Native folding of aggregation-prone recombinant proteins in *Escherichia coli* by osmolytes, plasmid- or benzyl alcohol–overexpressed molecular chaperones

Ario de Marco,[1] Laszlo Vigh,[2] Sophia Diamant,[3] and Pierre Goloubinoff[4]

[1]Protein Expression Unit, European Molecular Biology Laboratory, Meyerhofstrasse 1, D-69117, Heidelberg, Germany
[2]Institute of Biochemistry, Biological Research Center of the Hungarian Academy of Sciences, PO Box 521, H-6701 Szeged, Hungary
[3]Silberman Institute of Life Sciences, The Hebrew University of Jerusalem, 9404 Jerusalem, Israel
[4]Department of Plant Molecular Biology, Lausanne University, CH-1015 Lausanne, Switzerland

Abstract When massively expressed in bacteria, recombinant proteins often tend to misfold and accumulate as soluble and insoluble nonfunctional aggregates. A general strategy to improve the native folding of recombinant proteins is to increase the cellular concentration of viscous organic compounds, termed osmolytes, or of molecular chaperones that can prevent aggregation and can actively scavenge and convert aggregates into natively refoldable species. In this study, metal affinity purification (immobilized metal ion affinity chromatography [IMAC]), confirmed by resistance to trypsin digestion, was used to distinguish soluble aggregates from soluble nativelike proteins. Salt-induced accumulation of osmolytes during induced protein synthesis significantly improved IMAC yields of folding-recalcitrant proteins. Yet, the highest yields were obtained with cells coexpressing plasmid-encoded molecular chaperones DnaK-DnaJ-GrpE, ClpB, GroEL-GroES, and IbpA/B. Addition of the membrane fluidizer heat shock–inducer benzyl alcohol (BA) to the bacterial medium resulted in similar high yields as with plasmid-mediated chaperone coexpression. Our results suggest that simple BA-mediated induction of endogenous chaperones can substitute for the more demanding approach of chaperone coexpression. Combined strategies of osmolyte-induced native folding with heat-, BA-, or plasmid-induced chaperone coexpression can be thought to optimize yields of natively folded recombinant proteins in bacteria, for research and biotechnological purposes.

INTRODUCTION

The expression of recombinant proteins in *Escherichia coli* has become a prevalent method to recover large amounts of functional proteins for research, biotechnology, and pharmaceutics. Despite the simplicity of this approach, the yields of soluble and correctly folded, biologically active proteins are often hampered by misfolding and aggregation events and by degradation (Hanning and Mak-

Correspondence to: Pierre Goloubinoff, Tel: 41-21-692-4232; E-mail: Pierre.Goloubinoff@unil.ch.
Correspondence to: Ario de Marco, Tel: 49-6221 387 448; E-mail: ario.demarco@embl.de.

rides 1998). Depending on various factors such as the intrinsic nature of the recombinant protein, the rate of protein expression, the cellular concentrations of protein folding intermediates, the viscosity, and the temperature of the folding environment, compact large protein aggregates known as inclusion bodies may form. Inclusion bodies have long been considered as disordered, concentrated bodies of misfolded and aggregated nonnative proteins, some times found associated with molecular chaperones, such as the inclusion body binding proteins IbpA/B, DnaK, GroEL, and ClpB, and with proteases, such as Lon, ClpA/X/P, FtsH, and HslUV (Georgiou and Valax 1996; Tomoyasu et al 2001). Despite some impurities, nonnative recombinant proteins in inclusion bodies are highly concentrated and nearly pure, encouraging in

vitro refolding protocols to recover soluble native recombinant proteins from chemically solubilized inclusion bodies (Speed et al 1996; Middelberg 2002). Although the isolation and chemical solubilization of inclusion bodies are simple, protocols for subsequent in vitro refolding are often time consuming and laborious. Rarely do they produce high yields of soluble proteins, only a minute fraction of which are natively refolded proteins (Middelberg 2002). Recently, inclusion bodies were shown to have both amorphous and partially ordered structures. Different levels of organization were described, in which both amyloid fibrils and nativelike proteins coexist (Carrió et al 2005). The dynamic relationship between aggregates and soluble protein fractions has been described and exploited to improve the yield of soluble recombinant proteins by forcing the equilibrium toward the soluble state (Carrió and Villaverde 2001; de Marco and De Marco 2004; Schrödel and de Marco 2005).

Some eukaryotic proteins cannot correctly fold in *E coli* because the bacterial cytoplasm does not provide proper redox conditions, necessary posttranslational modifications, and optimal folding environment in terms of pH and ions or the presence of specific lipids (Rogl et al 1998). In addition, foldases, such as peptidyl prolyl isomerases (PPI) and protein disulfide isomerases (PDI), and molecular chaperones IbpA/B, DnaK, DnaJ, GrpE, GroEL, GroES, and ClpB may be necessary to passively prevent and actively reverse misfolding and aggregation events that may accidentally take place on the native-folding pathway of recombinant proteins in the foreign environment of the bacterial cell (Baneyx and Mujacic 2004). Similarly, single subunits of protein complexes may be prone to aggregation when expressed in the absence of their natural partners (Tang 2001). Unexpectedly, aggregation is also often observed with overexpressed bacterial proteins although they clearly have the capacity to natively fold in a natural bacterial context. Anfinsen (1972) demonstrated that under ideal conditions, proteins may natively fold without the assistance of external factors but the frequency of off-pathway misfolding and of consequent ''accidental'' aggregation events tends to increase with the protein concentration. Plasmid-encoded recombinant proteins expressed from very strong promoters can accumulate up to 50% of the total protein in the cells. Thus, under extreme unnatural conditions, such as massive overexpression of a single recombinant protein in the cell, the posttranslational folding machinery made of molecular chaperones that can normally prevent and correct protein misfolding, may become rapidly overwhelmed. To overcome this problem, protocols were applied to reduce the rate of protein synthesis and folding by lowering the growth temperature and the concentration of the inducer, by increasing the concentration of viscous osmolytes (Diamant et al 2003) or by increasing the levels of

molecular chaperones by plasmid coexpression (Goloubinoff et al 1989b; de Marco and De Marco 2004).

The solubility of recombinant proteins can also be improved by fusions with highly soluble partners that slow down the folding time and, as in the case of DsbA and DsbC, may even apply a direct foldase activity (Zhang et al 1998; Jurado et al 2002). However, while improving solubility, fusion protein may not necessarily improve native folding and often form nonnative soluble aggregates that are difficult to distinguish from the native soluble proteins of interest (Sachdev and Chirgwin 1999; Nominé et al 2001a).

The systematic comparison of several cloning and expression conditions remains the major approach to obtain optimal amounts of natively folded recombinant proteins from bacteria (Vincentelli et al 2003). Possible testable conditions are infinite and there is a need for protocols enhancing the natural posttranslational mechanisms of the *E coli* cell, osmolyte- and chaperone based, to favor correct folding, prevent misfolding and aggregation, and actively convert stable protein aggregates into nativelike proteins.

In this study, using typical folding-recalcitrant proteins, we compared the effect of salt-induced osmolyte accumulation and of plasmid- or chemically induced expression of molecular chaperones. Whereas osmolytes were significantly but mildly efficient, the combination of osmolyte and chaperone accumulation by plasmid coexpression or treatment with benzyl alcohol (BA) was most effective in expressing nativelike recombinant proteins, for 6 of the 8 different folding-recalcitrant proteins tested.

MATERIAL AND METHODS

Plasmid constructs and cell cultures

The sequences corresponding to O36 (developmental protein cactus from *Anophele gambiae*) and E8R (NP 063710) were cloned in the pET24d-modified vector pETM60 (for more information about the features of the pETM vectors: http://www-db.embl.de/jss/servlet/de.embl.bk.emblGroups.EmblGroupsOrg/g_143). In pETM60, the proteins included a N-term His-tag and a fusion to bacterial NusA protein for improved solubility. Another pET24d-modified vector (pETM43) expressed a His-tagged maltose-binding protein (MBP). The proteins O15 (CAB 66874), O46 (CAD 97928), and N34 (CAB 66539) were expressed using a pETM20 vector in which the target proteins are fused to a His-tag and to thioredoxine. The proteins N61 (AAH 32681) and O82 (CG2094) were expressed as glutathione-s-transferase (GST) fusions using a commercial pGEX-6 vector (Amersham Biosciences, Freiburg, Germany).

Plasmids were transformed in BL21(DE3) cells. The effect of the coexpression of recombinant chaperones was

analyzed using BL21(DE3) cotransformed with different combinations of chaperone-harboring plasmids. The genes encoding for DnaK, DnaJ, GrpE, ClpB, and GroES-GroEL were carried into the plasmids pBB540 (pSC101 origin, chloramphenicol resistance) and pBB542 (p15A origin, spectinomycin resistance), GroES-GroEL into pBB541 (p15A origin, spectinomycin resistance) (Tomoyasu et al 2001), and the small chaperones IbpA/B were cloned in a ColE origin plasmid (ampicillin resistance). The use of specific promoters and of plasmids with different copy number enabled to tune the expression levels of the chaperones and to obtain a ratio among them resembling the cellular ratio.

Single colonies from transformed strains were used to inoculate 3 mL of Luria-Bertani medium. Standard liquid cultures were initially incubated at 37°C. When cultures reached 0.6 OD_{600}, protein expression was induced for 2 hours with 0.4 mM IPTG or for 20 hours at 20°C with 0.1 mM IPTG. Any other variation, including culture times, temperature modifications, and conditions used for the chemical additions, is described case by case in the results. Thirty minutes before IPTG addition, betaine and K-glutamate were added to a final concentration of 5 mM together in the presence of a final concentration of 0.5 M NaCl in the medium. Twenty minutes before IPTG induction, BA was added to a final concentration of 10 mM to bacterial cultures grown at 20°C.

Protein purification, yield evaluation, Western blot, dynamic light scattering, and fluorimetric assay

Frozen bacterial pellets (100 mg) were resuspended in 350 mL of 20 mM K-phosphate buffer, pH 7.8, 2 mM phenylmethylsulfonyl fluoride, 0.05% Triton X-100, 20 mg/mL deoxyribonuclease, 2 mM $MgCl_2$, and 1 mg/mL lysozyme and incubated on ice for 30 minutes, with periodic stirring. The suspension was sonicated in water for 5 minutes, pelleted in a minifuge, the supernatant was added to 13 mL of prewashed Talon magnetic beads (Dynal AS, Oslo, Norway) for Ni-affinity purification. After incubation for further 30 minutes under agitation, the supernatant was removed. Beads were washed for 30 minutes with 20 mM K-phosphate buffer, pH 7.8, 300 mM NaCl, 15 mM imidazole, 8% glycerol, 0.1% Triton X-100, and later with phosphate-buffered saline (PBS) plus 0.05% Triton X-100. Finally, His-tagged proteins were eluted either in the presence of 250 mM imidazole or boiled in sodium dodecyl sulfate sample buffer. Samples were loaded onto a sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) using an Amersham minigel system. Proteins were detected after coloration with Simply Blue Safestain (Invitrogen, Karlsruhe, Germany) as per manufacturer's instructions, and the gels were recorded using a Umax Astra 4000U scanner. Large-

scale purifications were performed using the same buffers and HiTrap chelating columns (Amersham) prepared with $CoCl_2$. Western blot analysis was performed as described previously (de Marco et al 1999), using an Immobilon P membrane (Millipore, Cork, Ireland), an ECL™ detection kit (Amersham), and a Biomax MR imaging film (Kodak, Chalone-sur-Saône, France). The primary antibodies against DnaK, GroEL, and ClpB were a gift from Drs B. Bukau and E. Deuerling. Trypsin digestion of total soluble fractions was with pancreatic bovine trypsin of sequencing grade (Roche Applied Science, Mannheim, Germany), according to the manufacturer's instructions.

Dynamic light-scattering measurements were performed in 1- × 1-cm cell with a DynaProMS instrument (DynaPro, Santa Barbara, CA, USA) maintained at 20°C and equipped with a Dynamics V5 software (Protein Solution, Berkshire, UK) for the statistic analysis. The fluorimetric assay for the determination of the aggregation state of the purified proteins (Nominè et al 2001b) was performed at 25°C, using an AB2 Aminco luminescence spectrometer equipped with SLM software.

RESULTS

Soluble proteins are not necessarily natively folded

Measuring enzymatic activities or fluorescence vs light scattering are most reliable straightforward methods to estimate the native foldedness of recombinant proteins. However, activity or fluorescence is not always applicable to structural proteins or to individually engineered protein domains. Protein solubility after lysate centrifugation and filtration is the most common criterion used to estimate the partitioning between misfolded insoluble forms and soluble natively folded forms of recombinant proteins. But, some inactive low-mass aggregates are filtrated and can remain soluble after ultracentrifugation (Diamant et al 2000; Nominé et al 2001b; Ben-Zvi and Goloubinoff 2002; Schrödel and de Marco 2005). This implies that additional criteria must be applied to distinguish soluble native from soluble nonnative proteins.

The formation of soluble aggregates is particularly frequent in the case of fusions composed of aggregation-prone target proteins with very soluble partners (Sachdev and Chirwin 1999; Nominé et al 2001a). A fluorimetric analysis has been successfully applied to compare the aggregation complexity of recombinant fusion proteins recovered after different treatments (Nominé et al 2001b; Schrödel and de Marco 2005; Schrödel et al 2005). Aggregation may reduce the accessibility of some protein epitopes, such as His-tag tails, with a strong effect on the binding efficiency of the protein to a metal ion–activated resin. Therefore, affinity purification may also be used as
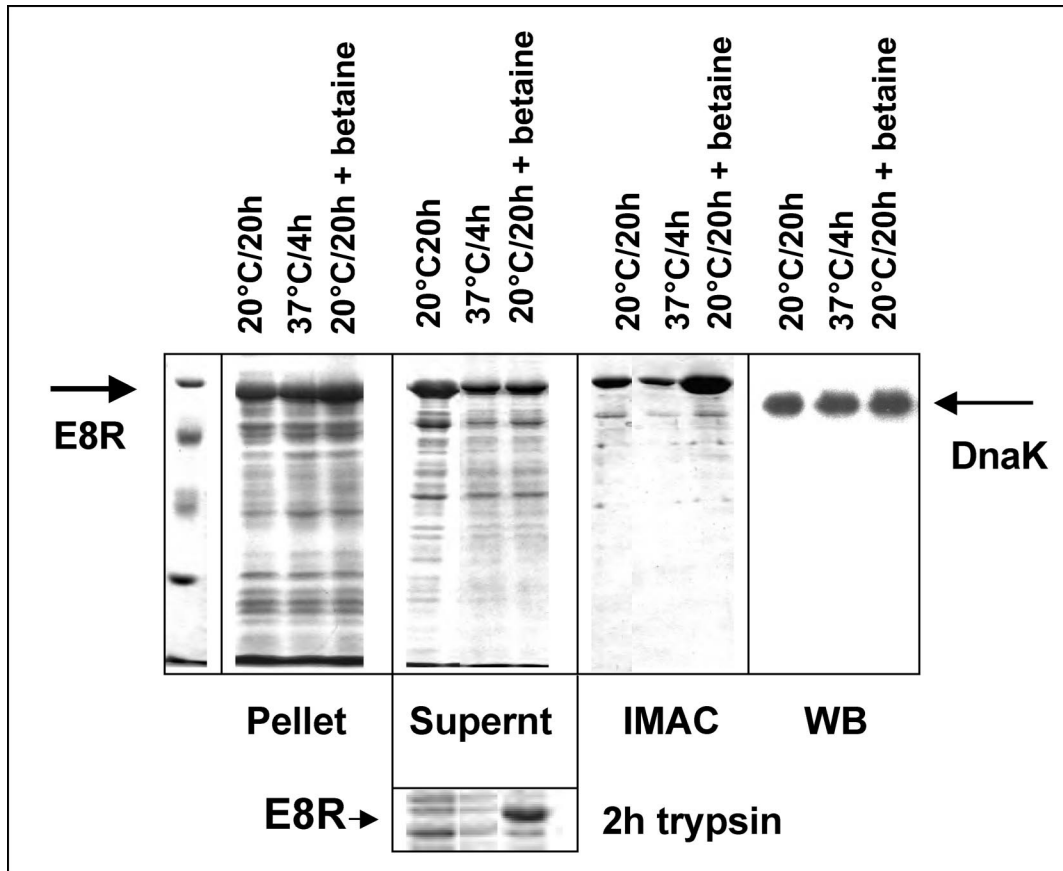
Fig 1. Affinity purification and protease sensitivity to estimate native folding after salt + osmolyte treatment. Cells expressing plasmid-encoded E8R-NusA fusion protein were grown in the presence of IPTG at 20°C for 20 hours or at 37°C for 4 hours, as indicated. Salt treatment: 30 minutes before the IPTG induction, 0.5 M NaCl, and 5 mM betaine were supplemented to the medium. After cell lysis, homogenates were centrifuged. The insoluble pellets (left panel), supernatants (central left panel), or the ion metal affinity–purified fractions (central right panel) were separated on SDS-PAGE and stained by Coomassie blue. Right panel: total homogenates separated by SDS-PAGE and immunoblotted with DnaK antibodies. Lower panel: supernatant fractions (100 μg) were incubated with trypsin (5 μg) for 2 hours at 24°C before separation by SDS-PAGE and Coomassie blue staining. SDS-PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis.

a means to discriminate soluble nativelike proteins from soluble nonnative aggregates, behaving, respectively, as mono- and multidispersed species in dynamic light-scattering measurements (data not shown).

In this study, we first tested the possibility of discriminating between soluble native and soluble nonnative species of an overexpressed recombinant viral protein, E8R, that normally accumulates entirely into the insoluble fraction of the inclusion bodies (data not shown). When expressed fused in frame with bacterial NusA (De Marco et al 2004), 25–40% of the E8R was found in the soluble fraction after 30 minutes centrifugation at $100\,000 \times g$, depending on the growth conditions (Fig 1). The bacteria were grown with inducer either at 20°C for 20 hours or at 37°C for 4 hours, with and without addition of 0.5 M NaCl and 5 mM glycine betaine (betaine) to cause accumulation of betaine in the bacteria (Diamant et al 2003). Western blot analysis showed that NaCl treatment did not induce significant variations in the expression of the total

E8R protein or of the molecular chaperone DnaK (Fig 1). At 20°C, the highest yields of soluble protein was obtained without treatment. Yet, the highest amount of immobilized metal ion affinity chromatography (IMAC)–purified E8R protein was obtained in cells treated with salt + betaine (10% of the total E8R). In contrast, cells grown at 37°C expressed nearly as much soluble but significantly less IMAC-purifiable protein (1–2% of the total E8R).

Dynamic light-scattering analysis of the 3 IMAC-purified fractions identified a single monodispersed species with an expected molecular mass of 85 kDa, and fluorimetric analysis confirmed that the purified protein was not aggregated (data not shown). In comparison, urea-solubilized E8R inclusion bodies diluted in PBS indicated, by this method, nonhomogeneous particles with an average mass of several millions of Daltons. Thus, IMAC purification can effectively select soluble nativelike proteins from other soluble misfolded forms of the E8R protein. In addition, the soluble fractions from urea-dena-

IPTG

30 min	2hs x 37°C	DnaK x	O36	x
37°C		1		1
37°C + NaCl		2		3.5
37°C + NaCl + bet		2.5		5.5
37°C + NaCl + bet + glut		2.5		7
37°C + NaCl + glut		2.5		3
42°C		4		2.5
20°C + benzyl alcohol	20hs x 20°C	3.5		48

Fig 2. Osmolytes or BA treatments but not heat shock improve IMAC yields of O36. Thirty minutes before induction, 0.5 M NaCl, 5 mM betaine, or K-glutamate (or all) were added as indicated. Cells were harvested 2 hours after induction and extracts were centrifuged, IMAC purified, separated on SDS-PAGE, and stained by Coomassie blue as in Figure 1. Total homogenates were separated by SDS-PAGE and immunoblotted with DnaK antibodies. Relative amounts of DnaK and O36 were estimated with a densitometer and the values reported in the table. A heat shock effect was induced by raising the growth temperature to 42°C for 30 minutes before IPTG induction. BA (10 mM) was added to a culture grown at 20°C, 20 minutes before IPTG induction (0.1 mM). Bacteria were recovered after 20 hours with BA at 20°C. Experiments were repeated at least thrice. BA, benzyl alcohol; IMAC, immobilized metal ion affinity chromatography; SDS-PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis.

tured PBS-diluted samples showed very low affinity to the resin, confirming a correlation between protein misfolding and poor IMAC binding. When the different total soluble fractions were treated with trypsin, only that from the salt + betaine treatment resulted in a significant amount of undigested E8R protein (Fig 1). This confirms a correlation between improved IMAC binding and increased protease resistance, which are both characteristic of native proteins.

Osmolyte accumulation and BA treatment increase the IMAC yields of the O36 protein

The presence of high concentrations of osmolytes, such as betaine or K-glutamate, can strongly stabilize heat-labile proteins during heat stress in vivo and favor native refolding of urea-denatured proteins in vitro (Diamant et al 2001, 2003). The results shown in Figure 1 confirm that the cellular accumulation of a viscous osmolyte such as betaine, during massive E8R synthesis, may strongly favor the funneling of early folding intermediates into the native-folding pathway. We next confirmed and generalized this observation using a different folding-recalcitrant protein, O36, testing also another type of osmolyte, K-

glutamate known as a noncompatible osmolyte involved in short-term adaptation of organisms to salt stress. Before protein synthesis was induced, bacteria were grown at 37°C from OD_{600} of 0.3 to 0.6 in the presence of medium supplemented with NaCl to the final concentration of 0.5 M, without or with 5 mM betaine or K-glutamate. This is a sublethal salt concentration inducing the uptake or synthesis of osmolytes in the bacteria, up to cellular concentrations in the molar range (Dinnbier et al 1988).

As in the case of E8R, the presence of salt and osmolytes in the medium increased the amount of IMAC-purified O36 by 3- to 7-fold (Fig 2). Uptake of osmolytes present in the medium is less energy consuming than osmolyte synthesis by the bacteria. In agreement, the presence of osmolytes in the medium produced more IMAC-soluble proteins than with salt stress alone. Moreover, the effect of betaine and K-glutamate in the medium was cumulative. Interestingly, salt stress only mildly increased the cellular concentration of DnaK, the major molecular chaperone involved in protein disaggregation (Diamant et al 2000; Ben-Zvi et al 2004). In comparison, a short heat shock (42°C for 30 minutes) before IPTG induction resulted in a 4-fold increase of the DnaK concentration but only a 2.5-fold increase of the IMAC yields.

This confirms that, independent of the aggregation-prevention and curing activity of this molecular chaperone, chemical chaperones can efficiently favor the native-folding pathway of protein in vivo as well as in vitro (Diamant et al 2001, 2003).

In *E coli* as well as other organisms, BA can artificially induce a strong heat shock response at low temperature, by induced membrane fluidization (Horvath et al 1998; Saidi et al 2005; Shigapova et al 2005). We examined in this study the effect of BA treatment at 10 mM, a concentration inducing acquired thermoresistance and a strong expression of the bacterial chaperone network (Shigapova et al 2005). Remarkably, BA treatment resulted in a 48-fold increase of the IMAC-soluble O36, corresponding to more than 50% of the total O36 expressed in the cell. DnaK levels were increased to the same mild 4-fold level as that after heat shock at 42°C without BA (Fig 2).

We, therefore, next compared the general heat shock–like effect of BA treatment (at 20°C) with plasmid overexpressing specific combinations of molecular chaperones, on the IMAC yields of the O15 protein.

Combined effects of osmolytes, BA induction, and overexpressed GroES-GroEL

Goloubinoff et al (1989b) first demonstrated that plasmid-encoded overexpression of plasmid-encoded molecular chaperones (the groES-groEL operon) can significantly prevent the misfolding and aggregation of a recombinant RubisCO enzyme from cyanobacteria and promote the correct folding and assembly of active RubisCO L8 and L8S8 complexes in *E coli*. In this study, we observed the same positive effect by plasmid-encoded GroES-GroEL overexpression on the IMAC yields of the O15 protein (Fig 3).

Whereas the factor of IMAC solubilization by salt + betaine, or BA treatment alone, was, respectively, 2- and 6-fold, it was 16-fold by plasmid coexpression of the GroES-GroEL operon. Combining plasmid expression of GroEL-GroES with BA application (but not with salt + betaine) was synergistic, with a remarkable 42-fold improvement of the IMAC yields, corresponding to 20% of the total O15 protein in the cell (Fig 3). This suggests that BA, but not salt + betaine, induces other endogenous molecular chaperones, such as IbpA/B, DnaK, and ClpB (Veinger et al 1998; Mogk et al 2003), acting in synergy with the GroES-GroEL chaperonins to prevent aggregation and to mediate active disaggregation and correct refolding of the O15 protein in the cell. We next tested the synergistic effects of various ensembles of plasmid coexpressed molecular chaperones.

Effect of chaperone networks and osmolytes on recombinant protein solubility

The major members of the *E coli* chaperone network, DnaK-DnaJ-GrpE, ClpB, GroES-GroEL, and IbpA-IbpB, efficiently collaborate in the passive prevention of protein aggregation, and significantly, in the active unfolding and disaggregation (leading to native refolding) of stably aggregated proteins in vitro and in vivo (Goloubinoff et al 1989a, 1989b, 1999; Veinger et al 1998; Ben-Zvi and Goloubinoff 2001; Mogk et al 2003; de Marco and De Marco 2004; Schrödel and de Marco 2005). To address the combined roles of overexpressed chaperones and osmolytes, the E8R protein was coexpressed without or with a number of individual compatible plasmids expressing either (1) the full DnaK-DnaJ-GrpE operon together with the ClpB gene (KJEB), (2) the GroES-GroEL operon (LS), or (3) the IbpA-IpbB operon I_{AB}.

As initially shown in Figure 1, without plasmid-expressed chaperones and without salt-induced osmolyte accumulation, IMAC yields of E8R were very low (Fig 4). Because arresting massive protein synthesis at the later stages of induction can improve yields of native recombinant proteins (Carrió and Villaverde 2001), we also tested the effect of salt-induced osmolyte uptake (without chaperone plasmids) under conditions where recombinant protein synthesis was arrested by IPTG removal. Compared with IPTG coapplication (Fig 1), synthesis arrest during the salt + betaine and salt + K-glutamate treatments significantly improved the IMAC yields of E8R, by 14- and 5-fold, respectively (Fig 4). This suggests that during synthesis arrest, the chaperone network becomes less engaged in fruitless interactions with de novo misfolding species and allows more effective scavenging and unfolding of stable E8R aggregates into natively refoldable species.

We next tested during recombinant synthesis arrest, the direct effect of plasmid-expressed chaperone networks on IMAC yields and found a 26- and 32-fold improvement in the presence of KJEB + LS or KJEB + LS + I_{AB} expressing plasmids, respectively, reaching up to 50% of the total E8R protein (Fig 4). No additive effect of salt + osmolyte treatment on chaperone overexpression was observed, possibly because of the masking effect by the very high yields already reached with chaperones alone. Thus, in the case of the E8R protein, the tedious procedure of plasmid-mediated chaperone coexpression can be substituted by a simpler procedure of salt + osmolytes addition, although producing slightly lower yields of native protein.

Improved yields by chaperone coexpression or by BA treatment

To address the generality of the process, 6 folding-recalcitrant recombinant proteins were tested in the presence
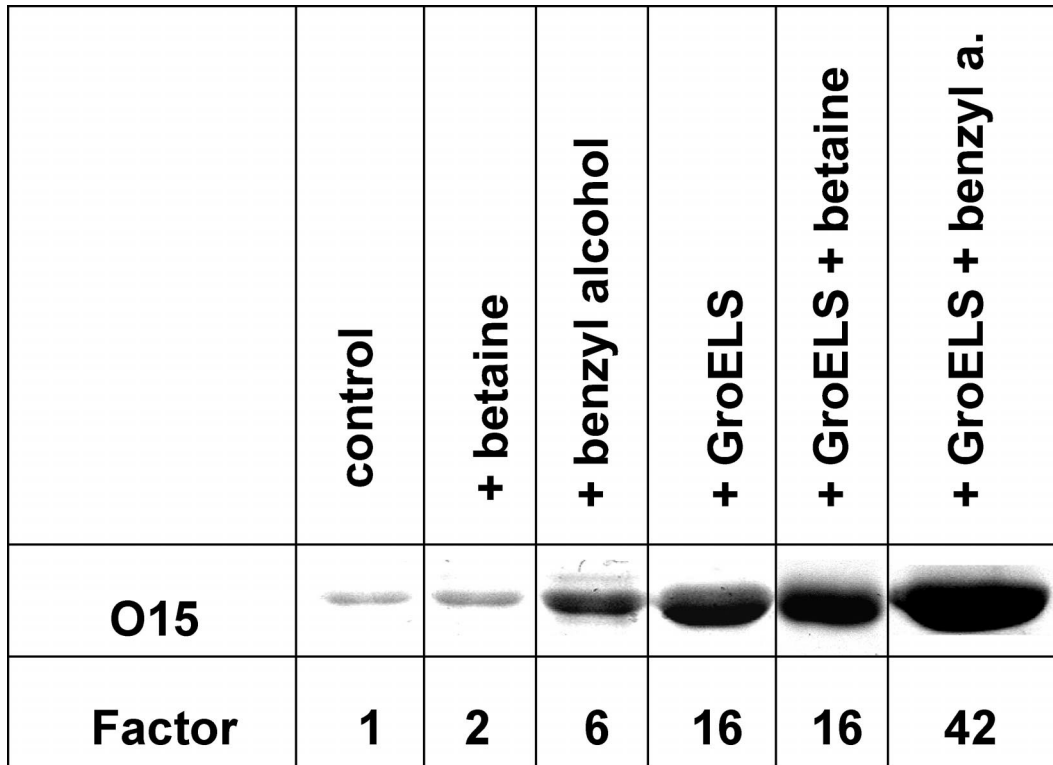
	control	+ betaine	+ benzyl alcohol	+ GroELS	+ GroELS + betaine	+ GroELS + benzyl a.
O15						
Factor	1	2	6	16	16	42

Fig 3. Complementary effects of GroES-GroEL, betaine, and BA on the solubility of the O15 protein. Cells expressing recombinant proteins O15, without or in the presence of a second compatible plasmid coexpressing the GroES-GroEL operon, were treated at 20°C for 20 hours in the presence of salt + betaine or of BA, as in Figure 2. Betaine and BA were added to the bacterial culture 30 and 20 minutes before IPTG induction, respectively. Soluble O15 protein was IMAC purified and detected as in Figures 1 and 2. Values are from 1 representative experiment out of 3. BA, benzyl alcohol; IMAC, immobilized metal ion affinity chromatography.
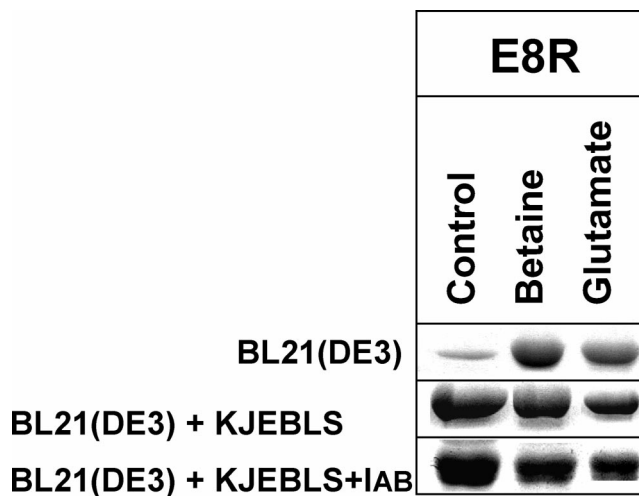

Fig 4. IMAC yields in the presence of different chaperones and osmolytes. E8R was expressed as in Figure 1 in the absence or presence of compatible plasmids coexpressing DnaK-DnaJ-GrpE and ClpB (KJEB), GroES-GroEL (LS), and IbpA/B (IAB). NaCl and osmolytes were added 30 minutes before IPTG induction (0.1 mM) and the cells were grown 20 hours at 20°C. Then, IPTG was removed by centrifugation from the culture medium and the cells were grown for 2 additional hours at 20°C to allow chaperone-mediated disaggregation. The E8R protein was purified by IMAC and detected as in Figures 1–3. Values are from 1 representative experiment out of 3. IMAC, immobilized metal ion affinity chromatography.

of plasmid-overexpressed chaperones (Fig 4). Using optimal chaperone combinations (as detailed in Fig 4), this approach was highly efficient with O46, MBP, E8R, and N61 but not with O82 and N34 (Fig 5). Because BA treatment induces in *E coli* the expression of endogenous chaperones (Shigapova et al 2005; Fig 2), the effect of simple BA application was compared with that of plasmid chaperone coexpression. Again, BA application was highly efficient with the same types of proteins as with chaperone coexpression (O46, MBP, E8R, and N61 but not with O82 and N34), suggesting that chaperones induced by BA or expressed from plasmids act on the same substrates. Moreover, IMAC yields of O46 were several folds higher with BA induction than with chaperone coexpression (Fig 5). This suggests that by causing a heat shock–like response at nonstress temperature, BA induces the network of endogenous chaperones that can optimally prevent aggregation and actively convert stable recombinant protein aggregates into native proteins. Regarding biotechnological applications, for most of the cases tested, a simple treatment with the membrane fluidizer, BA, causing synthesis of all the endogenous chaperones in optimal ratios, was as effective as the more intricate approach of plasmid coexpression of specific chaperone combinations.

Proteins	046	MBP	E8R	N61	082	N34
Control						
+ molecular chaperones	I	I	II	II		
+ benzyl alcohol						

Fig 5. Improved IMAC yields using molecular chaperones and BA. In a preliminary screening to identify optimal conditions for expressing soluble target proteins, cells harboring each of the 6 target proteins, were cotransformed with different combinations of chaperone plasmids (I: KJEB + LS, II: KJEB + LS + IAB), as in Figure 4. BA treatment, protein IMAC purification and detection was as in Figures 1–4. Values are from 1 representative experiment out of 3. BA, benzyl alcohol; IMAC, immobilized metal ion affinity chromatography.

It should be noted that during upscaling attempts from 3 to 500 mL bacterial cultures, no significant differences in IMAC yields were observed after salt + osmolyte treatments or plasmid chaperone coexpression. In contrast, BA treatment was often found to inhibit cell growth when the same conditions set for small-scale cultures were used for large flasks (data not shown). Therefore, an additional case-by-case analysis of the possible factors, such as changes of the BA concentration and tuning of the gas exchange by varying rotation speed and flask profile, was necessary to optimize BA treatment to large-scale bacterial cultures.

DISCUSSION

The massive overexpression and accumulation of soluble recombinant proteins in *E coli* is often hampered by off-pathway misfolding and aggregation events, leading to the formation of nonnative, nonfunctional compact as well as soluble protein aggregates (Schrödel and de Marco 2005). Whereas different protocols must be tested to improve native folding of recombinant proteins in the bacteria, it is not less necessary to use better criteria to discriminate between natively folded and stably misfolded recombinant proteins, especially when enzymatic or fluorescent assays are lacking.

The data in Figure 1 highlight the weakness of the solubility assay after ultracentrifugation, as a means to estimate the partitioning between misfolded and natively folded recombinant proteins. To better identify conditions favoring native folding, we describe in this study 2 simple steps that can be applied subsequent to centrifugation, one based on the specific high affinity of natively folded proteins for metal-activated resins and the other based on the higher aptitude of natively folded proteins to withstand a mild trypsin treatment, as compared with misfolded proteins.

Several strategies have been initially used to improve yields of natively folded recombinant proteins. Slower cell growth, lower amounts of inducers, and incubation at lower temperatures are simple methods to provide cells with better conditions to foster correct native folding of the recombinant proteins. In a preliminary screening, we compared and selected fusions between the target proteins and the carriers that, in each case, enabled the highest yields of soluble protein. Recently, the mere arrest of recombinant protein synthesis (by inducer removal) during the later phase of expression was shown to favor the native refolding of previously misfolded recombinant proteins, likely by allowing the recruitment of the endogenous chaperones to the active scavenging of already formed aggregates (Carrió and Villaverde 2001; de Marco and De Marco 2004).

The major members of the *E coli* chaperone machinery, DnaK-DnaJ-GrpE, ClpB, GroES-GroEL, and IbpA/IbpB, are also heat-inducible proteins (Hsps), whose accumulation can be induced by a simple heat treatment. However, we found that heat shock had no significant positive solubilization effect, probably because heat shock temperatures often have a stronger negative effect on the stability and the proper folding pathway of aggregation-prone proteins (Gur et al 2002). Moreover, although chaperone concentrations may rapidly decrease after heat shock, ongoing massive synthesis of recombinant proteins may yet necessitate high chaperone concentrations to foster optimal native folding. Thus, to gain control over chaperone expression during recombinant protein expression, a more laborious method of recombinant chaperone coexpression may be advantageous (Goloubinoff et al 1989b; de Marco et al 2004).

In this study, we propose 2 additional less-demanding procedures to improve the yields of several different natively folded recombinant proteins. In the specific case of the folding-recalcitrant protein E8R (Fig 1) and O36 and O15 (see Figs 2, 3), salt and osmolyte addition to the growth medium, causing accumulation of high cellular concentrations of chemical chaperones during protein synthesis, was shown to significantly improve the yields of the natively folded proteins. Because of the unchanged

DnaK levels, the mechanism of direct-assisted de novo folding may be independent from assisted refolding by way of active ClpB- and DnaK-mediated disaggregation (Goloubinoff et al 1999; Ben-Zvi et al 2004).

Plasmid-mediated overexpression of various ensembles of the chaperones was highly effective at improving IMAC yields in 4 (MBP, E8R, N61, and O46) of the 6 cases of folding-recalcitrant recombinant proteins. The membrane sensor hypothesis predicts that fluidity elevation of specific membrane domains by temperature or by chemical fluidizers may act as a cellular thermometer, generating a primary stress signal that culminates in the activation of heat shock genes, even in the absence of significant protein unfolding in the cell (Vigh et al 1998; Shigapova et al 2005). In accordance, BA addition resulted in a similar specificity of action, with significant IMAC purification of the MBP, E8R, N61, and O46 proteins but not of O82 and N34. This suggests that BA acts by inducing the expression of the endogenous chaperone network. Shigapova et al (2005) have shown in wild-type *E coli* cells not stressed by expressed recombinant proteins, a strong BA-induced accumulation of chaperone messenger RNA, which was not readily followed by increased chaperone protein levels (DnaK, GroEL) within 20 minutes of BA treatment. In this study, we observed in *E coli* cells stressed by massive expression of recombinant proteins that BA treatment resulted in a 3.5-fold increase of DnaK protein within 20 hours after BA addition. This suggests that after heat shock or after BA induction, sufficient time has to be given for the chaperones to accumulate in the cell, to prevent aggregation and actively solubilize aggregates during massive synthesis of recombinant proteins.

Relative advantages of molecular chaperones vs osmolytes

The understanding of the molecular mechanisms by which osmolytes and specific molecular chaperones act in stressed and nonstressed bacterial cells and in vitro is central to the design of protocols to produce optimal amounts of natively folded recombinant proteins.

The presence in the cell of physiological amounts of compatible osmolytes, such as proline, glycine-betaine, and trehalose, or of less-compatible osmolytes, such as K-glutamate, can significantly increase the stability of native thermolabile proteins. Moreover, physiological (molar) amounts of osmolytes can efficiently promote the in vitro spontaneous native refolding of urea-denatured proteins (Diamant et al 2001, 2003). This may be a direct consequence of osmolyte viscosity, which by slowing down rates of structure acquisition would reduce the thermodynamically less favorable, accidental misfolding events to a greater extent than the thermodynamically more fa-

vorable native-folding events (Anfinsen 1972). Thus, improving the yields of natively folded recombinant proteins with high osmolyte concentrations likely results from a direct effect of viscosity on the partitioning between misfolding and native-folding events. The same mechanism of direct-assisted de novo native folding is probably also at work when the viscosity of the cytoplasm is increased at lower temperatures.

Similar to the effect of osmolytes, the overexpression of small heat shock proteins (IbpA/B in *E coli*) was also shown to provide some protection to native proteins against thermal denaturation in various organisms and in vitro (Forreiter et al 1997; Veinger et al 1998; Haslbeck 2002). Significantly, small heat shock proteins are also powerful binders of misfolding proteins, which may thus reduce protein aggregation and optimally present misfolded proteins to scavenging molecular chaperones (Ehrnsperger et al 1997; Veinger et al 1998; Török et al 2001; Mogk et al 2003). The ATPase chaperones DnaK, GroEL, and ClpB may indirectly favor the native-folding pathway by preventing protein aggregation and, more significantly, by converting the energy of ATP hydrolysis into an unfolding energy that can extract stable misfolded intermediates from kinetic traps and relocate them back on the spontaneous refolding pathway leading to the native state (Shtilerman et al 1999; Ben-Zvi et al 2004; Weibezahn et al 2004; Schrödel and de Marco 2005).

In conclusion, to optimize the native folding of recombinant proteins, simple protocols can combine salt-induced uptake of specific osmolytes, to favor de novo native folding, and BA- or plasmid-mediated overexpression of chaperones, to favor native refolding of insoluble as well as soluble protein aggregates. The following simple and increasingly complex actions can be suggested to improve the solubility and native folding of overexpressed recombinant proteins in *E coli*. (1) The growth temperature and inducer concentration can be reduced (Steczko et al 1991). (2) Chaperone accumulation can be induced by addition of BA (this study). (3) Chaperone accumulation can be induced by heat (Makrides 1996). (4) Osmolyte accumulation can be induced by salt stress and osmolyte addition (Dinnbier et al 1998; Diamant et al 2003; this study). (5) A combination of salt-induced osmolyte accumulation and BA-induced chaperone accumulation can be applied (this study). (6) Coexpression of single plasmid–encoded single molecular chaperones (Goloubinoff et al 1989b). (7) Coexpression of multiple plasmids encoding several molecular chaperone systems (Mogk et al 2003; de Marco and De Marco 2004). (8) Engineering translational fusions with soluble protein carriers (LaVallie et al 1993; Pryor and Leiting 1997; De Marco et al 2004).

ACKNOWLEDGMENTS

REFERENCES

Anfinsen CB. 1972. The formation and stabilization of protein structure. *Biochem J* 128: 737–749.

Baneyx F, Mujacic M. 2004. Recombinant protein folding and misfolding in *Escherichia coli*. *Nat Biotechnol* 22: 1399–1408.

Ben-Zvi AP, De Los Rios P, Dietler G, Goloubinoff P. 2004. Active solubilization and refolding of stable protein aggregates by cooperative unfolding action of individual hsp70 chaperones. *J Biol Chem* 279: 37298–37303.

Ben-Zvi AP, Goloubinoff P. 2001. Review: mechanisms of disaggregation and refolding of stable protein aggregates by molecular chaperones. *J Struct Biol* 135: 84–93.

Ben-Zvi AP, Goloubinoff P. 2002. Proteinaceous infectious behaviour in non-pathogenic proteins is controlled by molecular chaperones. *J Biol Chem* 277: 49422–49427.

Carrió M, Gonzalez-Montalban N, Vera A, Villaverde A, Ventura S. 2005. Amyloid-like properties of bacterial inclusion bodies. *J Mol Biol* 347: 1025–1037.

Carrió M, Villaverde A. 2001. Protein aggregation as bacterial inclusion bodies is reversible. *FEBS Lett* 489: 29–33.

de Marco A, De Marco V. 2004. Bacteria co-transformed with recombinant proteins and chaperones cloned in independent plasmids are suitable for expression tuning. *J Biotechnol* 109: 45–52.

de Marco A, Guzzardi P, Jamet E. 1999. Isolation of tobacco isoperoxidases accumulated in cell-suspension culture medium and characterization of activities related to cell wall metabolism. *Plant Physiol* 120: 371–381.

De Marco V, Stier G, Blandin S, de Marco A. 2004. The solubility and stability of recombinant proteins is increased by their fusion to NusA. *Biochem Biophys Res Commun* 322: 766–771.

Diamant S, Ben-Zvi AP, Bukau B, Goloubinoff P. 2000. Size-dependent disaggregation of stable protein aggregates by the DnaK chaperone machinery. *J Biol Chem* 275: 21107–21113.

Diamant S, Eliahu N, Rosenthal D, Goloubinoff P. 2001. Chemical chaperones regulate molecular chaperones in vitro and in cells under combined salt and heat stresses. *J Biol Chem* 276: 39586–39591.

Diamant S, Rosenthal D, Azem A, Eliahu N, Ben-Zvi AP, Goloubinoff P. 2003. Dicarboxylic amino acids and glycine-betaine regulate chaperone-mediated protein disaggregation under stress. *Mol Microbiol* 49: 401–410.

Dinnbier U, Limpinsel E, Schmid R, Bakker EP. 1988. Transient accumulation of potassium glutamate and its replacement by trehalose during adaptation of growing cells of *Escherichia coli* K-12 to elevated sodium chloride concentrations. *Arch Microbiol* 150: 348–357.

Ehrnsperger M, Graber S, Gaestel M, Buchner J. 1997. Binding of non-native protein to Hsp25 during heat shock creates a reservoir of folding intermediates for reactivation. *EMBO J* 16: 221–229.

Forreiter C, Kirschner M, Nover L. 1997. Stable transformation of an Arabidopsis cell suspension culture with firefly luciferase providing a cellular system for analysis of chaperone activity in vivo. *Plant Cell* 9: 2171–2181.

Georgiou G, Valax P. 1996. Expression of correctly folded proteins in *Escherichia coli*. *Curr Opin Biotechnol* 7: 190–197.

Goloubinoff P, Christeller JT, Gatenby AA, Lorimer GH. 1989a. Reconstitution of active ribulose bisphosphate carboxylase from an unfolded state depends on two chaperonin proteins and Mg-ATP. *Nature* 342: 884–889.

Goloubinoff P, Gatenby AA, Lorimer GH. 1989b. GroE heat-shock proteins promote assembly of foreign prokaryotic ribulose bisphosphate carboxylase oligomers in *Escherichia coli*. *Nature* 337: 44–47.

Goloubinoff P, Mogk A, Zvi AP, Tomoyasu T, Bukau B. 1999. Sequential mechanism of solubilization and refolding of stable protein aggregates by a bichaperone network. *Proc Natl Acad Sci U S A* 96: 13732–13737.

Gur E, Biran D, Gazit E, Ron EZ. 2002. *In vivo* aggregation of a single enzyme limits growth of *Escherichia coli* at elevated temperatures. *Mol Microbiol* 46: 1391–1397.

Hanning G, Makrides SC. 1998. Strategies for optimizing heterologous protein expression in *Escherichia coli*. *Trends Biotechnol* 16: 54–60.

Haslbeck M. 2002. sHsps and their role in the chaperone network. *Cell Mol Life Sci* 59: 1649–1657.

Horvath I, Glatz A, Varvasovszki V, et al. 1998. Membrane physical state controls the signalling mechanism of the heat shock response in *Synechocystis PCC 6803*: identification of hsp17 as a "fluidity gene". *Proc Natl Acad Sci U S A* 95: 3513–3518.

Jurado P, Ritz D, Beckwith J, de Lorenzo V, Fernandez LA. 2002. Production of functional single-chain Fv antibodies in the cytoplasm of *Escherichia coli*. *J Mol Biol* 320: 1–10.

LaVallie ER, DiBlasio EA, Kovacic S, Grant KL, Schendel PF, McCoy JM. 1993. A thioredoxin gene fusion expression system that circumvents inclusion body formation in the *E. coli* cytoplasm. *BioTechnology* 11: 187–193.

Makrides SC. 1996. Strategies for achieving high-level expression of genes in *Escherichia coli*. *Microbiol Rev* 60: 516–538.

Middelberg APJ. 2002. Preparative protein refolding. *Trends Biotechnol* 20: 437–443.

Mogk A, Deuerling E, Vorderwulbecke S, Vierling E, Bukau B. 2003. Small heat shock proteins, ClpB and the DnaK system form a functional triade in reversing protein aggregation. *Mol Microbiol* 50: 585–595.

Nominé Y, Ristriani T, Laurent C, Lefevre J-F, Weiss E, Travé G. 2001a. Formation of soluble inclusion bodies by HPV E6 oncoprotein fused to maltose-binding protein. *Protein Expr Purif* 23: 22–32.

Nominé Y, Ristriani T, Laurent C, Lefevre J-F, Weiss E, Travé G. 2001b. A strategy for optimizing the monodispersity of fusion proteins: application to purification of recombinant HPV E6 oncoprotein. *Protein Eng* 4: 297–305.

Pryor KD, Leiting B. 1997. High-level expression of soluble protein in *Escherichia coli* using His6-tag and maltose binding-protein double-affinity fusion system. *Protein Expr Purif* 10: 309–319.

Rogl H, Kosemund K, Kuhlbrandt W, Collinson I. 1998. Refolding of *Escherichia coli* produced membrane protein inclusion bodies immobilised by nickel chelating chromatography. *FEBS Lett* 432: 21–26.

Sachdev D, Chirgwin JM. 1999. Properties of soluble fusions between mammalian aspartic proteases and bacterial maltose-binding protein. *J Protein Chem* 18: 127–136.

Saidi Y, Schaefer D, Finka A, Zryd JP, Goloubinoff P. 2005. Controlled expression of recombinant proteins in *Physcomitrella Patens* by a

conditional heat-shock promoter: a tool for plant research and biotechnology. *Plant Mol Biology,* in press.

Schrödel A, de Marco A. 2005. Characterization of the aggregates formed during recombinant protein expression in bacteria. *BMC Biochem* 6: 10.

Schrödel A, Volz J, de Marco A. 2005. Fusion tags and chaperone co-expression modulate both the solubility and the inclusion body features of the recombinant clipB14 serine protease. *J Biotechnol,* in press.

Shigapova N, Török Z, Balogh G, Goloubinoff P, Vigh L, Horvath I. 2005. Membrane fluidization triggers membrane remodelling which affects the thermotolerance in *Escherichia coli. Biochem Biophys Res Commun* 328: 1216–1223.

Shtilerman M, Lorimer GH, Englander SW. 1999. Chaperonin function: folding by forced unfolding. *Science* 284: 822–825.

Speed MA, Wang DIC, King J. 1996. Specific aggregation of partially folded polypeptide chains: the molecular basis of inclusion body composition. *Nat Biotechnol* 14: 1283–1287.

Steczko J, Donoho GA, Dixon JE, Sugimoto T, Axelrod B. 1991. Effect of ethanol and low temperature culture on expression of soybean lipoxygenase L-I in *Escherichia coli. Protein Expr Purif* 2: 221–227.

Tang S. 2001. A modular polycistronic expression system for over-expressing protein complexes in *Escherichia coli. Protein Exp Purif* 21: 224–234.

Tomoyasu T, Mogk A, Langen H, Goloubinoff P, Bukau B. 2001. Genetic dissection of the roles of chaperones and proteases in protein folding and degradation in the *Escherichia coli* cytosol. *Mol Microbiol* 40: 397–413.

Török Z, Goloubinoff P, Horvath I, et al. 2001. Synechocystis HSP17 is an amphitropic protein that stabilizes heat-stressed membranes and binds denatured proteins for subsequent chaperone-mediated refolding. *Proc Natl Acad Sci U S A* 98: 3098–3103.

Veinger L, Diamant S, Buchner J, Goloubinoff P. 1998. The small heat-shock protein IbpB from *Escherichia coli* stabilizes stress-denatured proteins for subsequent refolding by a multichaperone network. *J Biol Chem* 273: 11032–11037.

Vigh L, Maresca B, Harwood J. 1998. Does the membrane physical state control the expression of heat shock and other genes? *Trends Biochem Sci* 23: 369–373.

Vincentelli R, Bignon C, Gruez A, Canaan S, Sulzenbacher G, Tegoni M, Campanacci V, Cambillau C. 2003. Medium-scale structural genomics: strategies for protein expression and crystallization. *Acc Chem Res* 36: 165–172.

Weibezahn J, Tessarz P, Schlieker C, et al. 2004. Thermotolerance requires refolding of aggregated proteins by substrate translocation through the central pore of ClpB. *Cell* 119: 653–665.

Zhang Y, Olsen DR, Nguyen KB, Olson PS, Rhodes ET, Mascarenhas D. 1998. Expression of eukaryotic proteins in soluble form in *Escherichia coli. Protein Expr Purif* 12: 159–165.