetre l'extracte cel·lular a una centrifugació en gradient de sucrosa, que separa els cossos d'inclusió dels altres components cel·lulars (Bowden *i col.*, 1991).

3. SOLUBILITZACIÓ DEL LA PROTEÏNA DELS COSSOS D'INCLUSIÓ

Per a poder replegar la proteïna recombinant agregada en la seva estructura nativa, aquesta ha de ser prèviament

solubilitzada. Hi ha diferents mètodes que es poden utilitzar per a solubilitzar cossos d'inclusió. El més comú, és la utilització de forts agents desnaturalitzants, com 6M de Gdn-HCl o 8M urea, amb els quals la proteïna queda totalment desnaturalitzada i es trenquen totes les interaccions intermoleculars dels agregats. En aquest cas, el Gdn-HCl és preferible que la urea per dues raons (Rudolph i Lilie, 1996): 1) és un agent caotrópic més fort i pot solubilitzar cossos d'inclusió fortament agregats, que la urea n'és incapaç i 2) les solucions d'urea tenen isocianat, que poden portar a la carbamilació dels grups amino lliures dels polipèptids, especialment en temps d'incubació llargs a pHs alcalins (Hagel, 1971). Un altre mètode que es pot utilitzar per a la solubilització dels agregats és sotmetre'ls a pH extrems, i/o temperatures extremes, però aquesta aproximació no és vàlida per proteïnes que pateixen modificacions químiques irreversibles sota aquestes condicions.

Els detergents també dissocien completament la proteïna agregada. Els més utilitzats són el SDS i el n-cetyl trymethilammonium bromide (CTAB) (Qiu *i col.*, 1997; Rudolph *i col.*, 1997; Stempfer i Rudolph, 1996). L'avantatge d'utilitzar detergents és que la proteïna solubilitzada ja té activitat biològica, i per tant, s'evita el pas posterior de replegament. Tanmateix, és important eliminar les proteases de membrana associades als cossos d'inclusió, pel que finalment es requereix una sèrie de rentats òptims per eliminar les restes de detergents (Burgess, 1996). L'aplicació d'una elevada pressió hidrostàtica (1-2 kbar), en

presència d'agents reductors i baixes concentracions d'agents desnaturalitzants suposa un mitjà alternatiu per solubilitzar els agregats (De Bernandez Clark *i col.*, 1998).

Durant el procés de solubilització, és convenient l'addició d'un agent reductor en excés com β -mercaptoetanol, DTT (1,4-ditiotreita) o DTE (1,4-ditioeritrita) per a mantenir els residus de cisteïnes en un estat reduït i, així, prevenir la formació no nativa de ponts disofre. També és aconsellable incloure un agent quelant d'ions, com per exemple l'EDTA (àcid etilenediaminotetraacètic) o el EGTA (àcid 3,6-dioxaoctano-1,8-diamin.N,N,N',N'-tetraacètic) per prevenir la oxidació de les cisteïnes catalitzada pels metalls.

Per estar segur de la qualitat de la proteïna solubilitzada, sovint s'inclou un pas més de purificació, previ al replegament. Aquesta purificació es pot fer per intercanvi iònic (Ahn *i col.*, 1997; Cardamone *i col.*, 1995), per separació per grandària (Simmons *i col.*, 1997), afinitat per metalls (Negro *i col.* 1997) o per cromatografia de fases (Cerletti *i col.*, 1997). Tots aquests processos tenen en comú que utilitzen tampons que deixen la proteïna en un estat desnaturalitzat reduït.

En aquest pas, si la proteïna solubilitzada es vol guardar, es conserva en un tampó més àcid, com 10% d'àcid acètic o 5-10mM d'àcid clorhídric (Cerletti *i col.*, 1997; Kim *i col.*, 1997) i es congela en sec. S'ha de tenir en compte, però, que en algunes proteïnes, pH tan baixos poden provocar la formació d'intermediaris de plegament incapaços de replegar-se en la seva estructura nativa (Buchner i Rudolph,

1991). En aquest cas, la proteïna liofilitzada haurà d'ésser abans del replegament.

4. REPLEGAMENT DE LA PROTEÏNA SOLUBILITZADA

Els mètodes que s'utilitzen per a replegar la proteïna desnaturalitzada es basen en eliminar o reduir les altes concentracions d'agents desnaturalitzants i reductors utilitzats per a la solubilització, creant un medi idoni per a que les proteïnes es repleguin espontàniament. Entre aquests hi ha la dilució, la diàlisi, la diafiltració, la gel filtració i la immobilització en un suport sòlid.

El mètode més utilitzat en estudis de replegament a petita escala és el de diluir la proteïna solubilitzada directament en un tampó de renaturalització. El més important és controlar la taxa d'addició de la proteïna solubilitzada al tampó, per tal de mantenir una concentració baixa de proteïna desnaturalitzada i evitar l'agregació. Això s'aconsegueix diluint en múltiples passos (Herman, 1993; Maachupali-Reddy *i col.*, 1997) o utilitzant un sistema continu de "fed-batch" (Creighton, 1997). El principal desavantatge d'utilitzar aquest mètode és que requereix passos posteriors de concentració.

Una altra manera d'intercanviar el tampó per a eliminar els agents desnaturalitzants i reductors és mitjançant la diafiltració (Yon, 1996) i la diàlisi (Goldberg *i col.*, 1991), utilitzant membranes d'ultrafiltració. Un dels obstacles d'aquest mètode és que les proteïnes solubilitzades s'uneixen a la membrana, perdent-ne gran quantitat. Degut a això és aconsellable utilitzar materials altament hidrofílics, com l'acetat de

cel·lulosa, amb els que es redueix la unió de la proteïna desplegada a la membrana (Goldberg *i col.*, 1991) i dialitzar contra un tampó amb baixes concentracions d'agent desnaturalitzant, que porta a la configuració nativa de "molten-globule" (estructura secundària, pròxima a la tridimensional) (Goldberg *i col.*, 1991). Alternativament, es pot utilitzar una columna de cromatografia d'exclusió per grandària (Kim *i col.*, 1997; Simmons *i col.*, 1997; Speed *i col.*, 1996).

Una manera de minimitzar les interaccions intermoleculars que porten a l'agregació durant el procés de renaturalització és mantenir la proteïna solubilitzada immobilitzada, unida a una matriu. Per poder utilitzar aquest mètode, la proteïna ha d'estar fusionada amb un pèptid d'unió a la matriu, com pot ser el His-tag (De Bernandez Clark *i col.*, 1998) o el domini d'unió a la cel·lulosa (Speed *i col.*, 1997). Quan la proteïna solubilitzada s'uneix a la matriu a través del pèptid fusionat, la columna s'equilibra amb un tampó de renaturalització i la proteïna replegada és eluida (Thomas *i col.*, 1997; Herman, 1993; Lilie *i col.*, 1998). Aquest mètode té un rendiment força elevat.

Com es pot promoure la formació de ponts disofre durant la renaturalització de proteïnes?

Per al replegament de les proteïnes que tenen ponts disofre en la seva estructura nativa, s'han d'utilitzar tampons de renaturalització suplementats amb un sistema redox per crear un medi oxidatiu apropiat per la formació de ponts disofre. El mètode més simple i barat és utilitzar aire en presència d'un metall catalitzador i un agent reductor per a facilitar el remodelament de

punts disofre (Dorin *i col.*, 1996). Les taxes d'oxidació es poden accelerar utilitzant un sistema "oxido-shuffling", que consisteix en una mescla d'agents tiols de baix pes molecular de caràcter oxido-reductor, com el glutatión, la cisteïna i la cisteïna, que proporcionen un potencial redox. L'oxidació també es pot accelerar utilitzant una petita molècula que mimetitzava la proteïna disulfid isomerasa (Katzav-Gozansky *i col.*, 1996). L'inconvenient d'aquests sistemes és que les substàncies que s'utilitzen són molt cares.

El repliegament oxidatiu no és efectiu per a proteïnes que presenten una via de plegament difícil, amb intermediaris de baixa solubilitat. En aquests casos, el repliegament es fa en dues etapes. Primer la proteïna desnaturalitzada és totalment oxidada en presència d'un excés de glutatión que converteix tots els grups SH- a disulfits i deixa els residus de la proteïna carregats, afavorint la solubilitat. Després, es dilueix en presència d'agents reductors per promoure la formació dels punts disofre i el repliegament (Maachupali-Reddy *i col.*, 1997).

L'eficiència de la renaturalització depèn de la competència entre el plegament i l'agregació. L'agregació de la proteïna recombinant desplegada augmenta amb la presència d'altres proteïnes contaminants que es troben a la preparació i que per elles mateixes ja tenen tendència a agregar.

Com es pot evitar la reagregació de proteïnes durant el procés de repliegament?

El principal problema que fa disminuir el rendiment dels processos de repliegament

de la proteïna és la formació de polipèptids mal plegats i la seva agregació. L'estratègia que més s'utilitza per a suprimir aquesta agregació és l'addició de molècules de baix pes molecular que inhibeixen la formació d'interaccions intermoleculares. Aquestes petites molècules són relativament fàcils d'eliminar un cop la proteïna ha estat repliegada. Els additius més comuns són la L-arginina (0,4-1M), petites concentracions d'agents desnaturalitzants, com l'urea (1-2M) i Gdn-HCl (0,5-1,5M) i detergents (Chaps, SDS, CTAB i tritón X-100). També s'utilitzen cofactors específics, com el Zn^{2+} o el Ca^{2+} , que estableixen proteïnes en estat d'intermediaris de plegament i així prevenen l'agregació.

1.3. ESTUDI DELS COSSOS D'INCLUSIÓ COM A MODEL PER A ESTUDIAR LES BASES MOLECULARS QUE PORTEN A L'AGREGACIÓ DE PROTEÏNES

Les proteïnes acabades de sintetitzar han de seguir la ruta de plegament que les dirigeixi cap a la seva estructura nativa única, funcional i energèticament més estable. Les vies de plegament que segueix cada cadena polipeptídica són específiques i tot i que ja se'n coneixen moltes, es desconeix com la proteïna n'utilitza una i descarta les altres. Hi ha un gran nombre de proteïnes que es pleguen immediatament després de ser sintetitzades de manera espontània, però n'hi ha d'altres que requereixen un patró de plegament més complex, en el que es formen estats conformacionals transitoris metaestables, que es coneixen com a intermediaris de plegament. Aquestes últimes

INTRODUCCIÓ

són més susceptibles a patir errors en algun moment del procés i això pot comportar greus disfuncions, degut a que les proteïnes mal plegades poden ocasionar toxicitat o pèrdua de la seva funcionalitat cel·lular.

Les proteïnes que no aconsegueixen plegar-se correctament, generalment són degradades pel sistema proteolític cel·lular per evitar que provoquin disfuncions. Una de les conseqüències més traumàtiques del mal plegament proteic és l'agregació de proteïnes. Aquest fenomen es dona quan intermediaris estructurals que exposen les regions hidrofòbiques a la superfície o segments desestructurats interaccionen entre ells a través de forces hidrofòbiques o ponts d'hidrogen (Dobson i Karplus, 1999; Ellis i Hemmingsen, 1989; Frydman, 2001; Gething i Sambrook, 1992; Hartl, 1996; Radford, 2000), en comptes d'encaixonar-se per formar l'estructura nativa. Per això, condicions en les que es genera una gran població d'intermediaris

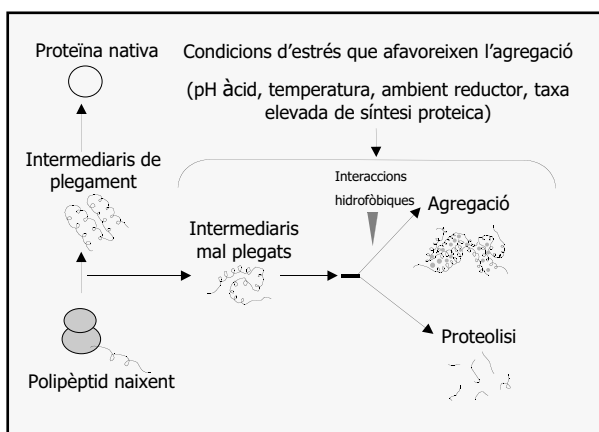


Figura 2. Els cossos d'inclusió es formen per l'agregació d'intermediaris mal plegats que apareixen fora de la correcta via de plegament. Aquests són els precursors de l'agregació i estan en competència amb la proteolisi. Les condicions de sobreexpressió d'una proteïna dificulten el plegament proteic i afavoreixen l'acumulació d'intermediaris mal plegats.

parcialment plegats, tendeixen a fomentar l'agregació.

Durant el procés de plegament, la via que ha de seguir el polipèptid per assolir l'estructura nativa està en competència amb l'agregació i la degradació d'estructures aberrants per les proteases cel·lulars (Figura 2). L'èxit d'aquest procés depèn de l'eficàcia del sistema de control de qualitat cel·lular (compost per xaperones i proteases) que s'encarrega d'assistir el plegament d'altres polipèptids i evitar l'agregació. Per això l'agregació no es dona en situacions normals, sinó en situacions d'estrés, quan aquest sistema està saturat. L'agregació es veu afavorida per la sobrepoblació d'intermediaris de plegament i per condicions desestabilitzants com elevades temperatures, pH àcid, presència d'agents desnaturitzants com l'urea, etc.

La natura dels intermediaris precursors de l'agregació difereix en els diversos tipus d'agregats, poden ser molècules totalment desplegadas o colapsades en estat de "molten-globule". L'estat de "molten globule" pot procedir del procés de plegament, com sembla que és en el cas dels cossos d'inclusió, però també pot derivar de la proteïna nativa, com és el cas dels intermediaris que formen les plaques amiloidees extracel·lulars (Khurana i col., 2001). Hi ha un ampli ventall de possibles vies de plegament i per tant, de vies que porten cap a l'agregació, però en tots els casos, la finalitat del procés és que el polipèptid naixent assoleixi un estat energètic més favorable. Així doncs, quan els intermediaris de plegament no aconsegueixen adquirir l'estructura nativa, o bé són eliminats per proteolisi o són conduïts cap a l'agregació, que

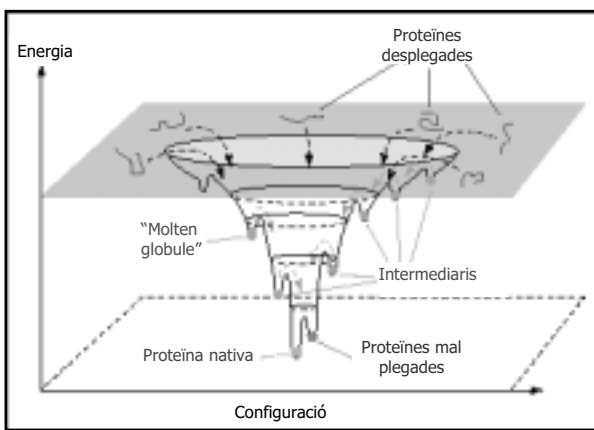


Figura 3. Diagrama esquemàtic del camp energètic del plegament. Les proteïnes desnaturalitzades, poden seguir diferents rutes de plegament per adquirir la seva forma nativa, que resulten de buscar l'estat de mínima energia lliure. Les diferents vies poden passar per formes intermediàries, que depenen de la proteïna (Radford, 2000).

proporciona un estat més estable, més proper al natiu (Fig.3) (Radford, 2000).

Es coneixen agregats proteics ben estructurats, com els amiloides, que presenten una organització fibril·lar ordenada, formada per un procés de polimerització nuclear (creixement ordenat a partir d'un nucli d'agregació) (Jarret *i col.*, 1993). Per altra banda, també existeixen agregats amorfs, com els cossos d'inclusió, dels quals se'n desconeix el seu procés de formació. Tanmateix, es creu que les bases moleculars que porten a la formació

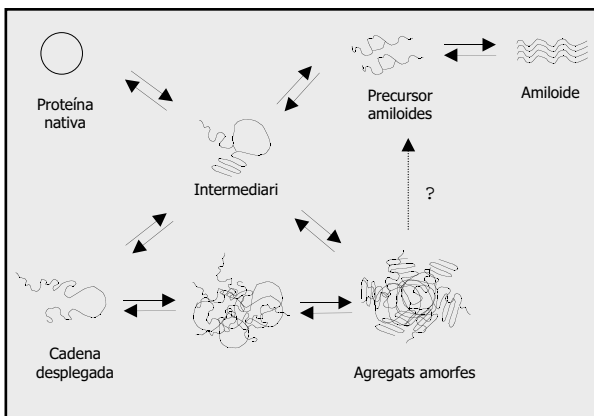


Figura 4. Esquema de les possibles vies que porten a la formació d'agregats amorfs i ordenats. La proteïna desplegada pot formar agregats amorfs, o pot seguir la ruta de plegament, que sovint passa per estats intermediaris. Aquests que poden donar lloc a precursors de l'agregació i formar agregats de tipus amiloide, o agregar de forma desordenada. Fink *i col.* proposen que en el procés de formació dels amiloides, hi ha una fase prèvia d'agregació desordenada (Fink, 1998).

de diferents agregats són comunes o similars (Khurana *i col.*, 2001).

Fins fa poc, s'assumia que l'agregació era un procés passiu, dirigit per les forces químiques existents entre els intermediaris de plegament. Però recentment s'ha demostrat que alguns agregats intracel·lulars es formen com a una resposta fisiològica a l'excés de polipèptids mal plegats. Es tracta d'un mecanisme actiu que depèn del transport retrògrad a través dels microtúbuls, de la proteïna agregada cap a un gran i únic agregat, anomenat agrosoma (Kopito, 2000). El descobriment de l'existència d'un mecanisme cel·lular que s'encarrega de dirigir la formació dels agrosomes obre noves perspectives sobre el paper que pot jugar l'agregació proteica a la cèl·lula i suggereix que podria representar un procés inclòs dins de la maquinària de resposta a l'estrès, i no ésser una conseqüència aberrant del mal plegament proteic, com es creia.

1.3.1 Agregats proteics responsables de patologies

L'interès que ha sorgit darrerament per l'agregació de proteïnes es deu a la seva associació a varies malalties humanes amb una gran incidència a la població. Les diferents patologies estan relacionades amb la formació de grans agregats de naturalesa fibril·lar, ben estructurats, que es coneixen com amiloidosis. Aquest grup de malalties, de les quals se n'han descrit unes 20, inclouen majoritàriament desordres neurodegeneratius, com l'Alzheimer, el Parkinson, les encefalopaties espongiformes provocades per prions, etc. (la taula 1 resumeix les propietats d'aquestes malalties que es caracteritzen per

INTRODUCCIÓ

formar agregats proteics), però també es sumen altres desordres, com la diabetes de tipus II. Tot i que les proteïnes implicades amb la formació dels amiloides no estan relacionades entre elles, els agregats presenten una estructura molt similar entre ells, tenen uns 10 nm de gruix i la proteïna es troba en estructura secundària de tipus fulla beta, orientada perpendicularment a l'eix de la fibra. Per això es creu que l'agregació en forma d'amiloides no és específica de la seqüència, sinó que es creu que és deguda a una propietat genèrica de totes les cadenes polipeptídiques, sota unes condicions determinades (Ellis i Pinheiro, 2002).

Els agregats fibrilars que tenen tan els amiloides com els prions es formen per un procés de polimerització nuclear (Jarrett i Lansbury, 1993). Aquest nucli es forma quan polipeptíds mal plegats escapen del sistema de control de qualitat de la cèl·lula i s'acumulen en quantitats suficients per a engegar el procés d'agregació, que a partir d'aquest moment és ràpid (Taylor *i col.*, 2002). En molts casos, abans de formar-se les fibres amiloidees altament ordenades, es passa per un estadi d'agregat amorf que, curiosament, presenta una estructura més propera a la nativa (Khurana et al., 2001)(veure Fig.5).

Moltes de les malalties neurodegeneratives anomenades estan associades a la mutació d'un gen que genera la producció d'una proteïna anòmala que agrega i provoca la patogènesi. La malaltia d'Alzheimer està causada per la formació d'agregats de la proteïna β -amiloid (β A). Es creu que la formació d'aquestes plaques amiloidees són

les responsables de l'inici de la patologia, i que totes les altres propietats de la malaltia són secundàries (Hardy i Higgins, 1992). Per això les teràpies en desenvolupament estan enfocades a evitar l'agregació proteica.

En la malaltia del Parkinson, la patogènesi també s'inicia per la formació d'agregats proteics citoplasmàtics en neurones dopaminèrgiques. La proteïna responsable és l' α -sinucleïna i els agregats que es formen es coneixen com a cossos "Lewy". Es creu que l'agregació de l' α -sinucleïna està relacionada amb una deficiència en la degradació proteica, ja que aquesta interacciona normalment amb una lligasa del sistema del proteosoma-ubiquitina (proteïna parkin) (Shimura *i col.*, 2001). Així doncs, l' α -sinucleïna anòmala, podria provocar una disfunció en la degradació proteica, que podria ser la responsable de la formació dels agregats (Farrer *i col.*, 2001; Shimura *i col.*, 2001).

La progressió de les malalties prioniques es diferencien dels altres desordres neurodegeneratius per què la proteïna responsable de la patologia, la Prp, té propietats infectives. La proteïna Prp cel·lular (Prp^C) és soluble i rica en hèlix alfa, la forma patològica (Prp^{SC}) es diferencia en la seva estructura tridimensional, que presenta un contingut elevat de fulles beta, i és insoluble. La presència de Prp^{SC} promou la conversió de la forma normal Prp^C a la forma anòmala Prp^{SC} i així s'inicia la propagació de la patologia (Kaneko *i col.*, 1997).

Hi ha un altre grup de malalties neurodegeneratives relacionades amb la formació d'agregats proteics, que estan

causades per una mutació dominant del gen responsable de la malaltia, basada en l'expansió de repeticions de poliglutamines. Aquest tipus de mutació provoca canvis conformacionals de la proteïna, que porten a l'agregació proteica, la formació de cossos d'inclusió i la mort de cèl·lules neuronals (ref). Aquest grup de desordres inclouen la malaltia de Huntington, l'atrofia muscular espinobulbar, i les ataxies espinocerebelars de tipus 1, 2, 3, 6 i 7. La localització dels agregats que es produeixen difereix en funció de la proteïna, en alguns casos es tracta d'agregats nuclears i en d'altres citosòlics (Taylor *i col.*, 2002). Les seves propietats estructurals s'assemblen molt a les dels agregats amiloidogènics, en que es formen per un procés de nucleació, que presenten una estructura secundària rica en fulles betes i que els agregats estan organitzats per filaments semblants a les fibres amiloidees (Chen *i col.*, 2002). Alguns estudis suggereixen que la formació dels

agregats no són els responsables de la patologia, sinó que són el resultat de la progressió de la malaltia (Klement *i col.*, 1998; Sisodia, 1998).

1.3.2. L'agregació de proteïnes no sempre provoca disfuncions

Hi ha unes proteïnes de llevats que es classifiquen com a prioniques per tenir un comportament similar als prions animals. Els prions de llevats, així com els de mamífer, confereixen canvis fenotípics a través d'interaccions entre proteïnes, en les quals una proteïna prionica, formada de manera anormal, promou la conversió de les noves proteïnes homòlogues cap a l'estructura anormal, que es caracteritza per formar grans agregats. Així com en mamífers l'aparició de l'estructura anòmala de la proteïna i la seva agregació provoca patologies mortals, en llevats confereixen un canvi fenotípic únic i viable, alterant el

Taula 2. Malalties provocades per agregats proteics

Malaltia	Agregats proteics	Proteïna tòxica	Gens implicats en la malaltia
Malaltia d'Alzheimer	Plaques extracel·lulars	A β	APP Presenilin 1 † Presenilin 2 †
Malaltia de Parkinson	Agregats intracel·lulars Cossos Lewy	tau α -Synuclein	α -synuclein* Parkin † UCHL1 †
Malalties prioniques	Plaques prioniques	Prp ^{SC}	PRNP*
Malalties per poliglutamines	Inclusions nuclears i citoplasmàtiques	Proteïnes que contenen poliglutamines	9 gens diferents amb expansions repetides CAG* tau*
Tautopaties	Agregats citoplasmàtics	tau	tau*
Esclerosi lateral amiotrofia familiar	Cossos Bunina	SOD1	SOD1*

* Les mutacions patògenes estan associades a la toxicitat de la proteïna

† Les mutacions patògenes estan associades a una pèrdua de funció

patró de lectura dels codons d'aturada (en el cas de la proteïna Sup35) o modificant la via d'assimilació del nitrogen (en el cas de la proteïna Ure2) (Serio *i col.*, 2001). Aquest fenotip és dominant i s'hereda citoplasmàticament. Aquest fenomen representa un mecanisme molt interessant ja que indica que una proteïna pot funcionar com un element d'herència genètica, mitjançant canvis conformacionals (Serio *i col.*, 2001), que promouen l'agregació proteica. En aquest cas, la formació d'agregats proteics no s'entén com una aberració cel·lular.

Els agregats que formen els prions de llevats mostren un patró característic de fibres amiloidees (Glover *i col.*, 1997; Serio i Lindquist, 2000). Darrerament, s'està estudiant les similituds que presenten algunes proteïnes responsables de malalties humanes, com la malaltia de Huntington, amb aquestes proteïnes prioniques de llevats (Sup35 i Ure2), ja que sembla que comparteixen algunes característiques (Uptain *i col.*, 2001).

Un altre fet interessant d'aquest mecanisme és el paper que hi juga la xaperona Hsp104. Aquesta xaperona té la capacitat de desfer agregats creats per xoc tèrmic (Parsell *i col.*, 1994), però en aquest context afavoreix el procés de polimerització dels agregats i la seva propagació citoplasmàtica. El creixement dels grans agregats prionics es pot bloquejar degut a interaccions amb altres proteïnes o amb fragments proteolítics de la mateixa, la xaperona Hsp104 trenca els agregats en petits troços i el fet de deixar les proteïnes prioniques lliures i exposades

afavoreix el seu creixement (Kushinrov i Ter-Avanesyan, 1998).

1.3.3. Propietats estructurals de la proteïna agregada en forma de cossos d'inclusió bacterians. Comparació amb altres agregats descrits

Els cossos d'inclusió bacterians són un tipus d'agregats que han estat molt poc estudiats a nivell estructural i semblen presentar algunes propietats diferents dels descrits fins ara. Per això seria molt interessant de conèixer com es formen i com s'organitzen per a entendre millor el fenomen de l'agregació proteica

Quan inicialment es van descriure els cossos d'inclusió, es creia que eren partícules de proteïna agregada denses i compactes, formades per polipèptids mal plegats de manera irreversible. La seva resistència proteolítica, s'associava a la inaccessibilitat de les proteases a la proteïna agregada, degut al seu caràcter dens (Enfors, 1992). Però algunes observacions posteriors suggereixen que tenen una estructura porosa, amb cavitats internes en les que el solvent hi pot penetrar (Bowden *i col.*, 1991; Taylor *i col.*, 1986). Es va calcular el voidatge de diferents cossos d'inclusió, que és el volum de solvent associat a les partícules, i van observar que és més gran que el contingut de solvent que s'observa en els cristalls de proteïna, la qual cosa suggereix que les cadenes polipeptídiques en els cossos d'inclusió, s'organitzen d'una manera més aleatòria que en els cristalls (Matthews, 1974).

Les bases moleculars que porten a l'agregació *in vivo* no es coneixen amb

profunditat, però hi ha evidències de que els cossos d'inclusió es formen per l'associació d'intermediaris de plegament amb un alt grau d'estructura secundària. Oberg *i col.* van estudiar l'estructura secundària de les proteïnes incloses en els cossos d'inclusió d'interleukina-1 β mitjançant FTIR (Espectroscopia d'infraroig per la transformada de fourier), i van observar que presenten una estructura secundària pràcticament igual que la proteïna nativa (Oberg *i col.*, 1994). També van analitzar agregats formats per desnaturalització i d'altres formats durant el replegament de proteïna desnaturalitzada i tots presentaven espectres similars. Es coneixen altres casos, com els cossos d'inclusió de cel·lulasa, en els que les proteïnes mostren un estat conformacional molt pròxim al natiu, però d'altres en que l'estructura es troba una mica més pertorbada, com en els agregats de β -lactamasa (Pryzbyecien *i col.*, 1994). Les perturbacions estructurals que es troben en diferents agregats són molt similars, consisteixen en un augment de fulles beta a expenses del contingut en hèlix alfa de l'estructura nativa. El fet que els intermediaris de plegament que formen els agregats presentin estructura secundària nativa, significa que l'agregació es dona en estadis avançats de la ruta de plegament (Oberg *i col.*, 1994). La presència d'estructura secundària en les partícules també encaixa amb l'estructura porosa dels agregats (Bowden *i col.*, 1991). A més, la identificació d'activitat enzimàtica en cossos d'inclusió de β -galactosidasa (Worrall i Goss, 1989) i de l'endoglucanasa D de *Clostridium thermocellum* (Tokatlidis *i col.*, 1991) suggereix que com a

mínim, hi ha d'haver una fracció de la proteïna agregada amb una estructura molt propera a la nativa. Worrall *i col.* van proposar el 1989 que la proteïna activa present als cossos d'inclusió procedia de proteïna agregada després d'ésser correctament plegada.

Tot i que s'ha proposat que les interaccions intermoleculares que porten a l'agregació són específiques i es donen entre regions homòlogues (Speed *i col.*, 1996), la presència d'altres proteïnes coagregades suggereix que el grau d'especificitat de les interaccions varia en funció de la proteïna.

1.4. PAPER DE LES XAPERONES I LES PROTEASES CEL·LULARS EN EL PLEGAMENT I L'AGREGACIÓ DE PROTEÏNES

En *E.coli*, el sistema cel·lular que s'encarrega d'assistir el plegament de noves proteïnes i d'actuar com a control de qualitat proteic consisteix en dos tipus de proteïnes: les xaperones moleculars, que s'encarreguen de prevenir el mal plegament i l'agregació de proteïnes, i les proteases, que s'encarreguen d'eliminar estructures defectuoses. Les xaperones moleculars creen un entorn idoni per a que les noves cadenes polipeptídiques puguin adquirir la seva estructura tridimensional, però no aporten informació conformacional sobre el procés de plegament (Ellis, 2002).

La maquinària cel·lular de xaperones actua durant el plegament de noves proteïnes per evitar la seva agregació en condicions normals i, més acusadament, en condicions d'estrés, com les temperatures elevades. Hi ha moltes xaperones i

INTRODUCCIÓ

proteases, que a part d'ésser expressades constitutivament en condicions normals, sota condicions d'estrés cel·lular o de xoc tèrmic s'indueix la seva producció en grans quantitats, per això es classifiquen com a proteïnes d'estrés o de xoc tèrmic ("heat-shock") (Goff i Goldberg, 1985). En *E. coli*, les proteïnes de resposta a l'estrés estan regulades per un mateix reguló anomenat σ^{32} . En situacions extremes es sintetitza el factor de transcripció σ^{32} , que indueix l'expressió dels promotors de xoc tèrmic (Hartl, 1996; Yura, 1999). En els mutants deficients en el factor σ^{32} , s'observen molts més agregats proteics (Tomoyasu *i col.*, 2001), fet que evidencia la importància d'aquesta xarxa de xaperones i proteases sobre el plegament proteic.







Les xaperones assisteixen principalment a proteïnes que requereixen un procés de plegament llarg i complicat, en les que es formen estructures intermediàries que inclouen estats mal plegats que exposen regions hidrofòbiques al solvent i tendeixen

a agregar. Aquestes reconeixen residus hidrofòbics i/o regions desestructurades, de manera que s'hi uneixen i dirigeixen els polipèptids cap a la via productiva del plegament, evitant l'agregació (Deuerling *i col.*, 1999; Hartl i Hayer-Hartl, 2002).

En *E. coli*, les xaperones citosòliques participen en el plegament de noves proteïnes, bàsicament per dos mecanismes diferents: el sistema Hsp70 (DnaK/DnaJ/GrpE) i el "trigger factor" (TF), que actuen sobre la cadena polipeptídica recent sintetitzada, proporcionant un estat competent pel plegament i el sistema Hsp60 (GroEL/GroES), que actua en un estadi posterior, oferint un espai compartimentat, físicament aïllat del citosol, dins del qual proteïnes completes o dominis proteics es poden plegar completament. Aquestes dues classes de xaperones estan conservades en els tres dominis de la vida (Frydman, 2001; Langer *i col.*, 1992).

A part d'aquestes funcions bàsiques, les xaperones esmentades i altres participen en

Taula 3. Família de xaperones d'*Escherichia coli*: estructura i funció

			Membre d' <i>E. coli</i>	Número de còpies per cèl·lula	Fenotip del mutant nul	Funció
			ClpA		No fenotip	Degradació de proteïnes, juntament amb la ClpP
			ClpB	500	Termotolerància perjudicada	Desagregació d'agregats proteics, junt amb el sistema DnaK
Hsp90	dímer 	+	HtpG	1050	Es redueix la taxa de creixement a 44°C	Desconeguda
Hsp70	monòmer 	+	DnaK Co-xaperona DnaJ/GrpE	9900	Sensibilitat a la temperatura pel creixement (39°C)	Plegament de proteïnes recent sintetitzades Prevenició de l'agregació proteica a elevades temperatures Desagregació d'agregats proteics, junt amb la ClpB
Hsp60	14-mer 	+	GroEL Co-xaperona GroES	1230	Letal	Plegament de proteïnes recent sintetitzades Prevenició de l'agregació proteica a elevades temperatures
sHsp	8-24-mer 		IbpA IbpB	< 600	No fenotip	Prevenició de l'agregació proteica a elevades temperatures
Trigger-factor	monòmer 		Trigger-factor	20000	No fenotip	Xaperona associada al ribosoma, plegament de noves proteïnes

una gran diversitat de funcions relacionades amb la maquinaria de plegament de les proteïnes cel·lulars, per tant tenen una gran importància en el procés d'agregació (Fayet *i col.*, 1989; Horwich *i col.*, 1993; Gething i Sambrook, 1992; Ben-Zvi i Goloubinoff, 2001; Schein, 1991). Tanmateix, la intervenció de les xaperones en l'agregació de proteïnes es coneix poc i menys en la formació de cossos d'inclusió. Les propietats de les principals xaperones d'*E.coli* es descriuen breument a la taula 2 i el seu paper en el plegament i l'agregació de proteïnes està representat a la Figura 5.

1.4.1. El "Trigger Factor" i el sistema Hsp70

El "trigger factor" és una proteïna bacteriana de 48 kDa, que s'uneix als ribosomes i interactua amb les cadenes polipeptídiques naixents curtes (fins a 57 residus) (Hesterkamp *i col.*, 1996). El complexe cadena naixent-TF es dissocia de manera ATP independent, quan la cadena es desprèn del ribosoma (Hesterkamp *i col.*, 1996). La funció del TF es sobreposa amb la del principal sistema bacterià de DnaK/DnaJ, que s'encarrega d'estabilitzar les noves cadenes d'aminoàcids i proporcionar-les-hi un estat competent per a plegar-se correctament (Deuerling *i col.*, 1999; Teter *i col.*, 1999). Les dues xaperones tenen activitat isomerasa, que es creu que permet mantenir les noves cadenes sintetitzades en un estat flexible, que els hi permeti plegar-se ràpidament just després de la dissociació de la xaperona (Bukau *i col.*, 2000). Els eucariotes no tenen TF, però en canvi, tenen un complex heterodimèric d'associació a ribosoma

anomenat complex associat a la cadena naixent o NAC (de Nascent-chain Associated Complex) (Wiedmann, 1994).

La família Hsp70 existeix en el citosol d'eubacteries, eucariotes i algunes arqueobacteries. L'estructura i els aspectes mecànics del sistema Hsp70 que es coneixen millor són els d'eubacteris, que inclouen la Hsp70, DnaK, la coxaperona Hsp40, DnaJ i el factor GrpE d'intercanvi de nucleòtids. La xaperona DnaK consisteix en un domini ATPasa N-terminal de 44 kDa i un domini C-terminal de 27 kDa, d'unió a pèptid (Bukau i Horwich, 1998). La xaperona s'associa al pèptid substrate mitjançant interaccions hidrofòbiques i ponts d'hidrogen (Zhu, 1996). La unió al pèptid es produeix quan la DnaK està unida a ATP, això provoca la hidròlisi d'ATP junt amb un canvi conformacional de la xaperona. El cicle de DnaK entre aquests estats està regulat per la DnaJ i la GrpE. El domini N-terminal de la DnaJ s'uneix a DnaK i accelera la hidròlisi d'ATP (Rudiger *i col.*, 2001; Sha *i col.*, 2000), el domini C-terminal reconeix pèptids hidrofòbics i GrpE completa

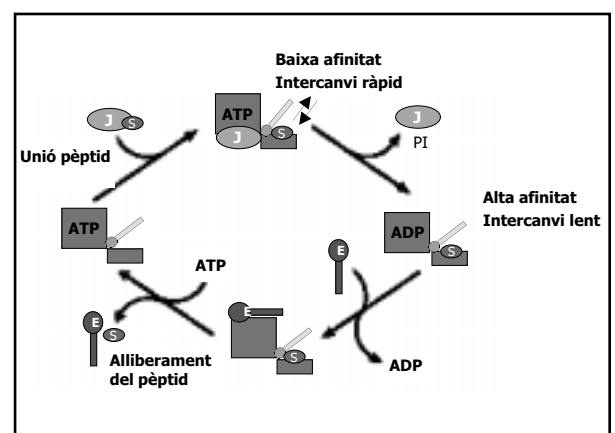


Figura 6. Cicle de funcionament del sistema DnaK/DnaJ/GrpE. DnaK està representada en groc i verd; DnaJ, J; GrpE, E i la proteïna substrate, S. L'intermediari DnaK-DnaJ-substrate-ATP probablement és molt transitori perquè és el pas més ràpid del cicle.

INTRODUCCIÓ

el cicle (Harrison *i col.*, 1997), induint l'alliberament de l'ADP i DnaK i dissociant el complexe DnaK-pèptid.

La DnaK és la xaperona per excel·lència encarregada d'evitar l'agregació, interaccionant amb els intermediaris propensos a l'agregació, i conduint-los cap a la via productiva del plegament. Una altra funció important que la DnaK exerceix sobre l'agregació és la de dissociar agregats produïts per xoc tèrmic, en cooperació amb la ClpB (Ben-Zvi i Goloubinoff, 2001), que s'explica més endavant.

1.4.2. El sistema Hsp70 - Xaperonines GroEL/ES

Un altre tipus de xaperones més grans i complexes són les xaperonines. Aquestes presenten una estructura cilíndrica, formada per un doble anell, amb una cavitat interna, que proporciona un compartiment aïllat físicament del medi citosòlic, on proteïnes senceres o dominis proteics es poden plegar correctament de manera més eficient. Les proteïnes que presenten una estructura no nativa són capturades per interaccions hidrofòbiques i dirigides cap a la cavitat central on es poden plegar, protegides de l'agregació amb altres proteïnes del medi.

GroEL i el seu cofactor GroES, representen el sistema més estudiat de xaperonines. La GroEL està composta per dos anells heptamèrics idèntics, de 57 kD, col·locats un davant de l'altre. Cada subunitat consisteix en tres dominis, l'equatorial, que té el lloc d'unió a ATP, que està unit al domini apical a través del domini tipus frontissa ("hingelike"). El domini apical s'encarrega d'obrir el cilindre i exposar els residus hidrofòbics que atrauen a

les proteïnes substrat. GroES és un anell homoheptamèric de 10 kDa que tanca i obre el cilindre de GroEL mitjançant la GroEL ATPasa. La superfície hidrofòbica exposada de la part apical de GroEL interacciona amb els residus hidrofòbics d'intermediaris de plegament. La GroES té seqüències bucle mòbils, que contacten les regions d'unió a substrat amb els dominis apicals de GroEL i

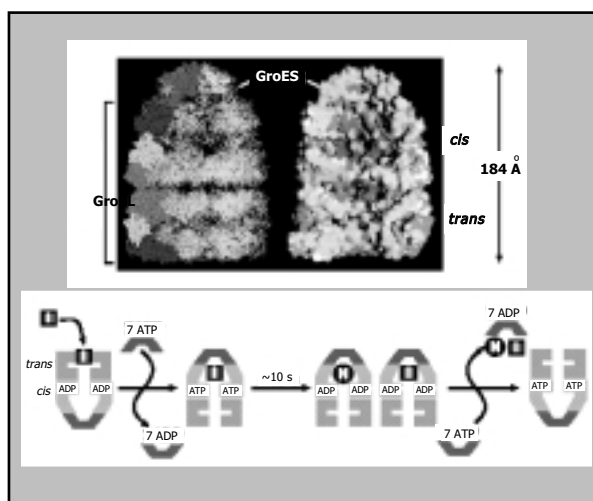


Figura 7. Sistema de xaperonines GroEL-ES. El panel superior mostra el complex asimètric GroEL-GroES(ADP)₇. Els dominis equatorial, intermediari i apical de cada subunitat de la GroEL estan representats en color rosa, groc i blau, i la subunitat GroES en vermell. En la imatge de la dreta, els àtoms polars carregats estan en color blau, els àtoms hidrofòbics, en groc, els de l'esquelet, en blanc i la superfície exclosa pel solvent, en gris. El panel B esquematitza el mecanisme de plegament proteic en la cavitat GroEL/ES, essent I, intermediari de plegament, que s'uneix al domini apical de GroEL i N, la proteïna nativa que es plega dins la cavitat. Sovint es necessiten varies rondes d'internalització del substrate a la xaperonina per assolir el plegament natiu (Hartl i Hayer-Hartl, 2002).

mediem la dissociació del substrat (Fig.7) (Bukau i Horwich, 1998; Hartl, 1996; Sigler, 1998).

Fins fa poc es creia que la GroEL només podia assistir el plegament de proteïnes que es poguessin encapsular a la cavitat formada per GroEL/GroES (>60 Kda), però recentment s'ha descrit un mecanisme d'acció de la GroEL per proteïnes més grans,

en el que no es requereix que aquestes entrin a la cavitat interna (Chaudhuri *i col.*, 2001).

Últimament s'ha descrit una nova propietat de la GroEL, que és particularment interessant en el context d'aquest treball. S'ha observat que la xaperonina pot induir l'agregació de la proteïna priònica Prp recombinant desnaturalitzada *in vitro*. Els agregats que es formen són resistents a proteolisi i presenten propietats molt similars als agregats de Prp^{SC} o Prp^{res}, associats a patologies priòniques (Stöckel i Hartl, 2001). Aquesta nova troballa suggereix que l'agregació és un procés actiu, mediat per factors cel·lulars. A continuació es mostra el model descrit a partir d'estudis *in vitro*.

1.4.3. La família de les Hsp100 o Clp ATPases

Les proteïnes Hsp100/Clp constitueixen una família de xaperones ATP-dependents molt conservades en la natura, trobades tan en eucariotes com en procariotes. Aquestes xaperones estan involucrades en varis processos cel·lulars que inclouen la reactivació proteica, la degradació de proteïnes, la replicació de DNA, la regulació de l'expressió gènica, la termotolerància, l'herència de factors tipus priònic i la translocació a través de membranes (Bukau i Horwich, 1998; Chernoff *i col.*, 1995; Mhammedi-Alaoui *i col.*, 1994; Parsell *i col.*, 1994; Schewder *i col.*, 1996; Tobias *i col.*, 1991; Vogel *i col.*, 1995). Molts membres d'aquesta família, com la ClpA, ClpX i ClpY, són components reguladors de les proteases dependents d'ATP, que

s'associen amb les subunitats peptidases ClpP i ClpQ, encarregades de la degradació proteica. La funció de les Clp ATPases en aquests complexos és reconèixer i desplegar els substractes proteics per a transferir-los a la subunitat peptidasa, que les proteolitzarà. El fet que les funcions ATPasa i peptidasa estiguin separades en diferents subunitats, permet que aquestes també puguin actuar independentment, amb altres proteïnes.

A diferència de les altres Clps, la ClpB d'*E. coli* i de *Thermus thermophilus*, i el seu homòleg en llevats, Hsp104, no estan involucrades en la degradació proteica, sinó en la desagregació i reactivació de proteïnes mal plegades i agregades (Parsell *i col.*, 1994; Goloubinoff *i col.*, 1999; Ben-Zvi i Goloubinoff, 2001).

1.4.4. La ClpB en cooperació amb la DnaK funcionen com un mecanisme de desagregació

La ClpB és un membre de la família de les Hsp100, molt conservada en la natura. Quan es va començar a estudiar, es va associar a la termotolerància per la seva capacitat de desagregar cúmuls proteics per temperatura. Darrerament s'ha descrit una acció cooperativa de la ClpB, conjuntament amb la DnaK, DnaJ i GrpE que facilita la desagregació d'agregats proteics estables. La funció concreta de cada xaperona en particular no està massa clara, però sembla que es tracta d'un mecanisme seqüencial, en el que la ClpB s'uneix inicialment a la proteïna substracte, seguit de la unió de la DnaK, que facilita la

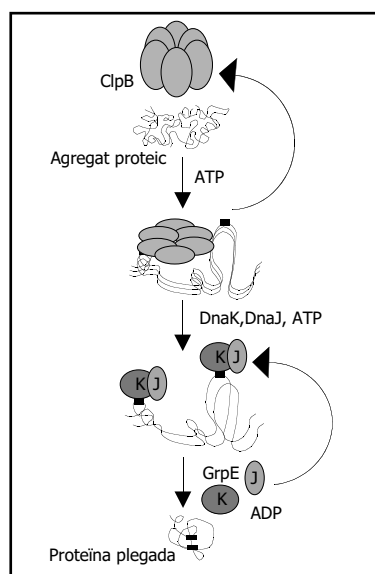


Figura 8. Mecanisme seqüencial de desagregació del sistema DnaK-ClpB. La ClpB reconeix els agregats proteics i amb l'ajuda d'ATP els desagrega, la proteïna desplegada és transferida al sistema DnaK/J/E per ser plegada a la seva forma nativa. Esquema adaptat de Goloubinoff *i col.*, 1999.

dissociació

dels agregats (Goloubinoff *i col.*, 1999; Ben-Zvi i Goloubinoff, 2001; Enfors, 1992; Zolkiewski, 1999). El mecanisme està representat a la figura

1.45. Les IbpAB

Les Ibps pertanyen a la família de les proteïnes de xoc tèrmic petites (small heat shock proteins). Representen les proteïnes de resposta a l'estrés més desconegudes i amb menys homologia entre elles. Se n'han identificat unes 25 en diferents organismes i totes elles tenen un pes molecular baix, entre 15 i 30 Kda. En bacteries s'han descrit les Ibps (Inclusion Bodies Protein), que es van denominar així per trobar-se associades als cossos d'inclusió. Les Ibps estan codificades pels gens IbpA i IbpB, de 14 i 16 Kda (Allen *i col.*, 1992). Tot i que la seva funció encara es desconeix, s'ha observat que s'uneixen a polipèptids desnaturalitzats i els estableixen i aquests poden ésser replegats espontàniament o transferits al sistema DnaK/DnaJ/GrpE per a plegar-se a través d'un procés dependent d'ATP (Thomas i Baneyx, 1998; Veinger *i col.*, 1998). *In vitro*,

IbpA i IbpB poden mantenir intermediaris de plegament en una conformació idònia per a ésser reactivats per la DnaK, la qual cosa suggereix que actuen com a reservoris de proteïna desnaturalitzada per estrès cel·lular, de manera que posteriorment pugui ser reactivable (Laskowska *i col.*, 1996; Thomas i Baneyx, 1998).

1.4.6. Paper de la proteòlisi en l'agregació de proteïnes

La proteòlisi juga un paper molt important en el sistema de control de qualitat cel·lular del plegament proteic, ja que s'encarrega d'eliminar els polipèptids mal plegats del citosol, que representen els precursors de l'agregació. Així doncs, si aquesta no funciona eficientment, s'afavoreix l'agregació.

Els gens que codifiquen per la majoria de les proteases d'*E.coli* també estan regulades pel factor de transcripció σ^{32} i s'indueixen per xoc tèrmic i condicions d'estrés (Rosen *i col.*, 2002; Yura, 1999). Formen part del mateix mecanisme de regulació que les xaperones, tenint com a substracte els mateixos polipèptids mal plegats. Encara és una incògnita si les vies alternatives (degradació/replegament) competeixen entre elles.

Les proteases citoplasmàtiques encarregades de degradar les proteïnes aberrants són dependents d'energia. Dins d'aquest grup de proteïnes hi ha la ClpP, ClpX, Lon, i HflB, essent HflB la única essencial. Les proteases i peptidases independents d'energia, assumeixen l'eliminació de petits fragments peptídics (Gottesman, 2002; Suzuki *i col.*, 1997).

En *E.coli* també hi ha una gran varietat de

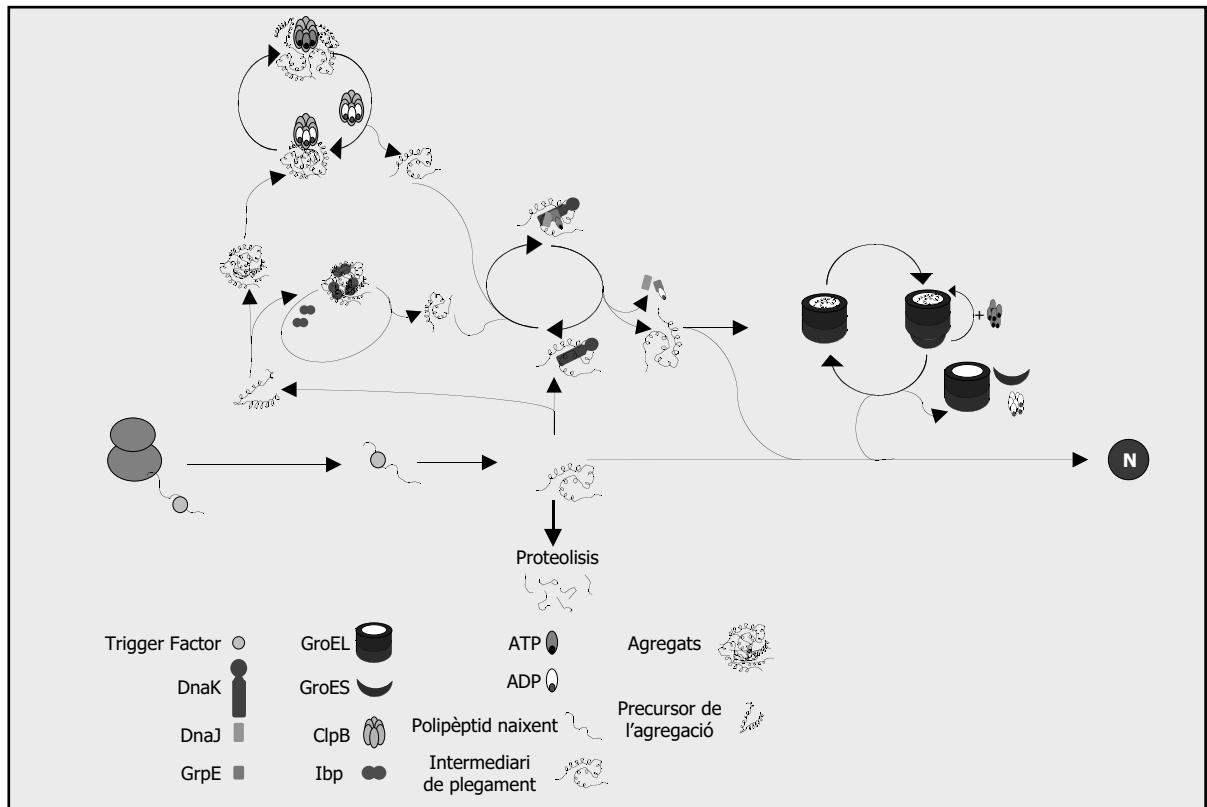


Figura 9. Esquema general del funcionament de les xaperones en el plegament proteic. Les cadenes polipeptídiques recent sintetitzades poden plegar-se espontàniament o poden entrar dins del sistema de xaperones, DnaK/J/E i GroEL/ES, que els hi proporciona un ambient adequat per a ser plegades correctament. Les cadenes polipeptídiques parcialment o mal plegades poden ésser proteolitzades o poden agregar. En alguns casos les petites xaperones, Ibps, coagreguen amb aquests i per això es creu que podrien tenir la funció de proporcionar una agregació reversible. La xaperona ClpB pot desagregar agregats proteics, deixant la proteïna desplegada, per a ser transferida al sistema de plegament DnaK/J/E.

proteases i peptidases extracitoplasmàtiques, que s'encarreguen de processar pèptid senyals de proteïnes de secreció i d'eliminar proteïnes mal plegades a l'espai extracitoplasmàtic (Miller, 1996).

1.4.7. Inducció de la resposta al xoc tèrmic durant la sobreproducció de proteïnes recombinants i formació de cossos d'inclusió

Les condicions de sobreproducció d'una proteïna recombinant indueixen fortament l'expressió de l'operó σ^{32} , que regula l'expressió de xaperones i proteases (Dong *i col.*, 1995; Rinas, 1996; Kanemori *i col.*, 1994; Parsell i Sauer, 1989; Hoffman i Rinas, 2000). La

temperatura elevada és el paràmetre que estimula més ràpidament i acusadament la seva expressió (Lemaux *i col.*, 1978). La presència de polipèptids mal plegats també la indueixen, però de manera més lenta i constant. La regulació del factor σ^{32} és dependent de la concentració de DnaK lliure, que estabilitza el factor σ^{32} . Quan apareixen molts polipèptids mal plegats, la DnaK té més afinitat per aquests, i això desestabilitza σ^{32} , que serveix com a senyal per la inducció del sistema de resposta al xoc tèrmic (Yura i Nakahigashi, 1999; Bukau, 1993).

Els components involucrats en la resposta a l'estrés, s'expressen immediatament després d'augmentar la temperatura, amb un temps màxim de 5 a 15 minuts, i es manté

INTRODUCCIÓ

durant una hora, en els cas dels sistemes majors de xaperones Hsp70 (DnaK, DnaJ i GrpE) i Hsp60 (GroEL i GroES) i les Clps. Les lbps es comporten de manera diferent, tenen un pic d'expressió en el mateix moment que les altres, però la seva expressió no es manté, i es deixen de produir quan la proteïna recombinant es comença a acumular en cossos d'inclusió (Hoffman, 2000).

La ràpida activació dels sistema de xaperones en aquestes ocasions, ens indica que quan s'inicia el procés de formació dels cossos d'inclusió, el nivell de xaperones cel·lulars es troba en el seu òptim (Georgiou i Valax, 1999).

Aquest treball va sorgir d'una línia de recerca del laboratori dedicada a estudiar els processos de producció de proteïnes recombinants en *Escherichia coli*, i la influència de les condicions de sobreproducció i de les propietats de la proteïna recombinant en el rendiment d'aquests processos (Corchero *i col.*, 1996; Corchero *i col.*, 1997; Vila *i col.*, 1997). Per altra banda, també s'havien explorat les respostes fisiològiques de la cèl·lula a la sobreproducció d'una proteïna heteròloga.

2.1. PROTEÏNES RECOMBINANTS UTILIZADES

Es varen construir una sèrie de proteïnes recombinants de fusió entre l'enzim β -galactosidasa d'*Escherichia coli* i la proteïna VP1 de la càpside del virus de la febre aftosa, amb l'objectiu d'estudiar la influència de diferents paràmetres en la seva producció. Les proteïnes de fusió s'han usat àmpliament com a una estratègia per a millorar la producció del pèptid heteròleg fusionat o com a base per a dissenyar proteïnes multifuncionals, o en la producció de pèptids amb interès vacunal, per augmentar el tamany i estructura del producte final i així millorar la seva immunogenicitat. En *E.coli*, l'enzim β -galactosidasa és l'encarregat d'hidrolitzar la lactosa en glucosa i galactosa, i això fa que aquesta espècie pugui utilitzar la lactosa com a font de carboni. La temperatura òptima d'aquesta reacció és de 35°C i el pH òptim de 7.2, tot i que l'enzim es manté estable entre pHs de 6 i 8 (Creuger i Creuger, 1984). A més del substrate natural, la lactosa, la β -galactosidasa pot hidrolitzar altres

substractes, com orto-nitrofenil-b-D-galactopiranòsid (ONPG), clorofenol galactopiranòsid (CRPG) i el 5-bromo-4-cloro-3-indol-b.D-galactopiranòsid (X-gal), donant lloc a compostos que presenten coloració. Gràcies a aquesta propietat, l'activitat enzimàtica pot ser fàcilment detectada. La forma activa de l'enzim és un homotetràmer de 464 kDa (Fowler i Zabin, 1978), que es forma a través de l'ensamblatge de quatre monòmers, codificats pel gen *lacZ*, i que tenen una seqüència de 1023 aminoàcids (Kalnins *i col.*, 1983). La seva estructura tridimensional està descrita i hi predominen les hèlix alfa (Jacobson *i col.*, 1994). La presència d'una proteïna de fusió als extrems C- o N-terminal d'aquest enzim no interfereix en la seva activitat enzimàtica (Fowler i Zabin, 1978; Winther *i col.*, 1986), i per això representa un bon model per a estudiar les propietats de les proteïnes de fusió.

La proteïna estructural VP1 de la càpside del virus de la febre aftosa conté els 2 llocs antigènics principals del viriò. En el serotip C, el principal determinant antigènic s'anomena lloc A i es troba entre els residus 138-150 dins del llaç G-H exposat al solvent (Strohmaier *i col.*, 1982). El lloc A inclou varis epítops (Mateu *i col.*, 1987; Mateu *i col.*, 1989; Mateu *i col.*, 1990) i a més es troba la seqüència Arg-Gly-Asp (RGD). Aquest tripèptid, conservat en tots els serotips, és la zona d'unió al receptor cel·lular (Fox *i col.*, 1989). Degut a la complexitat biològica que presenta l'estructura antigènica del lloc A d'aquest virus i a la importància econòmica de la febre aftosa en la medicina veterinària i diagnòstica, es va utilitzar com a model

d'estudi de la producció de proteïnes víriques en forma de fusió a l'enzim β -galactosidasa. Així doncs, d'entre altres construccions, es va fusionar la proteïna VP1 a l'extrem carboxi i amino terminal de la β -galactosidasa, i les proteïnes resultants es van anomenar VP1LAC i LACVP1 (Corchero *i col.*, 1996).

2.2. UTILITZACIÓ D'UN SISTEMA D'EXPRESSIÓ TERMOINDUÏBLE PER A LA PRODUCCIÓ DE LES PROTEÏNES DE FUSIÓ VP1LAC I LACVP1

L'elecció del promotor que dirigirà l'expressió d'un gen recombinant és de crucial importància en l'èxit d'un procés de producció, ja que influeix sobre factors com l'estabilitat del mRNA, la fidelitat en la incorporació d'aminoàcids, la susceptibilitat del producte a proteòlisi i la formació de cossos d'inclusió.

Els promotors que es troben sota el control de repressors termosensibles presenten una regulació molt fina dels nivells d'expressió a través de l'ús de la temperatura com a inductor (Villaverde *i col.*, 1993) i permeten assegurar que l'inductor es mantingui constant al llarg del procés, cosa que no és fàcil d'aconseguir quan s'utilitzen inductors químics. En aquest context, l'expressió dels gens des dels promotors lítics del fag lambda p_R i p_L recombinants es controla mitjançant el repressor termosensible *ci857*, que s'inactiva a mesura que augmenta la temperatura, de 30°C a 42°C. Normalment, una desavantatge d'utilitzar aquests sistemes és que s'augmenta la tendència de que els

productes recombinants s'acumulin en forma de cossos d'inclusió (Strandberg i Enfors, 1991), però en el context d'aquest treball ens interessa afavorir la seva formació.

Les proteïnes escollides per aquest estudi VP1LAC i LACVP1, es produeixen a partir del vector d'expressió pJLA602, basat en els promotors p_L i p_R del fag λ , col·locats en tàndem i sota el control del repressor *ci857*, el gen del qual també es troba en el plàsmid. Aquest plàsmid també consta d'un lloc múltiple de clonatge i d'un gen de resistència a ampicilina com a marcador (Shauder *i col.*,

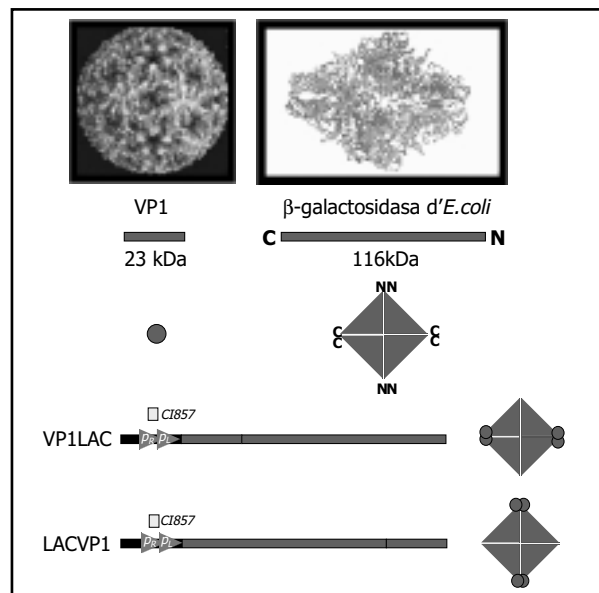


Figura 1. Esquema de la construcció de les proteïnes de fusió VP1LAC i LACVP1 amb el pèptid VP1 del virus de la febre aftosa i l'enzim β -galactosidasa d'*Escherichia coli*.

1987).

2.3. SOQUES PRODUCTORES

Les dues proteïnes de fusió esmentades, VP1LAC i LACVP1, es varen produir en dues soques d'*Escherichia coli*, la MC1061, K12, *hsdR2*, *mcrB*, *araD139*, $\Delta(\text{araABC}^- \text{leu})7679$, ΔlacX74 , *galU*, *galK*, *rpsL*, *thi*, *strA* (Sambrook *i col.*, 1989) i la BL21, deficient en la proteasa cel·lular La, amb el genotip *Lon- OmpT- hsdS*

gal (Studier i Moffatt, 1986).

2.4. AGREGACIÓ I ESTABILITAT DE LES PROTEÏNES VP1LAC I LACVP1

Els estudis de producció de VP1LAC i LACVP1 mostraven que la seva producció variava molt en funció de la soca productora. En la soca deficient en la proteasa La, BL21, la proteïna recombinant es trobava majoritàriament acumulada en cossos d'inclusió, degut a que la proteïna que no podia ser plegada correctament, no era degradada pel sistema proteolític cel·lular i era dirigida cap a l'agregació (Corchero *i col.*, 1996).

Un fet que ens va cridar l'atenció, va ser que els cossos d'inclusió de les dues proteïnes VP1LAC i LACVP1, presentaven aspectes molt diferents, essent els de VP1LAC molt més grans que els de LACVP1. Aquestes diferències es deuen a que la proteïna LACVP1 és més sensible a proteòlisi, possiblement degut a una major quantitat de dianes de proteases a la regió d'unió. Això suggereix que l'estabilitat proteolítica d'una proteïna determina (almenys, en part) la seva agregació. Per tant, no és del tot cert que l'agregació determini l'estabilitat proteolítica, com s'havia proposat (Hellebust *i col.*, 1989). Per altra banda, es desconeix si els fragments de degradació que es troben en els cossos d'inclusió han agregat després d'ésser proteolitzats, o resulten de la proteòlisi dels polipèptids agregats.

Quan les proteïnes VP1LAC i LACVP1 es produeixen en la soca MC1061, agreguen immediatament després d'ésser sintetitzades, però no formen cossos

d'inclusió. Es va estudiar l'estabilitat d'aquesta fracció de proteïna agregada i es va veure que té una vida mitja més curta que la proteïna soluble i és degradada per proteòlisi limitada, que allibera el fragment heteròleg i permet el replegament i la solubilització dels fragments de degradació de β -galactosidasa. Els resultats observats suggerien que aquests agregats no estaven protegits de la proteòlisi, com s'havia descrit pels cossos d'inclusió. Per això es va proposar que en les primeres fases de l'agregació, s'activen les dianes de proteases de la proteïna mal plegada generant una sensibilitat a la proteòlisi transitòria, que es podia perdre en agregats més densos com els cossos d'inclusió (Corchero *i col.*, 1997). Per altra banda, aquest fenomen també predia una associació entre la proteòlisi i la solubilització, que era prou interessant per a ser estudiada amb més profunditat.

Arrel d'aquests resultats va sorgir l'interès d'estudiar més detalladament les característiques que presentaven els cossos d'inclusió, el seu procés de formació i esbrinar si realment es tractaven d'agregats irreversibles i resistents a proteases, com es creia (Marston, 1986, Hellebust *i col.*, 1989).

2.5. LA PROTEÏNA DE FUSIÓ TSPA COM A MODEL PER A ESTUDIAR LA FORMACIÓ DE COSSOS D'INCLUSIÓ

Entre la varietat de proteïnes de fusió dissenyades i produïdes en el laboratori de Villaverde i col. (Corchero *i col.*, 1996), se'n va dissenyar una sèrie amb la proteïna de la cua del bacteriòfag P22, coneguda com "tailspike protein" (TSP), fusionada amb el pèptid GH23

ANTECEDENTS DEL TREBALL DE RECERCA

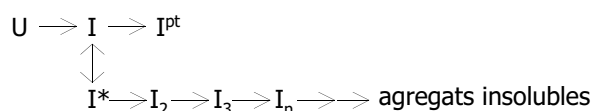
de la proteïna VP1 (de 23 aminoàcids, que inclou el lloc A). Es van realitzar fusions a l'extrem amino i carboxi terminal, anomenades ATSP i TSPA, respectivament (Carbonell i Villaverde, 1998). Aquestes proteïnes es produeixen per un sistema d'expressió induïble per IPTG p_{trc} en la soca d'*Escherichia coli* BL26.

La TSP és la proteïna que forma la cua dels bacteriòfags P22 i presenta dues activitats. Per una banda, el reconeixement i unió a les càpsides víriques dels bacteriòfags P22 durant la darrera fase de morfogènesi del virió, mitjançant el domini amino terminal. Per altra, l'activitat endoramosidasa, que uneix i hidrolitza els enllaços α -ramnosil-1,3-galactosa presents en la unitat repetitiva de l'antigen O de la cadena de lipopolisacàrid (LPS) de la superfície bacteriana (Iwashita i Kanegasaki, 1976; Wollin *et al.*, 1981). Es tracta d'una proteïna homotrimèrica de 666 aminoàcids (72 kDa), on predomina l'estructura en fulla beta. Cada monòmer presenta la forma general d'un peix, i està compost per sis segments que s'han denominat cos principal, boca, aleta dorsal i aleta caudal. Aquesta última està dividida en tres segments, i conté tres fulles betes paral·leles (anomenades A, B i C pels autors) i dos antiparal·leles (C i D). A l'aleta dorsal es troben una fulla beta antiparal·lela força recargolada (full bF) i només cinc hèlix α .

La TSP representa un bon model per estudiar agregació, ja que les rutes de plegament i d'agregació estan ben caracteritzades. La proteïna nativa és termoestable i resistent a SDS i proteòlisi, la qual cosa facilita l'estudi i l'aïllament dels

intermediaris que intervenen en les vies de plegament i agregació (Speed *et al.*, 1996).

La cadena polipeptídica recent sintetitzada, en estat desplegat (U), adopta una conformació monomèrica parcialment plegada que s'anomena intermediari parcialment plegat (I), el qual ja presenta una forta estructura secundària. Els intermediaris I interaccionen a través del domini carboxiterminal, formant un segon intermediari protrimèric (I^{pt}), que es continua plegant fins l'estadi de protrimer (pt) en el que es troben les tres cadenes associades, sense plegar-se completament. Aquest intermediari és metaestable i a baixes temperatures ($<10^{\circ}\text{C}$) es manté en un estadi de preplegament durant varies hores. A temperatures per sobre dels 30°C el protrimer té una vida mitja de minuts (Danner *et al.*, 1993; Fuchs *et al.*, 1991; Goldenberg i King, 1982) i acaba plegant-se en la forma nativa (N) seguint una cinètica de primer ordre. A temperatures restrictives de $38-42^{\circ}\text{C}$ es desestabilitza l'intermediari I de la via de plegament i es desvia cap a la via d'agregació. Així doncs, la forma I està en equilibri amb l'espècie I^* , que es dirigeix eficientment cap a l'agregació originant intermediaris multimèrics que poden ser identificats en gels no desnaturitzants de poliacrilamida. Els temps mitjans per a les reaccions d'agregació són similars als observats en la via de plegament. Danner i



Seckler (1993) determinen les vides

		U	→→	M	→→	pt	→→	N
$t_{1/2}$	10°C	7 min.		~30 min.		~10 h.		
	25°C	~1,5 min.		~10 min.		~30 min.		
		U= cadenes desplegadas		pt= protrimer				
		M= monòmers parcialment plegats		N= trimer natiu				

mitjanes ($T_{1/2}$) d'aquestes espècies per fluorescència del triptòfan i SDS-PAGE.

Degut a la seva tendència a agregar, la producció de TSP en *E.coli* sovint dona lloc a la formació de cossos d'inclusió (Carbonell i Villaverde, 1998; King *i col.*, 1996; Speed *i col.*, 1996), que representen un bon model per estudiar els intermediaris de plegament que intervenen en el procés d'agregació. La complexitat del seu plegament també ofereix un interès per a estudis de replegament de la proteïna dels cossos

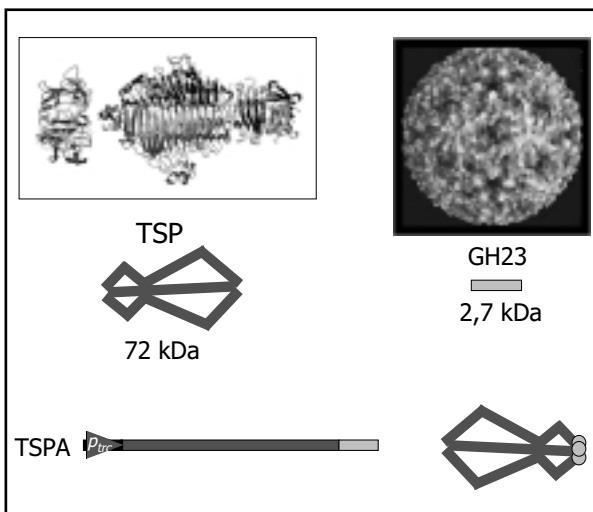


Figura 2. Esquema de la construcció de la proteïna de fusió TSPA amb la proteïna TSP del bacteriòfag P22 i el pèptid GH23 de la proteïna VP1 de a càpside del virus de la febre aftosa.

d'inclusió. Per a aquest estudi s'han utilitzat els cossos d'inclusió de la proteïna de fusió TSPA (Carbonell i Villaverde, 1998).

ANTECEDENTS DEL TREBALL DE RECERCA

OBJECTIUS

A. OBJECTIUS INICIALS DE LA TESI

1. Estudiar cinèticament la formació dels cossos d'inclusió de les dues proteïnes VP1LAC i LACVP1 en el citoplasma d'*Escherichia coli*.
 - 1.1. Descriure com creixen els agregats en el citoplasma d' *E.coli* per microscòpia òptica. Examinar la seva morfologia, el seu creixement volumètric, la localització i distribució cel·lular.
 - 1.2. Caracteritzar la composició proteica dels agregats durant el seu creixement i la taxa d'agregació de proteïna recombinant.
2. Intentar descobrir quin paper té la formació dels cossos d'inclusió en la biologia de l'estrés cel·lular. Per això es volen identificar les xaperones que estan involucrades en el procés de formació dels cossos d'inclusió de VP1LAC, utilitzant una sèrie de soques deficientes en les xaperones DnaK, GroEL, GroES, ClpA, ClpB, IbpAB i DnaKClpB
 - 2.1. Analitzar les diferències dels cossos d'inclusió produïts en les diferents soques, per a determinar la influència de les xaperones estudiades sobre el seu procés de formació.
3. Determinar l'estat estructural dels polipèptids que formen els cossos d'inclusió de VP1LAC i LACVP1
 - 3.1. Analitzar el procés de digestió proteolítica de les proteïnes recombinants que formen els cossos i fer el seguiment de la dissociació dels grans agregats a través de microscopia electrònica, per a tenir una idea de la seva arquitectura interna.
 - 3.2. Examinar les propietats estructurals dels polipèptids dels cossos d'inclusió mitjançant tècniques d'espectroscopia com la fluorimetria i el dicroïsme circular i veure si es produeixen canvis conformacionals durant el seu creixement.
4. Desenvolupar un mètode de recuperació de proteïna nativa dels cossos d'inclusió que sigui més eficaç dels descrits fins ara, amb un major rendiment, més econòmic, que es pugui escalejar i es minimitzi l'ús d'agents químics.
 - 4.1. Examinar la possibilitat de desagregar els cossos d'inclusió i recuperar la proteïna activa aplicant una elevada pressió hidrostàtica.

B. NOUS OBJECTIUS QUE HAN APAREGUT DURANT EL TREBALL EXPERIMENTAL

El fet de descobrir que l'agregació en forma de cossos d'inclusió és reversible, ens va portar a plantejar-nos nous objectius durant la realització de la tesi.

5. Investigar l'evolució dels cossos d'inclusió de VP1LAC i LACVP1 en el citoplasma cel·lular quan s'atura la síntesi de proteïna recombinant per a conèixer més a fons el fenomen de la reversibilitat de l'agregació.
 - 5.1. Determinar la proteïna recombinant agregada que es perd i la que es troba a la fracció soluble.
 - 5.2. Esbrinar si la proteïna que es solubilitza és replegada en la seva forma activa, analitzant la seva activitat enzimàtica.
6. Identificar les xaperones que participen en el mecanisme de desagregació dels cossos d'inclusió per entendre quina funció té aquest mecanisme en la resposta d'estrés cel·lular i per explotar aquesta funció per a recuperar proteïna dels cossos d'inclusió. Per això, s'estudiarà aquest procés en la mateixa col·lecció de soques deficientes en xaperones.
7. Estudiar la possibilitat de mimetitzar el procés de solubilització dels cossos d'inclusió que es dona *in vivo*, *in vitro*. Examinar el potencial d'utilitzar extractes cel·lulars sotmesos a un xoc tèrmic per a recuperar la proteïna nativa dels cossos d'inclusió.

4.1. ESTUDI CINÈTIC DE LA FORMACIÓ DELS COSSOS D'INCLUSIÓ EN EL CITOPLASMA D'*Escherichia coli*

En aquest treball es va analitzar la formació dels cossos d'inclusió de VP1LAC en la soca d'*E.coli* BL21 (Lon-) durant les primeres 5 hores d'inducció de la proteïna recombinant i al cap de 24 hores. L'estudi es va fer des d'un punt de vista morfològic, analitzant el seu creixement volumètric i la seva distribució a la cèl·lula, a partir d'imatges de microscòpia òptica. Per altra banda, es va analitzar la taxa d'agregació de la proteïna recombinant en els cossos d'inclusió. Aquest estudi descriptiu mostra que durant la fase exponencial de síntesi proteica, hi ha una acumulació constant de proteïna recombinant als cossos d'inclusió, que es correspon amb l'increment de volum.

RESULTATS



Dynamics of in vivo protein aggregation: building inclusion bodies in recombinant bacteria

M.M. Carrió, J.L. Corchero, A. Villaverde *

*Institut de Biologia Fonamental and Departament de Genètica i Microbiologia, Universitat Autònoma de Barcelona, Bellaterra,
08193 Barcelona, Spain*

Received 21 June 1998; received in revised form 16 September 1998; accepted 17 September 1998

Abstract

Time-dependent aggregation of a plasmid-encoded β -galactosidase fusion protein, VPILAC, has been carefully monitored during its high-rate synthesis in *Escherichia coli*. Immediately after recombinant gene induction, the full-length form of the protein steadily accumulates into rapidly growing cytoplasmic inclusion bodies. Their volume increases during at least 5 h at a rate of $0.4 \mu\text{m}^3 \text{h}^{-1}$, while the average density remains constant. Protein VPILAC accounts for about 90% of the aggregated protein throughout the building process. Minor components, such as DnaK and GroEL chaperones, have been identified in variable, but low concentrations. The homogeneous distribution of inclusion bodies among the cell population and the coexistence of large, still growing bodies with newly appearing aggregates indicate that the aggregation cores are mutually exclusive, this fact being a main determinant of the in vivo dynamics of protein aggregation. © 1998 Federation of European Microbiological Societies. Published by Elsevier Science B.V. All rights reserved.

Keywords: Inclusion body; Recombinant protein; Protein aggregation; β Galactosidase; Misfolding; *Escherichia coli*

1. Introduction

Protein aggregation is involved in important human and animal diseases and it has become the object of intense investigation. In bacteria, the overproduction of heterologous proteins often results in the formation of insoluble, quite homogeneous protein aggregates called inclusion bodies (IBs) [1]. While in a production context, IB formation represents a

major obstacle that can be partially overcome by the solubilization and refolding of active protein [2], it also offers an intriguing model to explore the molecular basis of misfolded protein-protein interactions [3,4]. Although structural analyses of IBs abound, the in vivo kinetics of IB formation has been poorly studied. In this work, we have explored the construction of IBs in *Escherichia coli* cells producing the model chimeric protein VPILAC [5]. The aggregation of VPILAC molecules in the cell cytoplasm takes place at nearly constant rates, around nucleation cores that are mutually exclusive. The regularity of protein aggregation permits a steady growth of IBs throughout the productive phase of the culture.

* Corresponding author.
Tel.: +34 (93) 5812148; Fax: +34 (93) 5812011;
E-mail: a.villaverde@cc.uab.es

2. Materials and methods

2.1. Plasmids and bacterial strains

The *E. coli* strain *BL21 lon ompT hsdS gal* [6] was used for recombinant gene expression from plasmids pJCO46 and pJVPILAC [4]. The absence of the protease La in this strain reduces the proteolysis of recombinant proteins and favors the generation of larger IBs [5]. Briefly, pJCO46 is a pJLA602 [7] derivative and harbors the *E. coli lacZ* gene, the CT857^o repressor gene and an ampicillin resistance marker. pJVPILAC [4] is a pJCO46 derivative that carries the *VP1* gene of foot-and-mouth disease virus (FMDV) fused upstream of the bacterial *lacZ* gene. In both cases, the expression of recombinant genes is directed by the bacteriophage lambda *p_L* and *p_R* promoters placed in tandem. More details of these constructions have been given elsewhere [5,8].

2.2. Cell culture and conditions for recombinant gene expression

Luria-Bertani (LB) medium [9] was used for cell growth. Aliquots from overnight cultures were inoculated at 1:50 to 50 ml of fresh medium with 100 µg ml⁻¹ ampicillin, in 500-ml shaker flasks, and incubated at 28°C and 250 rpm. When the cultures reached an optical density at 550 nm (OD₅₅₀) of 0.3 units, flasks were transferred to a pre-warmed bath at 42°C to promote recombinant gene expression, and they were further incubated for several hours. β-Galactosidase enzymatic activity was determined as previously described [10].

2.3. Purification of inclusion bodies

Culture samples were taken at different times during the induction period. For the purification of inclusion bodies, cells were lysed and fractionated by a lysozyme-NP40 standard procedure [9]. In a final step, purified inclusion bodies were washed twice with Triton X-100 and resuspended in lysis buffer (Tris-HCl 50 mM, EDTA 1 mM, NaCl 100 mM).

2.4. SDS-PAGE analysis

Purified IBs were boiled for 10 min in denaturing

buffer with 4 M urea [11]. Volumes corresponding to equal cell density were loaded on 3.5% stacking and 7% separation gels. After SDS-PAGE, gels were stained with Coomassie blue. Alternatively, proteins were electroblotted onto nitrocellulose filters to immunodetect β-galactosidase and VPILAC with a rabbit serum elicited against the native enzyme. Rabbit sera elicited against commercial DnaK and GroEL were also used to identify these chaperones. The relative amounts of recombinant proteins were calculated from densitometric analysis of β-galactosidase immunoreactive bands, always within a linear range. To determine the actual concentrations of recombinant proteins, known amounts of commercial, native β-galactosidase (from Boehringer Mannheim) were loaded as standards as previously described [12].

2.5. Image analysis

Samples of the VPILAC-producing cultures were taken at different times, fixed with 0.1% formaldehyde and stored at 4°C until use. Photographs were taken at ×1000 magnification factor, in an Olympus BH2 optical microscope. Cell images were digitized from slides with an ScanJet 4c/T (Hewlett Packard) and further analyzed with the UTHSCSA ImageTool program to calculate volumes of both IBs and cells. Ovoid and cylindrical shapes, both commonly observed in IBs, were considered, and volumes (*v*) were calculated according to the following standard equations:

$$v_{IB} = [4uw^2l]/24$$

$$v_{IB} = (\pi/4)w^2(l - w/3)$$

w and *l* being width and length respectively. For the measure of cell volume, we used the perimeter (*p*) and the area (*a*) of the scanned cell images, since these parameters have been shown to be the most reliable for such a determination [13]. Therefore, *w*, *l* and *v* were calculated according to the previously described [13] equations:

$$l = [p + \sqrt{(p^2 - 16a)}]/4$$

$$w = [p - \sqrt{(p^2 - 16a)}]/4$$

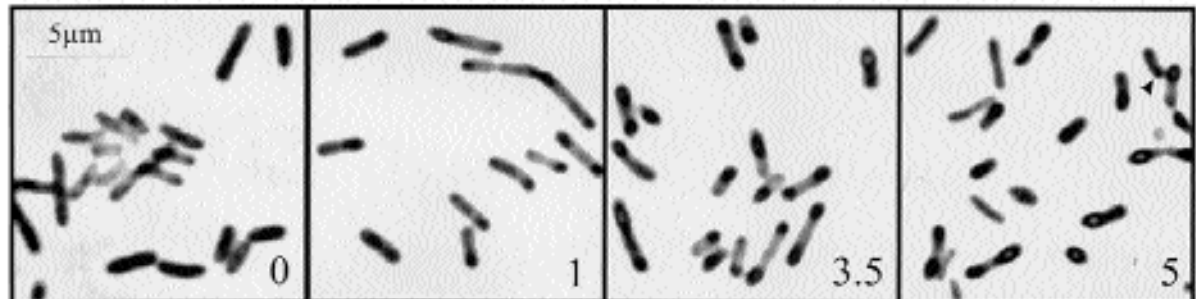


Fig. 1. Micrographs of formaldehyde-fixed BL21 cells producing VP11.AC. Numbers indicate hours after induction. The arrow at time 5 h indicates a small IB. A representative experiment is shown here and also in the following figures unless otherwise indicated.

$$V_{cell} = (\pi/4)w^2(l - w/3)$$

About 100 cell images were processed for each sample.

3. Results and discussion

3.1. Formation of IBs upon recombinant VP11.AC gene expression

Immediately after the temperature up-shift, polar, ovoid cytoplasmic particles were detected by optical

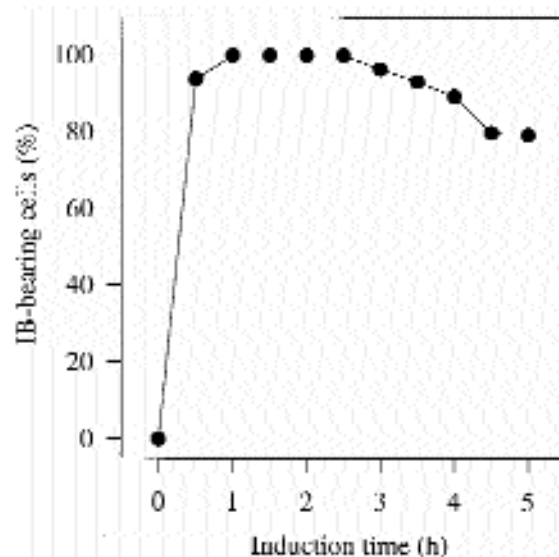


Fig. 2. Percentage of BL21/pVP11.AC cells carrying IBs in thermally induced batch cultures.

microscopy in BL21/pVP11.AC cells after 1 h, being refractile under phase contrast microscopic observation and showing a continuous increase in volume (Fig. 1). These aggregates were not observed in cells producing non-fused β -galactosidase from the parental vector pJCO46 (not shown). The formation of protein aggregates was not exclusive to specific cell sub-populations, since IB-free cells were not detected up to 3 h after recombinant gene expression (Fig. 2). However, from this time on, IBs were absent in an increasing fraction of cells (Figs. 1 and 2), probably caused by the decline of cell biosynthetic capabilities and a decrease in recombinant gene dosage.

3.2. IB composition in VP11.AC-producing bacteria

The time-dependent analysis of protein composi-

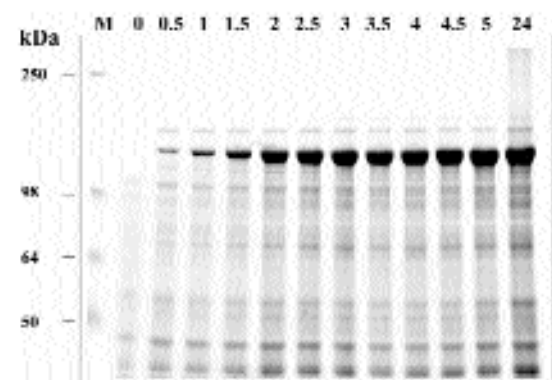


Fig. 3. Coomassie blue stained gel of purified IBs from VP11.AC-producing BL21 cells. Numbers indicate hours after thermal shift.

tion of purified IBs (Fig. 3) revealed the presence of a major band of about 140 kDa, corresponding to the intact form of VPILAC [14]. The amount of recombinant VPILAC in the insoluble cell fraction was undetectable before induction (see time 0 in Fig. 3) but it rapidly increased after the temperature shift at nearly constant rates and at least during 5 h (Figs. 3 and 4A). Despite this constancy, there is a significant variability in the VPILAC aggregation rates in different production experiments (between 4 and 25 ng ml⁻¹ h⁻¹), which among other reasons might be caused by differences in plasmid maintenance and cell density or physiological state during the production process. VPILAC was the major component of IBs, accounting for 80-100% of the total protein (Fig. 4B). This percentage remained almost stable from 2 h on and throughout the production period.

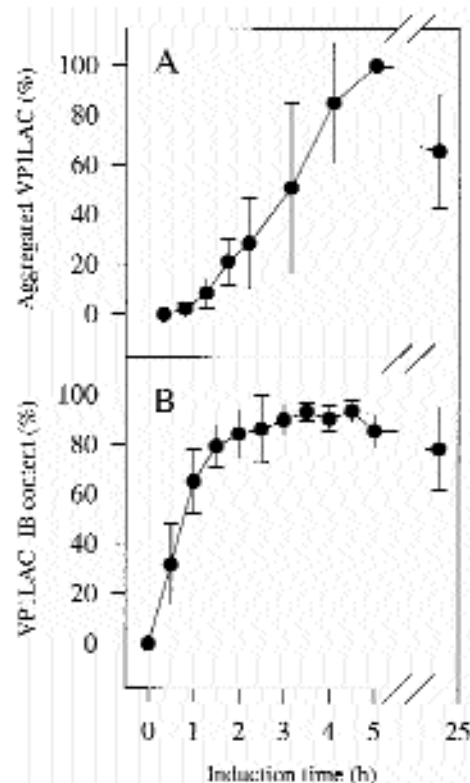


Fig. 4. A: Relative amounts of VPILAC in purified IBs at different times after induction, as estimated from Western blot analysis. Values are referred to those obtained at 5 h, namely 108, 55 and 18 ng ml⁻¹ culture respectively. B: Percentage of VPILAC in the total IB protein. The mean of three independent experiments is shown in both panels.

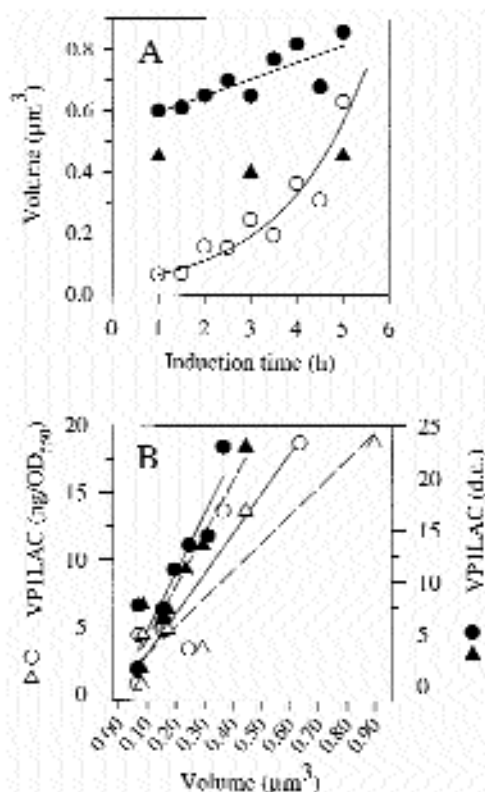


Fig. 5. A: Variation of cell (black circles) and IB volume (white circles) in VPILAC-producing cultures. The mean volume of cells producing a recombinant β -galactosidase is also shown (triangles). B: The amount of VPILAC in densitometric units (black symbols) and in estimated concentration (white symbols) is plotted versus IB volume. Both ovoid (circles, continuous lines) and cylindrical (triangles, dashed lines) shapes were considered for volume determination. Correlation coefficients (r) are higher than 0.92 in all cases.

despite the continuous incorporation of newly synthesized protein. Chaperones DnaK and GroEL were also identified among the minor IB components, always representing less than 5% of the total insoluble protein (not shown). These proteins could have been entrapped in VPILAC aggregates by a direct interaction with misfolded recombinant protein molecules, or alternatively they could have adsorbed unspecifically to pre-formed IBs during the purification process. However, it is noteworthy that although recombinant protein production stimulates synthesis of chaperones and other stress proteins [15] that interact with the inducing proteins [16,17], their presence in IBs is not a general event [18], suggesting

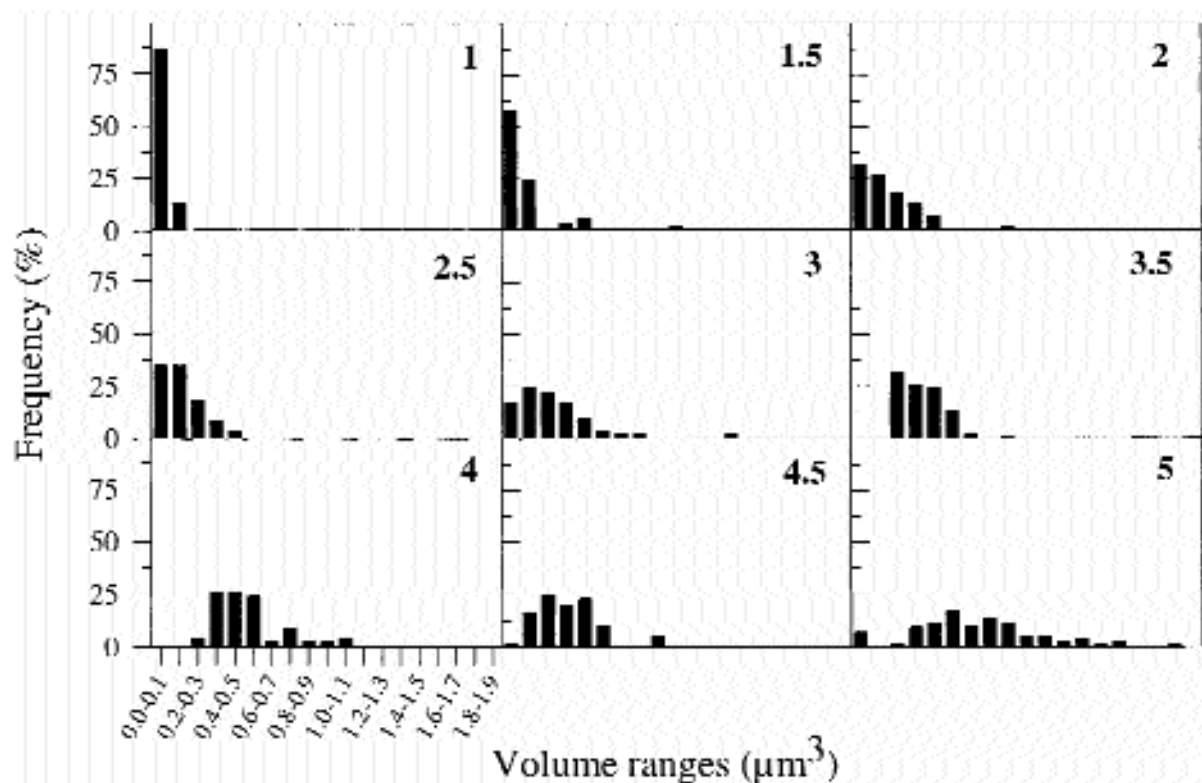


Fig. 6. Spectrum of IB volumes during production of VPILAC. Numbers indicate hours after induction of recombinant gene expression.

a protein-specific co-precipitation of chaperones. The loss of aggregated VPILAC observed between 5 and 25 h (Fig. 4A) might indicate that the IB composition suffers modifications in stationary-phase cells.

3.3. Continuous growth of IBs in producing bacteria

A continuous increase of the average IB volume was observed throughout the synthesis of VPILAC, ranging from about $0.05 \mu\text{m}^3$ after 1 h induction to more than $0.6 \mu\text{m}^3$ after 5 h (Fig. 5A). Although it has been postulated that the IB size depends on the protein amino acid sequence and genetic features of the producer strain [18], the data presented here demonstrate a permanent growth of IBs while the recombinant protein is produced. This suggests that the aggregation rate could be determinant of the observed IB volume at particular stages of formation. The minor enlargement of the average cell volume, from 0.6 to $0.85 \mu\text{m}^3$, which is not observed in

cells producing a recombinant β -galactosidase (Fig. 5A), could be partially accounted for by the growth of cytoplasmic IBs.

Density values of IBs are available for some model proteins [19], but their putative variation during IB formation had not been analyzed. When plotting the amount of insoluble VPILAC versus average IB volume, a linear correlation is observed considering both ovoid and cylindrical IB shapes (Fig. 5B). Since the recombinant protein is the main component of IBs (Fig. 4B), this correlation suggests that VPILAC IB density does not suffer dramatic variations during the building process. The increase in the average IB volume is accounted for by a wider volume distribution in older cell samples (Fig. 6), in which a fraction of small, probably newly generated IBs can still be observed (see also Fig. 1, arrow). The permanent presence of small particles indicates that new aggregates could be continuously generated, despite old ones still growing. By assuming that the largest IBs

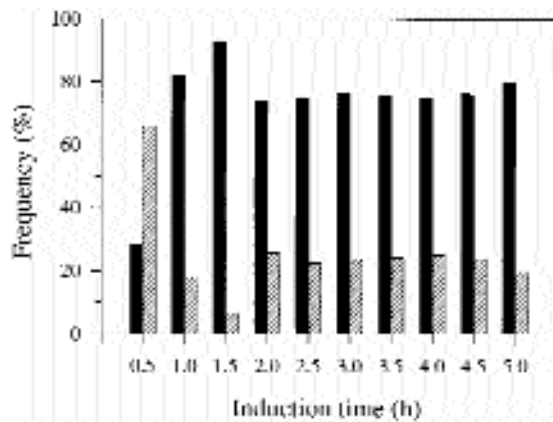


Fig. 7. Distribution of IB number in IB-carrying cells. Numbers indicate hours after induction. Black bars mean one IB per cell, and hatched bars two.

at time 5 h are the oldest ones, an average building rate of $0.4 \mu\text{m}^2 \text{h}^{-1}$ can be estimated.

3.4. Distribution of IBs in producing cells

The distribution of IBs in BL21 producing cells was monitored during the thermal induction of VPILAC gene expression. Up to two IBs were found in IB-bearing cells in all the scanned samples, those carrying only one being the most abundant. Surprisingly, from 1 h on, the percentage of cells carrying two IBs remained extremely constant around 25% (Fig. 7), suggesting that VPILAC aggregation cores are mutually exclusive and that their formation can be prevented by pre-formed ones. The high fraction of cells containing two IBs, observed immediately after induction but not later, might indicate that a minimal amount of already aggregated protein is required for efficient exclusion. A cause for IB interference could be a ribosome concentration by entrapping into aggregation cores, during the precipitation of elongating polypeptide chains. However, the presence of IBs in the totality of cells during almost 3 h (Fig. 2) indicates that ribosome distribution during cell division must not be dramatically unbalanced. Alternatively, the convergence of partially soluble aggregation intermediates into size-dependent nucleation cores might be the cause of the restricted number of IBs. The activation of protease target sites in VPILAC and related proteins, within the insoluble cell fraction [14,20], indi-

cates the existence of misfolded, partially precipitated protein forms that are still reactive before their complete aggregation. These loose forms could only be condensed in large IB cores, thus preventing the formation of multiple IBs in a single cell. Detailed kinetic analyses would be required to deeply elucidate the nature of the exclusive IB formation.

Acknowledgments

We are indebted to J. Checa and V. Ferreres for technical assistance. This work has been supported by Grants BIO95-0801, BIO98-0527 and by the M^o Francesca de Roviralta Foundation. J.L.C. is the recipient of a predoctoral fellowship from MEC.

References

- [1] Marston, P.A.O. (1986) The purification of eukaryotic polypeptides synthesized in *Escherichia coli*. *Biochem. J.* 240, 1–12.
- [2] Rudolph, R. and Lilic, H. (1996) In vitro folding of inclusion body proteins. *FASEB J.* 10, 49–56.
- [3] Spradl, M.A., Wang, D.F.C. and King, J. (1996) Specific aggregation of partially folded polypeptide chains: The molecular basis of inclusion body composition. *Nature Biotechnol.* 14, 1281–1287.
- [4] Yon, J.M. (1996) The specificity of protein aggregation. *Nature Biotechnol.* 14, 1231.
- [5] Corchero, J.L., Viaplana, E., Benito, A. and Villaverde, A. (1996) The position of the intrachain domain can influence solubility and proteolysis of β galactosidase fusion proteins in *E. coli*. *J. Biotechnol.* 48, 191–200.
- [6] Studier, F.W. and Moffat, B.A. (1986) Use of the bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* 189, 113–130.
- [7] Schauder, B., Blocker, H., Frank, R. and McCarthy, J.L.G. (1987) Inducible expression vectors incorporating the *Escherichia coli* *atpE* translational initiation region. *Gene*, 72, 279–283.
- [8] Corchero, J.L. and Villaverde, A. (1998) Plasmid maintenance in *Escherichia coli* recombinant cultures is dramatically, steadily, and specifically influenced by features of the encoded proteins. *Biotechnol. Bioeng.* 58, 625–632.
- [9] Maniatis, D., Fritsch, E.F. and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- [10] Miller, H.H. (1972) *Experiments in Molecular Genetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- [11] Tsurumi, H.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680–685.

- [12] Benito, A., Mateu, M.G. and Villaverde, A. (1995) Improved mimicry of a foot and mouth disease virus antigenic site by a viral peptide display on β -galactosidase surface. *Bio/Technology* 13, 801–804.
- [13] Bloom, J., Venuga, M. and Shephard, J. (1995) Fully automatic determination of soil bacterium numbers, cell volumes, and frequency of dividing cells by confocal laser scanning microscopy and image analysis. *Appl Environ Microbiol* 61, 926–936.
- [14] Corchero, J.L., Cuhari, R., Enfors, S.-O. and Villaverde, A. (1997) Limited in vivo proteolysis of aggregated proteins. *Biochim. Biophys. Res. Commun.* 237, 325–330.
- [15] Rinas, U. (1996) Synthesis rates of cellular proteins involved in translation and protein folding are strongly altered in response to overproduction of basic fibroblast growth factor by recombinant *Escherichia coli*. *Biotechnol. Prog.* 12, 196–200.
- [16] Thomas, J.G. and Bancya, F. (1996) Protein misfolding and inclusion body formation in recombinant *Escherichia coli* cells overexpressing heat shock proteins. *J. Biol. Chem.* 271, 11141–11147.
- [17] Gustavsson, K., Bergman, T., Veide, A. and Enfors, S. O. (1995) In vitro complex-formation between the molecular chaperone DnaK and staphylococcal protein A derivatives produced in *Escherichia coli* and its use in the purification of DnaK. *Biotechnol Appl Biochem* 25, 173–180.
- [18] Bowden, G.A., Paredes, A.M. and Georgina, G. (1991) Structure and morphology of protein inclusion bodies in *Escherichia coli*. *Bio/Technology* 9, 725–730.
- [19] Taylor, G., Hoare, M., Gray, D.R. and Marston, F.A.G. (1986) Size and density of protein inclusion bodies. *Bio/Technology* 4, 253–257.
- [20] Vinplow, F., Felin, F.X., Cuchman, H. and Villaverde, A. (1997) Reversible activation of a cryptic cleavage site within *E. coli* β -galactosidase in β -galactosidase fusion proteins. *Biochim. Biophys. Acta* 1343, 221–226.

RESULTATS

4.2. ELS COSSOS D'INCLUSIÓ SÓN ESTRUCTURES DINÀMIQUES: TRANSICIÓ ENTRE LA FORMA SOLUBLE I INSOLUBLE DE LA PROTEÏNA RECOMBINANT

En un segon treball, es va voler extendre l'estudi del creixement dels cossos d'inclusió utilitzant una altra proteïna, la LACVP1, que tot i ser molt semblant a VP1LAC, ja s'havia vist que era més sensible a la proteòlisi. L'objectiu d'aquest treball era veure si proteïnes amb una seqüència pràcticament igual, tenien la mateixa tendència a l'agregació. Tanmateix, la caracterització dels cossos d'inclusió d'aquestes proteïnes mostra que presenten propietats diferents, tant en morfologia com en composició. Examinant cinèticament el creixement dels cossos d'inclusió es van observar pèrdues de proteïna agregada i la posterior reagregació de la mateixa o dels seus productes de degradació. Això ens suggeria que la proteïna que forma els cossos d'inclusió es troba en un estat transitori entre la seva forma soluble i la insoluble.

RESULTATS



Proteolytic digestion of bacterial inclusion body proteins during dynamic transition between soluble and insoluble forms

María M. Carrió, José L. Corchero¹, Antonio Villaverde^{*}

Institut de Biologia Fonamental and Departament de Genètica and Microbiologia, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain

Received 15 February 1999; received in revised form 21 July 1999; accepted 28 July 1999

Abstract

Inclusion bodies formed by two closely related hybrid proteins, namely VP1LAC and LACVP1, have been compared during their building in *Escherichia coli*. Features of these proteins are determinant of aggregation rates and protein composition of the bodies, generating insoluble particles with distinguishable volume evolution. Interestingly, in LACVP1 and less perceptibly in VP1LAC bodies, an important fraction of the aggregated polypeptide is lost at a given stage of body construction. Stable degradation intermediates of the more fragile LACVP1 are concomitantly found embedded in the bodies. When recombinant protein synthesis is arrested in growing cells, the amount of aggregated protein drops while the amount of soluble protein undergoes a sudden rise before proteolysis. This indicates an architectural plasticity during the *in vivo* building of the studied inclusion bodies by a dynamic transition between soluble and insoluble forms of the recombinant proteins involved. During this transition, protease-sensitive polypeptides can suffer an efficient proteolytic attack and the resulting fragments further aggregate as inclusion body components. © 1999 Elsevier Science B.V. All rights reserved.

Keywords: β -Galactosidase; Misfolding; Protein aggregation; Proteolysis; Inclusion body; Recombinant protein

1. Introduction

Protein aggregation, despite being deeply involved in animal and human pathologies of growing impact, is still poorly understood at the molecular level. Bacterial inclusion bodies (IBs), which are frequently observed during the production of heterologous proteins [1], offer an interesting model to investigate the molecular basis of protein folding and aggregation.

Their formation takes place when misfolded proteins synthesised at high rates escape from the housekeeper proteolytic systems and aggregate around nucleation cores. Although some studies indicate that aggregation intermediates can be partially degraded [2], IB proteins are highly stable and resistant to proteases *in vivo* [1,3]. Therefore, during the production of recombinant proteins, the formation of IBs often permits a better protein recovery than that of soluble protein [3]. The continuous developing of subsequent refolding, product-adapted strategies [4] allows the final recovery of protein in its native conformation.

In this work, we have studied the aggregation kinetics of the recombinant, hybrid protein LACVP1 when produced in *Escherichia coli*. This construct is a β -galactosidase carrying the VP1 protein of foot and

^{*} Corresponding author. Fax: +34-93-5812011;
E-mail: a.villaverde@cc.uab.es

¹ Present address: VETIR, MailStop G 18, Centers for Disease Control, 1600 Clifton Road, Atlanta, GA 30333, USA.

mouth disease virus (FMDV) fused at its carboxy-terminus [5,6]. LACVP1 is more protease-sensitive than the related protein VPILAC, an N-terminal β -galactosidase fusion consisting of the same polypeptides arranged in the alternative positions [5,6]. The kinetic analysis of the *in vivo* LACVP1 IB building process described here shows an important, site-restricted proteolytic activity on LACVP1 during a late phase of IB formation, coincident with a reduction in the average volume of the bodies. A fraction of stable degradation fragments are found retained in purified IB particles, suggesting that proteolysis might take place directly on aggregated protein. Although this possibility cannot be excluded, a deeper investigation of this event prompts us to propose a dynamic protein exchange between soluble and insoluble cell fractions during the building of our model IBs, that under active protein synthesis is displaced towards protein aggregation and volumetric growth of IB particles. This dynamism, supporting an unexpected plasticity of IB architecture, could provide an appropriate molecular context for proteolysis of solubilised IB proteins and further aggregation of the resulting products.

2. Materials and methods

2.1. Bacterial strain, plasmids and culture conditions

The Lon *E. coli* strain BL21 [7] was used for protein production. Plasmids pJLACVP1 and pJVPILAC [5,6] encode two closely related hybrid proteins in which the VP1 capsid protein is fused to either the carboxy- or amino-terminus of β -galactosidase respectively. In these vectors, the lambda lytic promoters direct transcription of both recombinant genes, under the control of the temperature-sensitive C1857 repressor [8]. Luria-Bertani (LB) medium [9] plus 100 μ g/ml ampicillin was used for the culture of recombinant cells. Cultures were grown at 28°C and 250 rpm. At the beginning of the exponential phase (around 0.3 OD₅₅₀ units), flasks were transferred to a pre-warmed bath at 42°C to induce gene expression and they were further incubated at this temperature for 24 h. When necessary, β -galactosidase activity was measured by standard procedures [10]. For some experiments, expression of re-

combinant genes was stopped by transferring the cultures back to shaker baths at 28°C after growing for a period of time at 42°C.

2.2. Sampling, purification and electrophoretic analysis of IBs

Culture samples of 1.5 ml were periodically taken during the production phase. IBs were purified by repeated detergent treatment. Briefly, cells were harvested by centrifugation at 12000 \times g for 15 min and resuspended in 200 μ l of lysis buffer (50 mM Tris-HCl, 1 mM EDTA, 100 mM NaCl, pH 8) containing 30 μ l of 100 mM phenylmethylsulfonyl fluoride and 6 μ l of 10 mg/ml lysozyme. After 30 min of incubation at 37°C in agitation, NP-40 was added to 1%, and the mixture incubated at 4°C under gentle agitation for 30 min. Then, 3 μ l of 1 mg/ml DNase and 3 μ l of 1 M MgSO₄ were added and the mixture incubated again at 37°C for 30 min. After lysis, protein aggregates were separated by centrifugation at 12000 \times g for 15 min. Finally, IBs were washed twice with lysis buffer plus 0.5% Triton X-100, resuspended in lysis buffer and stored until analysis. IBs were boiled for 10 min in denaturing buffer (125 mM Tris-HCl, 6% SDS, 20% glycerol, 10% β -mercaptoethanol, pH 6.8) plus 8 M urea [11] and loaded on 3.5% stacking-7% separation gels. The presence of urea in this buffer is necessary to achieve a complete solubilisation of IB proteins and a correct migration of proteins as clear bands after electrophoresis (not shown). Protein bands were quantified after either Coomassie blue staining of SDS-PAGE gels or immunoblot with appropriate standards. In Western blot we used either an anti- β -galactosidase polyclonal serum or the anti-VP1 monoclonal antibody 3E5 [12]. For simultaneous analysis of protein content in soluble and insoluble fractions, cells from 3-ml samples were harvested by centrifugation at 12000 \times g for 10 min, resuspended in 300 μ l of Z buffer (0.06 M Na₂HPO₄, 0.04 M NaH₂PO₄, 0.01 M KCl, 1 mM MgSO₄, pH 7.2) and disrupted by ultrasonication at 50 W for 1.5 min in an ice bath. Soluble cell fraction was separated by centrifugation at 12000 \times g for 10 min at 4°C and processed for electrophoresis. Insoluble cell fraction was processed for IB purification and further SDS-PAGE as described above.

2.3. Image analysis

Morphological features of LACVPI IBs were determined as described for VPILAC [13]. Briefly, images of producing cells were analysed at 1000× magnification factor with the UTILSCSA ImageTool program. IB volumes were calculated according to the following equation:

$$v_{IB} = [4\pi w^2 l]/24$$

where w and l are width and length of an ovoid particle respectively. About 100 cell images were analysed in each sample.

2.4. Computing procedures

The deduced amino acid sequences for LACVPI and VPILAC were analysed using the GCG (Genetics Computer Group) Wisconsin Package version 9 for secondary structure and protease target site predictions respectively. Programs PeptideStructure and PeptideMap were run for these purposes.

3. Results

3.1. Distinct architecture and protein composition in LACVPI and VPILAC IBs

LACVPI and VPILAC are closely related hybrid proteins in which the same partner domain is fused at either the carboxy- or the amino-terminus of *E. coli* β -galactosidase [5,6]. Both constructs form IBs upon overproduction in *E. coli* BL21 [5] although they show distinct proteolytic sensitivity in vivo, LACVPI being more unstable and yielding a main degradation product similar in mobility to the native β -galactosidase [5]. Since the IB construction process remains essentially unexplored, we have recently investigated the kinetics of IB growth in VPILAC-producing cells [13]. The results obtained indicated a steady accumulation of the recombinant protein and a regular increase of IB volume while the cells remain biosynthetically active. To determine the putative influences of protein stability on the aggregation profile, LACVPI was subjected here to a similar analysis for comparison. Table 1 summarises the main features of both kinds of bodies. As observed, LACVPI IBs remain smaller throughout their growth, although the synthesis rate of both proteins

Table 1
Features of LACVPI and VPILAC IBs produced in *E. coli* BL21

Induction time (h)	IB volume (μm^3)		Recombinant protein IB content (%)		IB-bearing cells (%)		Cells containing 2 IBs (%)	
	LACVPI	VPILAC	LACVPI	VPILAC	LACVPI	VPILAC	LACVPI	VPILAC
0	nd	nd	nd	nd	nd	nd	nd	nd
0.5	nd	nd	5.48/18.67 ^b	32.1	nd	93.7	nd	65.6
1.0	nd	0.065 ^a	10.98/31.02	51.4	47.0	100	2.8 ^c	18.0
1.5	nd	0.069	18.53/55.25	73.7	nd	100	nd	6.3
2.0	nd	0.157	18.72/49.71	79.1	53.4	100	3.0	25.5
2.5	0.007	0.153	20.13/55.43	72.3	69.4	100	3.6	22.3
3.0	0.013	0.245	24.71/58.06	87.4	73.2	96.4	4.0	23.4
3.5	0.056	0.191	27.87/64.93	91.1	68.5	93.2	5.4	21.0
4.0	0.058	0.364	28.97/62.97	91.5	65.4	89.3	3.5	25.0
4.5	0.043	0.310	10.90/63.66	90.5	73.5	79.8	3.8	23.6
5.0	0.044	0.633	23.14/59.64	90.4	67.0	79.3	3.2	20.0
24.0	0.144	0.720	18.80/61.69	78.1	100	75.7	24.8	37.0

nd: not detectable.

^aThe given value is the median of the IB volume distribution.

^bFor LACVPI IBs, both the intact form and the total recombinant protein including the β -galactosidase- and VP1-like fragments have been considered.

^cThis percentage is given considering the total cell number (including both IB-bearing and IB-free cells).

is comparable [5]. The rapid proteolysis of LACVP1 before aggregation [5] may account for the smaller size of the resulting IBs. Interestingly, protein composition of LACVP1 and VP11AC IBs differs greatly. VP11AC is the main IB component accounting for 80–90% of the total protein, this percentage remaining stable throughout protein accumulation. However, LACVP1 represents less than 30% of IB protein, showing an important time-dependent evolution with a peak at about 4 h. If also considering detectable LACVP1 degradation intermediates present in IBs, the relative values of total recombinant protein are still lower than in VP11AC IBs, always below 65%. This suggests a higher heterogeneity in protein composition of LACVP1 IBs. Other parameters describing the incidence of IBs in the producing cultures are also compared in Table 1, in general being representative of a lower efficiency of LACVP1 aggregation. Note for example the lower number of IBs per cell compared to those of VP11AC. Intriguingly, in both LACVP1 and

VP11AC, a transient decrease in IB volume is observed at different times after induction, appearing more stable in LACVP1 bodies.

3.2. Proteolysis of IB LACVP1 protein

The low presence of recombinant protein in LACVP1 IBs was deeply investigated by SDS-PAGE of urea-denatured IBs, followed by Western blot analysis. We considered the intact form of LACVP1 as well as the main degradation intermediates, since they could be also incorporated to growing IBs. Both fragments, β -galactosidase- and VP1-like proteins are antigenically indistinguishable from the partners FM10V VP1 and *E. coli* β -galactosidase respectively (not shown), showing apparent molecular masses of about 18 and 117 kDa (Fig. 1). Although additional hydrolysis steps cannot be excluded, these fragment sizes are compatible with a discrete cleavage at the linker region. Western blots of LACVP1 IBs from a time course experiment are

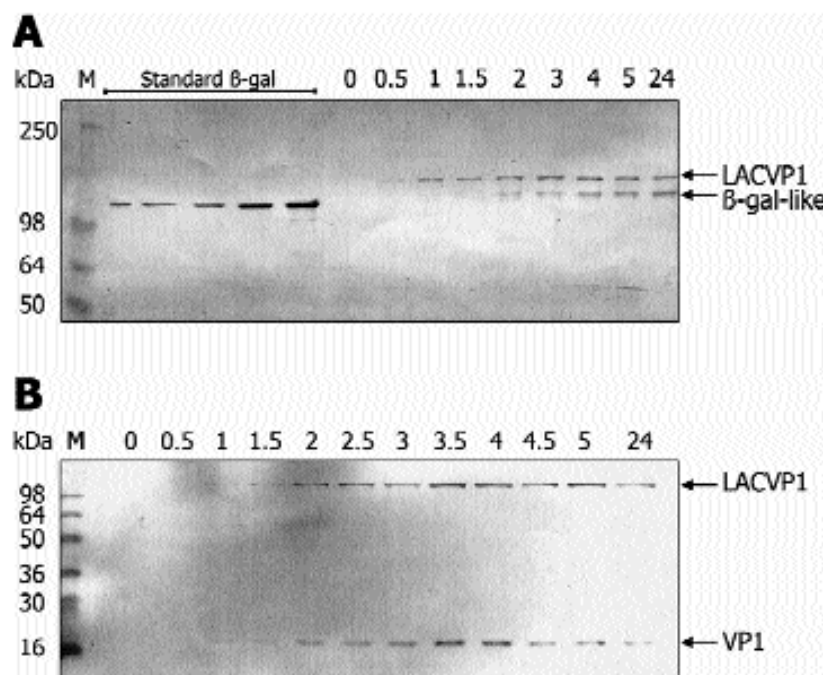


Fig. 1. Time dependent Western blot analysis of purified LACVP1 IBs. An anti β galactosidase serum (A) and the anti VP1 monoclonal antibody 3F5 (B) were used for immunodetection. Standards of pure β -galactosidase used for quantification are also shown. Numbers indicate hours after induction of recombinant gene expression.

shown in Fig. 1, and the estimated amounts of the identified components depicted in Fig. 2A. As occurs in VPILAC [13], the amount of IB LACVPI increases in a time-dependent fashion upon induction of gene expression. However, the apparent accumulation rate slightly declines after 3 h, and the accumulated amount of full-length LACVPI in isolated bodies drops from 4 h on. Coincident with the decline of the slope representing the intact LACVPI, there is an important increase in the amounts of β -galactosidase-like and VPI-like fragments. The evolution of intact IB LACVPI suggests that this protein can be partially lost from the growing IBs, and that this could be related to the decrease in IB average volume presented in Table 1, between 4 and 5 h after induction. On the other hand, the accumulation of the shortened forms would indicate that the full-length form of LACVPI suffers a site-specific digestion, giving fragments similar to those observed in the cell soluble fraction [5]. The intriguing increase of these fragments in purified IBs together with the decrease of the full-length form could indicate that the cleavage reactions could take place directly on protein associated with IBs. The comparison between the molar amounts of these polypeptides (Fig. 2B), however, suggests that a fraction of the degradation fragments derive from aggregation of soluble forms, since the digestion of aggregated LACVPI cannot account for the β -galactosidase-like accumulated from 3 h on. The steady increase of IB β -galactosidase-like fragment from 5 to 24 h also suggests its

stability in IBs, and this fact may account for the increase of IB volume within this period (Table 1) despite the loss of intact LACVPI. The accumulation kinetics of IB LACVPI is consistent when comparing different experiments (Fig. 2C), showing a peak at 4 h. A similar analysis of the related VPILAC protein does not show any perceptible decline up to 5 h during IB formation (Fig. 2C), although about 20% of the IB full-length protein is also lost between 5 and 24 h.

3.3. Dynamics of IB composition in the absence of recombinant protein synthesis

Results presented in Fig. 2 and Table 1 demonstrate a loss of IB components during *in vivo* body building. In addition, in the case of LACVPI, it seems that a limited cleavage is associated with protein release, but from the presented data it was not possible to discriminate if the proteolytic attack takes place over IB-associated protein or released polypeptides. While in the first case partially aggregated IB protein could expose protease target sites to the solvent at the IB interface, in the second possibility, degradation of LACVPI would imply protein solubilisation followed by digestion and further precipitation in IBs. In an attempt to discriminate between these possibilities, we have studied the *in vivo* evolution of already formed LACVPI and VPILAC IBs, after removing the inducing signal for transcription of recombinant genes. Fig. 3 shows the IB protein

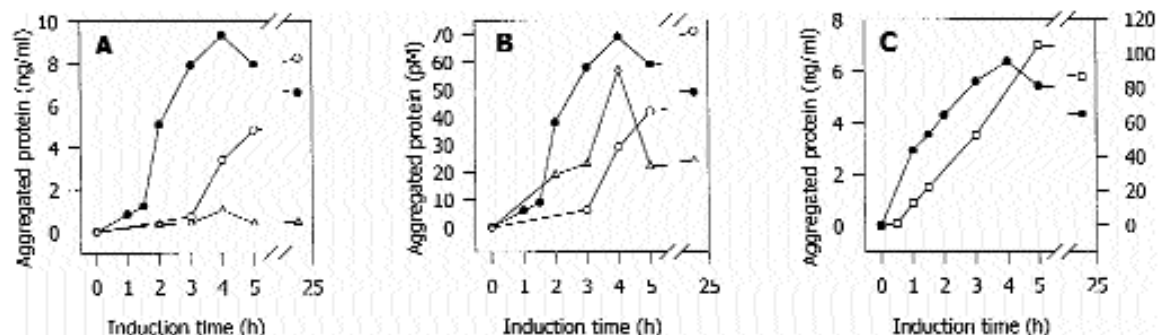


Fig. 2. Accumulation of recombinant proteins and protein fragments in isolated IBs from a growing culture. (A) The amounts of LACVPI (●) and β -galactosidase-like (○) and VPI-like (△) fragments are shown. (B) The same data are represented as molar amounts. (C) The accumulation of IB LACVPI from a different experiment (●, left scale) is compared with IB VPILAC (□, right scale).

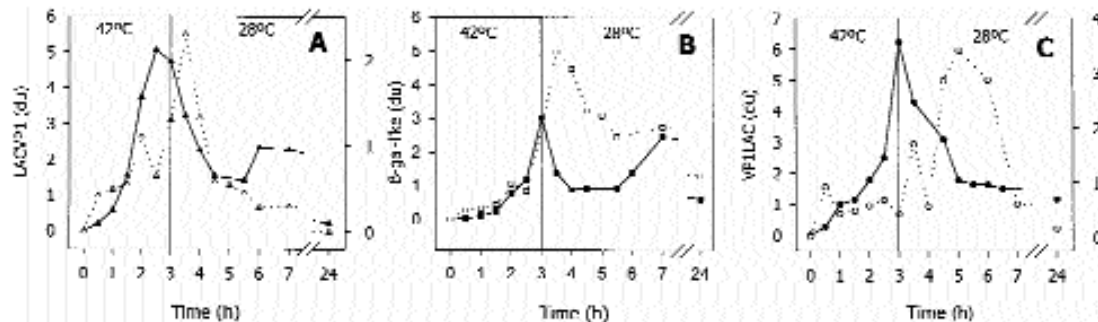


Fig. 3. Evolution of protein amounts of either soluble fraction (open symbols, right scale) and purified IBs (closed symbols, left scale) for intact LACVP1 (A), LACVP1-derived β -galactosidase-like fragment (B) and intact VP1LAC (C), as determined by densitometric analysis of Western blots. Vertical bars indicate the temperature down-shift in the bacterial cultures. The densitometric unit (du) values of soluble and IB fractions, as obtained from different blots, are not quantitatively comparable.

composition before, during and after the transition from productive to non-productive cell status for both kinds of bodies. After temperature decrease, a drop of intact LACVP1 (Fig. 3A), β -galactosidase-like LACVP1 fragment (Fig. 3B) and VP1LAC (Fig. 3C) takes place. The negative evolution of the IB protein content in biosynthetically competent cells, once the transcription from the recombinant promoters has been arrested, strongly suggests that the construction of these IBs is a dynamic process involving simultaneous incorporation and release of protein. Then, the transition between soluble and insoluble forms would be unbalanced towards protein aggregation when cells are actively overproducing recombinant protein, but this situation can be inverted in the absence of specific recombinant synthesis (Fig. 3) or in ageing cultures (Fig. 2).

For the experiments shown in Fig. 3, the soluble cell fraction was also monitored. In VP1LAC-producing cells, the concentration of the fusion protein found in the soluble cell fraction rises coincident with protein loss in purified IBs (Fig. 3C), confirming the presumed transition from aggregated to soluble forms. For LACVP1, the amounts of both the full-length fusion and the β -galactosidase-like fragment increase after arrest of recombinant synthesis. The rise of the intact LACVP1 under these conditions proves a release of undigested IB protein such as also occurs for VP1LAC. Although *in situ* cleavage of LACVP1 IB proteins cannot be excluded, these observations favour the hypothesis that at least an important fraction of intact LACVP1 is solubilised from IBs, proteolytically digested in solution and a

fraction of resulting fragments further aggregated as IB components.

4. Discussion

Failure of overproduced proteins in achieving their native conformation allows availability of protease target sites to the protease cell machinery [14]. Misfolded polypeptides escaping from degradation aggregate as amorphous, insoluble particles in bacterial compartments, namely cytoplasm or periplasm [1,13,15–17]. Cytoplasmic IBs grow steadily in biosynthetically competent cells, around nucleation cores whose formation is mutually exclusive [13]. While it is widely recognised that IB protein is stable in front of cell proteases, the results presented here prove that protein LACVP1, a protease-sensitive β -galactosidase fusion [5], is not irreversibly protected by aggregation as IBs, since a fraction of the recombinant protein is cleaved *in vivo* by a target site within or close to the hinge region. The kinetics presented in Fig. 2A,C shows that the site-restricted proteolytic attack becomes apparent in ageing, slowly dividing cells after several hours of IB growth. This is coincident with a partial protein loss from IBs resulting in a detectable reduction of the average IB volume (Table I and Fig. 2). This digestion generates two detectable fragments, the larger one being highly stable and retained in the aggregates. Protein VP1LAC is also lost from IBs, but later in the building process (Fig. 2C). No degradation products of this protein are found in IBs up to 24 h (not shown),

confirming the higher stability of this fusion [5]. A large number of protease target sites predicted in the highly hydrophilic hinge region of LACVP1, but absent in that of VP11AC (not shown), could account, at least partially, for this distinct proteolytic stability.

A kinetic analysis of IB protein composition after stopping recombinant gene transcription (Fig. 3) permits a clear monitoring of *in vivo* release of intact (as well as degraded) protein from both LACVP1 and VP11AC IBs, the solubilised protein being transiently found in the soluble fraction before further degradation. The growth kinetics of IBs during recombinant gene overexpression [13] can be seen as a balance between protein aggregation and solubilisation, which is naturally displaced towards protein release in ageing cells (Fig. 2 and Table 1), in the absence of synthesis of IB components (Fig. 3). This dynamic plasticity of our model IBs concerning protein permanence is an unreported event that deserves further analysis because of its possible biotechnological impact. The refolding potential of small heat shock proteins in the context of a multichaperone network [18] could provide a rationale for this intriguing finding.

The observed proteolysis of IB LACVP1 (Fig. 2) seems to take place during its release from the bodies, over soluble or partially solubilised polypeptide chains, giving fragments that can aggregate back into bodies. However, the results presented in Figs. 2 and 3 do not exclude direct proteolytic attack of polypeptides on the IB interface, by partial solvent exposure of target peptides. Since *in vivo* site-limited proteolysis of non-IB aggregated proteins has been reported [2], a restricted, but still relevant availability of protease target sites of IB proteins at the IB interface cannot be completely excluded.

Acknowledgements

We are indebted to J. Checa and V. Ferreres for technical assistance and to Domingo and E. Brocchi

for anti-VP1 monoclonal antibodies. This work has been supported by Grant BIO98-0527 (CICYT) and by the M^o Francesca de Roviralta Foundation. J.L.C. was the recipient of a predoctoral fellowship from MEC, Spain.

References

- [1] F.A.O. Marston, *Biochem. J.* 240 (1986) 1–12.
- [2] J.L. Corchero, R. Cubarsi, S.-O. Enfors, A. Villaverde, *Biochem. Biophys. Res. Commun.* 237 (1997) 325–330.
- [3] H. Hellebust, M. Marby, L. Abrahamson, M. Uhlen, S.-O. Enfors, *BioTechnology* 7 (1989) 165–168.
- [4] H. Rogl, K. Kossmund, W. Kahlbrandt, I. Collison, *FEBS Lett.* 452 (1998) 21–26.
- [5] J.L. Corchero, E. Viaplana, A. Benito, A. Villaverde, *J. Biotechnol.* 48 (1996) 191–200.
- [6] J.L. Corchero, A. Villaverde, *Biotechnol. Bioeng.* 58 (1998) 625–632.
- [7] F.W. Studier, R.A. Moffat, *J. Mol. Biol.* 189 (1986) 113–130.
- [8] A. Villaverde, A. Benito, E. Viaplana, R. Cubarsi, *Appl. Environ. Microbiol.* 59 (1993) 3485–3487.
- [9] D. Maniatis, J. Sambrook, E.F. Fritsch, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1982.
- [10] J.H. Miller, *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1972.
- [11] U.K. Laemmli, *Nature* 227 (1970) 680–682.
- [12] I. Capneri, E. Brocchi, F. De Simone, G.F. Punina, G.F., in: Report of a Session of the Research Group of the Standing Technical Committee of the European Commission for the Control of Foot-and-Mouth Disease, Food and Agriculture Organisation of the United Nations, Brescia, 1984, pp. 32–39.
- [13] M.M. Carrió, J.L. Corchero, A. Villaverde, *FEBS Microbiol. Lett.* 169 (1998) 9–15.
- [14] E. Viaplana, J.X. Feliu, J.L. Corchero, A. Villaverde, *Biochim. Biophys. Acta* 1343 (1997) 221–226.
- [15] G.A. Bowden, A.M. Paredes, G. Georgiou, *BioTechnology* 9 (1991) 725–730.
- [16] G. Taylor, M. Hoare, D.R. Gray, F.A.O. Marston, *BioTechnology* 4 (1986) 553–557.
- [17] J.G. Thomas, F. Bancys, *J. Biol. Chem.* 271 (1996) 11141–11147.
- [18] L. Veinger, S. Diamant, J. Buchner, P. Goloubinoff, *J. Biol. Chem.* 273 (1998) 11032–11037.

