Effects of overexpressing folding modulators on the *in vivo* folding of heterologous proteins in *Escherichia coli*

J Gerard Wall and Andreas Plückthun

University of Zürich, Zürich, Switzerland

Interest continues to increase in the use of folding modulators to overcome problems with heterologous protein folding in *Escherichia coli*. Currently, this approach, though highly successful with a number of individual proteins, remains a somewhat hit-and-miss affair. Ongoing research directed at unraveling the precise role and specificity of these folding modulators should generate a clearer understanding of the potential and limitations of overexpressing folding catalysts *in vivo*. This will facilitate the development, in the not too distant future, of a more structured and rational approach to improving the folding of heterologous gene products in *E. coli*.

Current Opinion in Biotechnology 1995, **6**:507–516

Introduction

It is now generally accepted that one of the limiting factors in the production of soluble functional heterologous proteins in *Escherichia coli* is the aggregation tendency during folding. The fundamental issue for many researchers, therefore, is why the folding of many recombinant proteins is only poorly guided to the native state by the folding modulators present in wild-type *E. coli*, whereas the folding of *E. coli*'s own proteins apparently does not constitute a problem.

Overexpression of a protein in *E. coli* entails the creation of an unnatural situation where folding intermediates are present at very high concentrations, which results in a greater tendency to aggregate. Many overexpressed proteins exist in their natural environments only in low amounts, and thus little evolutionary pressure may have existed to optimize their sequences for efficient folding. Recombinant proteins are often expressed at levels that are orders of magnitude higher than their normal expression levels. It is immediately obvious that the propensity to aggregate and the precise mechanistic reasons for aggregation will be strongly dependent on the particular recombinant protein being expressed, and that universally useful solutions are thus unlikely to exist.

In the light of recent research into the functioning of *E. coli*'s folding modulators, this review focuses on effecting improved folding of recombinant proteins. Particular emphasis is placed on the overexpression of these modulators in manipulating the *in vivo* folding of heterologous gene products expressed in the bacterium.

The panoply of factors that affect protein folding

Before discussing any subtleties of protein folding, such as possible constraints on the protein sequence, potential specificities of folding modulators, or the regulation of the chaperone machinery, a few global factors must be clarified. Successful protein folding requires, of course, that the end product is a thermodynamically stable entity. Many reports of only 'insoluble material' being produced upon expression of a particular protein in *E. coli* are not in the least surprising, in that something thermodynamically impossible was attempted. Almost invariably, truncated domains are severely destabilized and often totally unable to fold to monomeric protein; for example, a β-barrel from which, say, two strands are missing (e.g. because the domain was 'defined' as being located between conveniently spaced restriction sites) will generally not choose to go to the native state of the original protein. Similarly, dimeric complexes or multi-subunit assemblies may not tolerate the absence of a subunit, by virtue of a large hydrophobic subunit interface becoming exposed. In such cases, aggregation may be the only option for the protein to cover its hydrophobic surface. Furthermore, many heterologous proteins of interest are naturally secreted and contain disulfide bonds. The formation of these disulfides is often crucial for structure formation (i.e. for the stability of an intermediate) or at least for obtaining a minimal stability in the final product. In cases such as these, expression of functional protein in the cytoplasm of *E. coli* has very little chance of success, unless measures are taken

Abbreviations

BPTI—bovine pancreatic trypsin inhibitor; **Hsp**—heat-shock protein; **IL**—interleukin; **PPIase**—peptidyl prolyl *cis/trans* isomerase; **Rubisco**—ribulose-6-phosphate carboxylase/oxygenase; **TCR**—T-cell receptor.

that favor disulfide bond formation in this compartment, even though soluble, albeit inactive, protein is sometimes formed [1,2].

Moreover, post-translational modifications, notably glycosylation, may contribute to the solubility of crucial aggregation-prone intermediates in the folding pathway of some proteins, or even to the stability of the final product, as recently demonstrated in the case of human CD2 [3]. This severely restricts the ability to produce such products in a soluble functional form in *E. coli*, unless the protein can be engineered to exist without the sugars, as was possible with CD2 [3]. The importance of sugars for the stability and folding of proteins varies greatly, ranging from none to essential. Furthermore, it has been proposed that the presence of calnexin in the eukaryotic endoplasmic reticulum membrane (reviewed in [4]) may provide an 'anchoring' of polypeptides via sugar residues during folding, thus facilitating the subsequent transient binding of folding modulators and preventing premature aggregation. This of course is also something that prokaryotic systems cannot imitate. Thus, we may conclude that overexpression of folding modulators in *E. coli* only has a chance of having a beneficial effect in those cases where the final recombinant product is thermodynamically stable as a folded protein and does not depend on post-translational modification for either folding or stability.

In order to achieve correct folding of its own proteins, *E. coli* contains a number of helper proteins (collectively referred to as 'folding modulators' here) that catalyze certain folding steps or prevent the competing reaction of aggregation. It is likely that not all of these factors have been discovered, or at least recognized in their function in protein folding, up to now. One may separate these into 'molecular chaperones', which do not catalyze folding, but block the pathways of aggregation, and factors that facilitate a specific stage in the folding process, that is, 'folding catalysts' (see Fig. 1). Additionally, proteins exist that may prevent premature folding in order to facilitate the export of a polypeptide. Clearly, solving the paradox of the relative inefficiency of folding of many heterologous proteins compared with most 'native' *E. coli* proteins necessitates an understanding of how a protein can itself be optimized to fold well in the bacterial environment as well as the amounts and the specificity of the folding modulators it requires.

General advances in *E. coli* expression technology (e.g. vector improvement and refinement of environmental conditions for cell growth; reviewed in [5•]) have meant that, for many heterologous gene products, the yield-limiting step in their production in *E. coli* is the folding *in vivo*, rather than synthesis, of the polypeptide. This fact, coupled with the demonstration that the

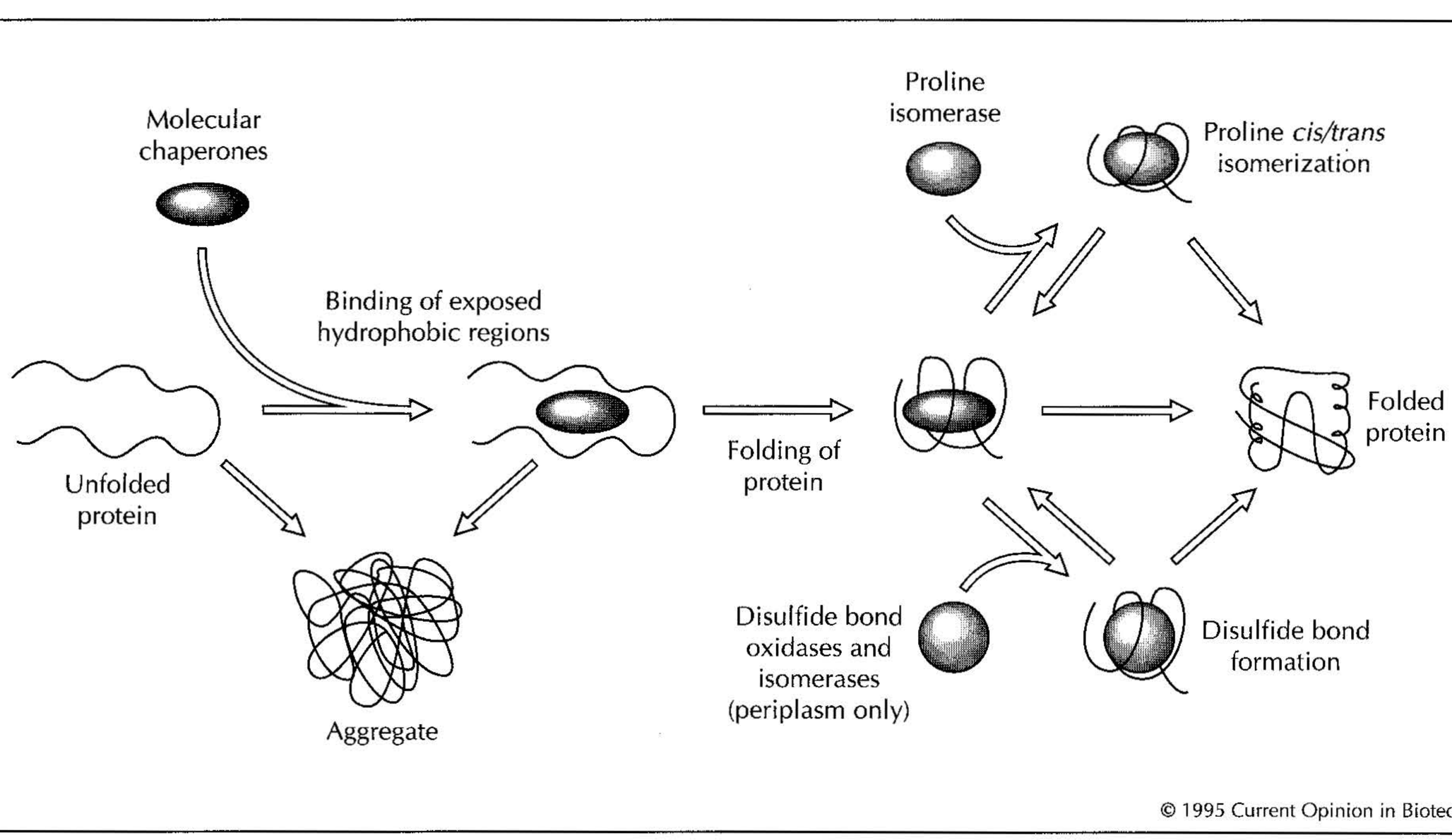

Fig. 1. Simplified overview of protein folding in *E. coli*. The demonstration that a protein contains all the information it needs to assume its correctly folded functional form [71] suggests that folding is essentially spontaneous *in vivo* as well, but is 'channeled' (i.e. side reactions are blocked). Furthermore, specific reactions may be catalyzed and simply occur more quickly. A variety of protein factors have a role in protein folding in *E. coli*, including molecular chaperones, which prevent aggregation of unfolded polypeptide chains resulting from exposed hydrophobic residues before folding or membrane translocation, and folding catalysts, which accelerate such specific steps in the folding process as proline *cis/trans* isomerization and disulfide bond formation [72]. A schematic view of this assisted folding of a protein in *E. coli* is represented in the diagram. A series of more complex interactions is also known to occur between the various components of the folding machinery (see Figs 2, 3) [72,73]. Note that disulfide bond formation only occurs in the periplasmic compartment and requires passage through the membrane, which is not shown here for the sake of simplicity (see Fig. 3).

ability to produce soluble functional proteins in *E. coli* is closely linked to inherent manipulatable properties of the proteins themselves [6–8,9•,10,11,12•], has led to renewed interest in the use of folding catalyst overexpression *in vivo* to improve the production of functional heterologous proteins in *E. coli*.

We should stress again that the principal physicochemical problems of folding are the same for *E. coli* proteins and heterologous proteins. *E. coli* may reasonably be assumed to have optimized the amount of its folding modulator content, and one should not be surprised if the simple overexpression of a single folding factor has frequently no effect, or even a deleterious effect. The effects of the various factors may also be synergistic and for this reason too, folding enhancements may not be apparent if only single components of the folding machinery are overexpressed. The greatest chance of success, therefore, may reside in the use of a 'shotgun' approach, in which a battery of folding modulators would be simultaneously overproduced in order to overcome different folding limitations in a variety of heterologous proteins. Furthermore, it may also be worth noting that the absence from this article of investigations of a particular folding catalyst(s) does not indicate that these studies have not been carried out, merely that no results for such experiments have been documented (i.e. the factor may have been tested but had no effect).

Manipulating the cytoplasmic folding process in *Escherichia coli*

Two major groups of molecular chaperones are now known to be involved in protein folding in the cytoplasm of *E. coli*: the Hsp70 and Hsp60 families. The functional mechanisms of these chaperones are represented in Figure 2. It is very likely, however, that other proteins will soon join the list of unambiguously identified folding modulators of *E. coli*.

The initial demonstration that GroEL and GroES are involved in cellular protein folding resulted from studying the effects of increasing and lowering their expression in the cell on intracellular folding of ribulose-6-phosphate carboxylase/oxygenase (Rubisco) [13]. Several other experiments designed to assess the effect of GroEL/ES overexpression on heterologous protein production in *E. coli* have also been reviewed [5•]. A number of recent studies have reinvestigated this approach with respect to the folding of various other proteins. The co-overexpression of GroES and GroEL results in increased solubility of several recombinant tyrosine kinases [14•,15•], and similar effects of GroEL/ES overproduction on solubility of the *E. coli* glutamate racemase protein have also been noted [16•]. The co-overproduction of GroEL/ES also leads to increased solubility and recovered activity of a heterologous dihydrofolate reductase in *E. coli* [17•]. Interestingly, this chaperone-mediated improvement in

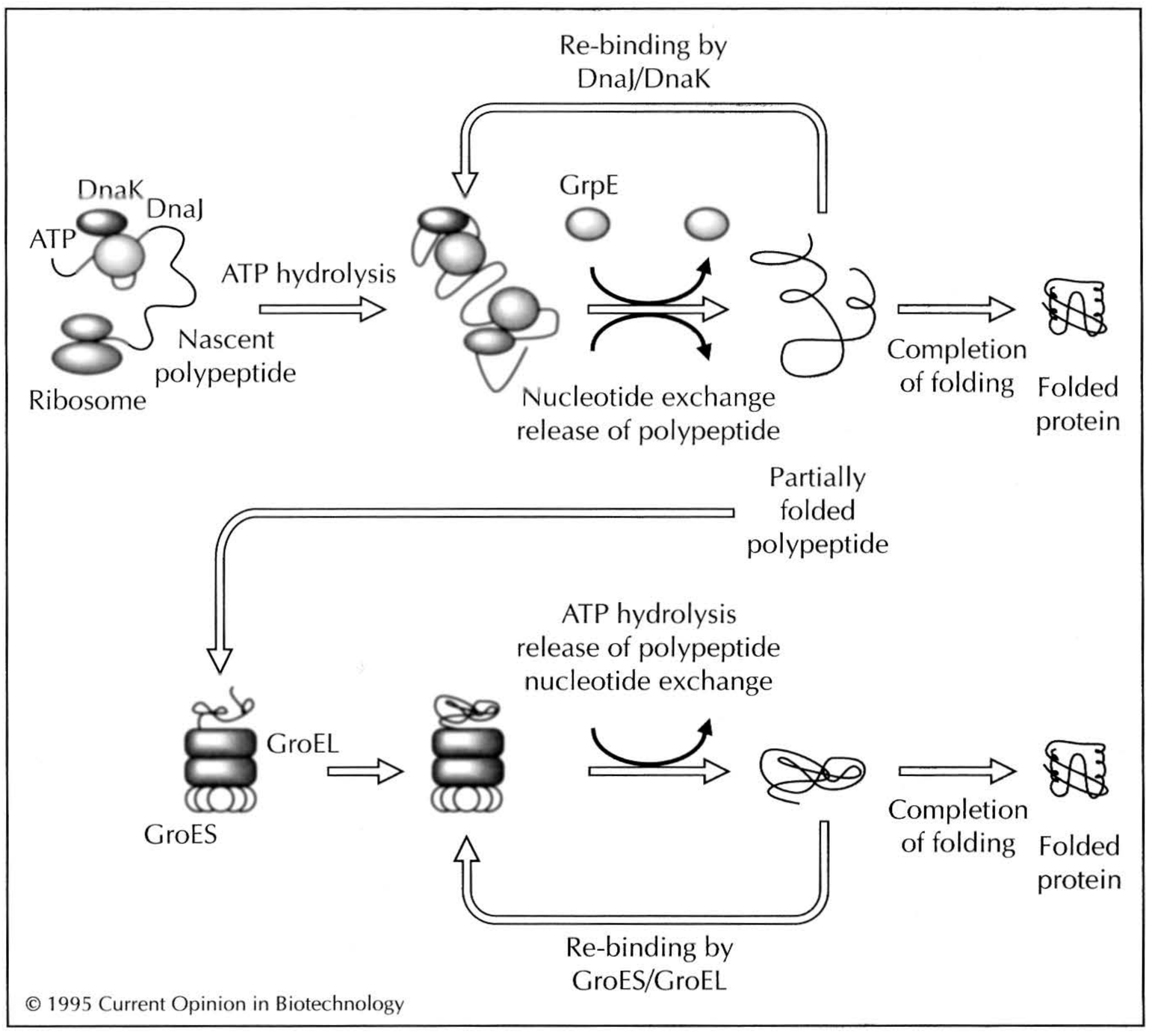

Fig. 2. Role of the Hsp70 machinery (DnaK, DnaJ and GrpE) and Hsp60 machinery (GroEL and GroES) [72,73]. DnaK and DnaJ recognize and bind hydrophobic regions in short peptides, thus preventing premature folding or aggregation during translation and translocation. The ATP-dependent binding and release of peptides uses GrpE for stimulating nucleotide exchange. Partially folded polypeptides may then be passed from the DnaK/DnaJ/GrpE machinery to the Hsp60 complex of *E. coli*, GroEL/ES, a large oligomeric complex that allows completion of folding of the polypeptides in the complex [74]. This simplified model ignores issues of stoichiometry of ATP and chaperones with respect to each other. With the possible emergence of new factors in protein folding, this diagram may be extended to encompass components outside the Hsp70/Hsp60 families. Despite the participation of ATP in various conformational transitions of the chaperones, it should be stressed again that protein folding *per se* occurs spontaneously and does not require a source of energy.

folding ultimately leads to a smaller increase in yield than engineering the protein for better folding [9•,17•] (for another such comparison, see [12•,18]).

Not all proteins require GroEL/ES for folding *in vivo*, however. This is demonstrated by the unchanged production of succinyl-CoA synthetase, an $\alpha_2\beta_2$ tetramer, at the restrictive temperature, in *E. coli* strains bearing GroES or GroEL temperature-sensitive mutations [19].

Other researchers have found no effect of GroEL/ES overproduction on secreted human procollagenase accumulation in *E. coli*. This would be expected because folding of this protein occurs in the periplasm, although a participation of GroEL/ES in the transport process itself cannot be excluded. In the absence of a signal peptide or the presence of a non-functional one, the accumulation of soluble and insoluble protein in the cytoplasm is increased in the GroEL/ES overproducers by increasing the half-life of the protein [20]. It is possible that a compact, more protease-resistant form of the protein can be reached, even though its disulfide bond cannot be formed. Similarly, in studying the effect of overproduction of GroES and GroEL on the expression of single-chain Fv antibody fragments [21], which cannot assume a native structure in the wild-type *E. coli* cytoplasm because of the absence of the disulfide bond (but see [1,2]), intracellular production of soluble, but inactive, fragment was found to be increased twofold. Refolding was subsequently required to obtain any antigen-binding activity, however, which re-emphasises the important point that soluble protein is not necessarily functional protein. No effect on the levels of secreted single-chain Fv fragments was observable, similar to the procollagenase experiments [20], because the production of antibody fragments appears to be limited by periplasmic folding [18].

The above experiments were based on the assumption that GroEL/ES is present in insufficient amounts. Another interesting issue is whether different homologs of GroEL are specific for particular substrates. If this were the case, particular homologs would have to be used instead of GroEL. Neither GroEL nor the plant chaperonin was found to be able to assemble active plant Rubisco in *E. coli*, however, whereas both could enhance the assembly of cyanobacterial Rubisco [22].

Many, though not all [19], recombinant proteins make use of GroEL/ES in the cell and need it to be present at wild-type levels. Such evidence for the general importance of GroEL and GroES in heterologous protein folding in *E. coli* comes from experiments in which their levels were reduced, rather than increased, and an absolute requirement for the two chaperones was demonstrated [13]. In more recent experiments, decreased levels of GroEL have been shown to lead to a reduced growth rate and increased amounts of a number of heat-shock polypeptides, apparently the result of an attempt by the cell to compensate for the GroEL effect [23].

An important feature of heat-shock protein Hsp70 family manipulation in *E. coli* is that overproduction of these proteins must be balanced. Blum *et al.* [24] demonstrated that overproduction of DnaK alone can lead to plasmid instability, defective filamentation and ultimately bactericidal effects in *E. coli*, whereas co-overproduction of DnaJ can partially or completely overcome all of these effects. In accordance with this is the finding that the simultaneous overproduction of DnaK, DnaJ and GrpE in *E. coli* appeared to lead to somewhat increased solubility of several co-expressed protein tyrosine kinases [14•], even though the total protein level is decreased. DnaK alone also leads to a decrease in the size of inclusion bodies of human growth hormone and increases the ratio of soluble/insoluble protein in directly comparable experiments by about a factor of two [25].

Other studies have investigated the effects of overexpression of DnaK alone on both heterologous protein secretion and folding. In cytoplasmically expressed procollagenase, DnaK overexpression increases the half-life and thus accumulation of both soluble and insoluble procollagenase. For secreted procollagenase, export is inhibited upon DnaK overproduction [20]. In contrast, DnaK and GroEL appear to lead to an increase in signal sequence processing of LamB–β-galactosidase fusions, which cannot normally be transported [26]. DnaK has no positive effect on penicillin-binding protein secretion, however [27•]. The effect of DnaK and DnaJ on secretion of human granulocyte stimulating factor has also been investigated [28]. In contrast to SecB and GroEL/ES, DnaK and DnaJ act synergistically to increase signal 5 sequence processing (and thus presumably transport). Even so, only a small fraction of the protein is soluble, and of this, only a small fraction is processed. In conclusion, therefore, overexpression of components of the Hsp70 complex can result in very heterogeneous effects.

Many other *E. coli* stress proteins induced by stress conditions, such as cold [29], superoxide [30], nutrient starvation [31], pressure [32,33], phage production [34,35], inclusion body formation [36], or universal stress conditions [37], have been reported and, indeed, continue to be identified and characterized. Small heat shock proteins [38] and a member of the Hsp90 class (HtpG) [39] have also been found in *E. coli*, as well as a homolog of DnaK [40] and a stationary phase analog of DnaJ [41]. Furthermore, novel specialized chaperones [42], notably those for pilus assembly (reviewed in [43]), continue to be discovered, but it is not yet clear which of these proteins have a direct or indirect role in general protein folding. Thus, we consider these to be currently outside the scope of this review because studies of the effects of their overproduction on heterologous protein folding in *E. coli* have not been reported to date.

Despite the suggestion by some authors that overproduction of GroEL/ES (even given its already very high concentration in the cell) or DnaK/DnaJ/GrpE might be of general use in increasing the solubility in *E. coli* of otherwise insoluble heterologous proteins, it is clear

that the effects reviewed here are highly protein-specific. Furthermore, it is likely that several factors may often need to cooperate for an effect on heterologous protein production to be seen. Thus, a dramatic general effect of overproducing any one or more of these chaperones on the folding of all heterologous proteins expressed in *E. coli* may probably not be realistically expected.

Improving the process of secretion

A potential problem inherent in the periplasmic expression of heterologous proteins in *E. coli* is translocation across the cytoplasmic membrane. Although transport of foreign proteins in *E. coli* has been exploited for some time, it is unclear whether the detailed requirements for the mature protein sequence of eukaryotic and Gram-positive proteins are precisely identical or only roughly similar to those for *E. coli* itself. The polypeptides must remain at least partially unfolded before membrane translocation, and in eukaryotes, transport is largely co-translational, whereas in *E. coli* both co- and post-translational export are known. In the case of most *E. coli* proteins, export from the cytoplasm to the periplasm depends on the machinery encoded by the '*sec*' genes (discovered as a result of mutations that block export), a secretion system composed of several proteins, mostly located in the membrane, which together facilitate membrane translocation by nascent polypeptides (see Fig. 3) [44•].

Manipulation of a number of components of the secretion system has been carried out in an attempt to improve heterologous protein production in *E. coli*. Export of a recombinant penicillin-binding protein in *E. coli* has been found to be facilitated by increasing the SecB content of the producing strain [27•], and co-expression of the *E. coli* SecB protein in *Bacillus subtilis* has been shown to result in stimulation of export of recombinant *E. coli* maltose-binding protein [45]. In another study, co-overexpression of the SecD and SecF proteins has been found to result in increased translocation of proteins with mutant signal sequences and faster secretion of wild-type proteins in *E. coli*, whereas depletion of the two proteins led to a greatly reduced level of translocation [46•]. SecD and perhaps SecF are suspected to participate in protein release steps, which may require an electrochemical gradient [46•,47,48]. A direct role in periplasmic folding may of course also be envisaged, and these proteins might

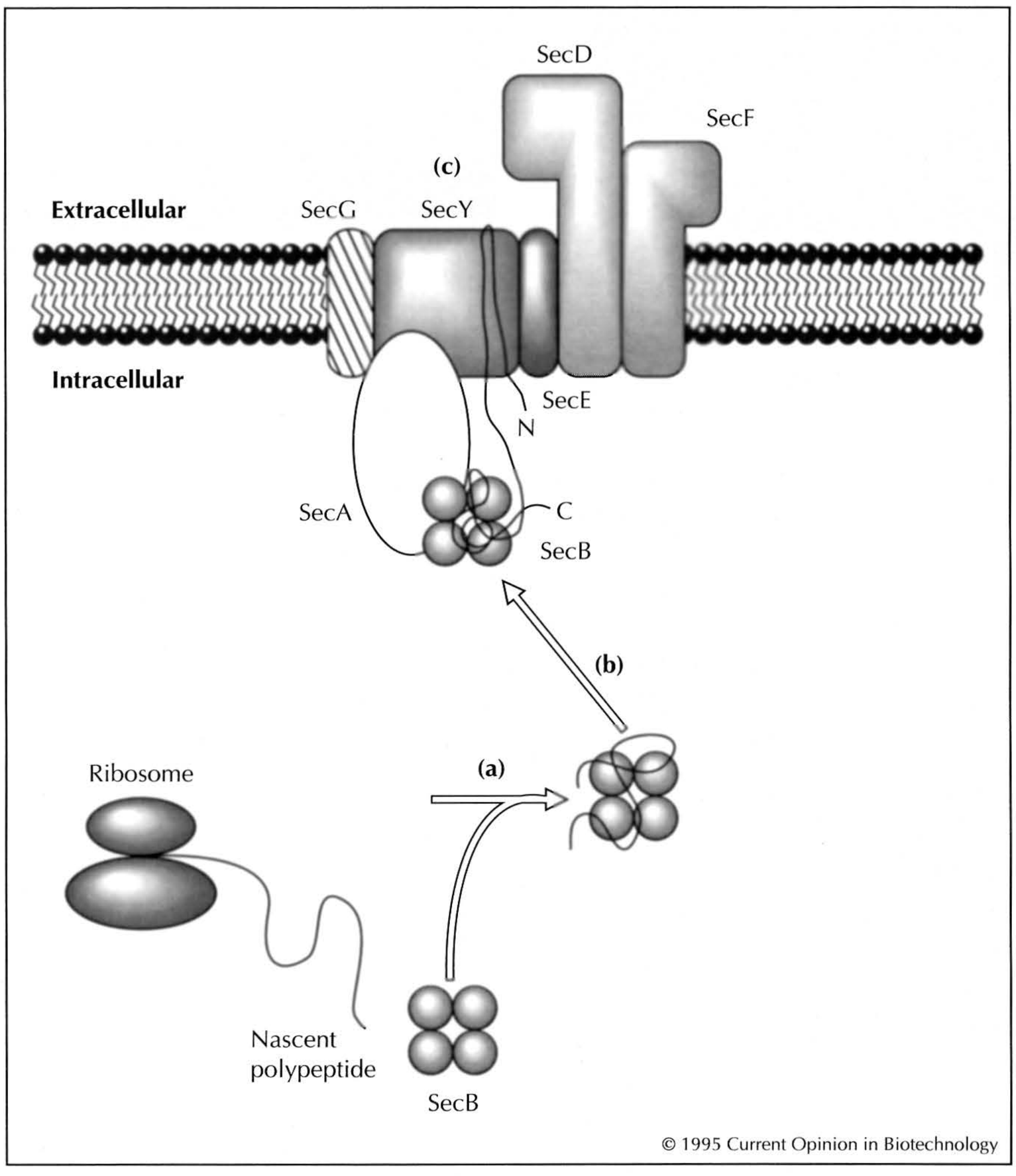

Fig. 3. Proposed model for polypeptide translocation of the cytoplasmic membrane in *E. coli*. Translocation of the *E. coli* membrane by secreted proteins is normally mediated by the *sec* system, although its universal use is not proven [56]. The Sec proteins SecB and (perhaps) SecA, as well as, most likely, other chaperones, maintain the nascent polypeptide in a translocation-competent state before directing its insertion into the membrane. SecB, a cytoplasmic factor involved in protein export, is thought **(a)** to first interact with free polypeptide chains in the unfolded protein and **(b)** to deliver the protein to SecA in a translocation-competent form [75], although it does not react with all secreted proteins. SecA is a peripheral membrane protein with ATPase activity that interacts with SecB and the polypeptide chain to be translocated. SecA then accompanies the chain into the membrane so that it transiently becomes a transmembrane protein itself [76]. **(c)** The 'translocase' component of the Sec machinery is believed to be made up of three proteins, SecY, SecE, and SecG. In eukaryotes, it has been shown by fluorescence labeling of precursors that proteins do indeed move through a gated aqueous pore [77]. Further proteins involved in the translocation process, probably at the release stage, or even the folding stage, are SecD and SecF.

function as periplasmic, though membrane-bound, chaperones with a large periplasmic domain [49•].

The influence of the sequence of heterologous proteins on their translocation efficiency in *E. coli* has also been repeatedly emphasized, with lysines and especially arginines at the beginning of the mature protein resulting in more difficult crossing of the membrane during *sec*-dependent translocation ([50] and references therein). The *prlA* suppressor mutations of the SecY protein have been demonstrated to dramatically alleviate filamentous phage growth defects associated with positively charged residues close to the signal sequence cleavage site of the mature pIII coat protein [51•]. Such *E. coli* suppressor strains might prove even more useful in general by overcoming secretion problems associated with the production in *E. coli* of heterologous proteins that carry positively charged residues near the amino terminus of the mature protein (see also [52••]).

Production of growth hormones in the periplasm of *E. coli* has been found to lead to a 30-fold higher amount of a human hormone than of a porcine homolog under identical conditions. By making hybrid proteins, two helices were found to be responsible for this effect. Unfortunately, from the published data it is difficult to distinguish between an effect on secretion and periplasmic folding [53], and thus the cause of the effect may be similar to that seen in [12•].

The *prl* (protein localization) mutations — named differently from the *sec* mutations by virtue of having been found from different genetic selections — have been discovered through their restoration of export to proteins with secretion defects. Several Prl proteins are now known to be identical to Sec proteins (PrlD, PrlG and PrlA are identical to SecA, SecE and SecY, respectively), though others are different. The effect of overproduction of one of these Prl proteins, PrlF, which was originally isolated as a suppressor of secreted fusion protein lethality in *E. coli*, on the production of a *Bacillus stearothermophilus* α-amylase enzyme has recently been demonstrated by Minas and Bailey [54••]. They found that although *E. coli* cell viability was severely impaired upon overproduction of the α-amylase enzyme, co-overexpression of the *prlF* gene led to improved cell viability, higher α-amylase yields, and greater specific enzyme production [54••]. PrlF may act by activating the Lon protease [55] to degrade the otherwise lethal accumulation of the precursor, even though higher total protease activity cannot be measured, and the phenotypic effect is different from increasing the Lon concentration by heat shock.

Another investigation with PrlA4, a mutant form of the SecY protein, has shown that its co-overexpression with SecE leads to an increased yield of recombinant human interleukin (IL)-6 in the periplasm of *E. coli* [52••]. From the presumed function of SecY and SecE in the actual transport process, this appears to be a case where the transport step itself is limiting.

The existence of a 'non-classical' or *sec*-independent pathway of secretion in *E. coli* has been proposed on the basis of the surprising homologies between several components and eukaryotic transport factors ([56] and references therein). To the best of our knowledge, however, no positive effect on heterologous protein production has been reported to date as a result of overexpression of any of these pathway components.

These various results indicate that, at least for some heterologous proteins for which translocation appears to be the limiting step in *E. coli* expression, overexpression of one or more secretory proteins or accessory factors, which may clean up the debris causing stress on the cell, might allow such problems to be partially overcome. The amount and/or proportion of correctly processed and folded gene products may thus be increased, or at least the time of growth might be extended. The limiting step in the production process frequently occurs not in membrane passage itself, however, but in the subsequent folding in the periplasm (see next section).

Finally, these results of manipulation of the secretion apparatus are by no means universal and the protein-specificity of the effects is demonstrated by those examples where the level of the recombinant protein could not be improved.

Folding in the periplasm

Following translocation of heterologous proteins through the cytoplasmic membrane, folding takes place in the periplasmic space. The periplasm is the compartment of *E. coli* in which, because of its non-reducing environment, disulfide-containing heterologous proteins must generally be expressed (for an alternative strategy, see [1,2]). To date, no general molecular chaperones that prevent non-productive folding reactions have been identified in the periplasm, though SecD and SecF, transmembrane proteins with extended periplasmic domains, might help to prevent aggregation of newly translocated proteins during the initial stages of folding [46•,49•]. Furthermore, ATP, which is so closely associated with chaperone action in the cytoplasm (see Fig. 2), is not thought to exist in the periplasmic space [49•].

Although a wide variety of heterologous proteins have now been produced in the periplasm of *E. coli*, studies of their folding are limited. In the case of antibody fragment [18] and T-cell receptor (TCR) fragment [57•] production, the limiting step has been shown to be folding of the fragments in the periplasm. In the latter study, the authors attempted to achieve a global activation of potential chaperones — including previously unidentified ones — by overexpression at low temperature of the heat-shock factor σ^{32} (encoded by *rpoH*), the sigma subunit of the RNA polymerase. Co-overexpression of σ^{32} and the *E. coli* periplasmic disulfide isomerase DsbA was found to result in an increased yield of one of the expressed TCR fragments, concomitant with reduced proteolytic degradation [57•,58].

A further, albeit related, problem in the expression of heterologous proteins in the *E. coli* periplasm is

that of degradation by host proteases. In a recent paper, Meerman and Georgiou [59•] have described a set of *E. coli* strains deficient in all (then) known cell envelope proteases as well as σ^{32} (*rpoH*), which controls several heat-shock regulated proteases such as Lon [59•]. They also report a dramatic increase in the expression of various heterologous proteins in these mutant strains of *E. coli*. Newly identified homologs of the periplasmic protease DegP (SWISS-PROT accession numbers: P31137, P39099 and P39436) may also be worth investigating for similar effects on heterologous protein production. When manipulating host cell proteases, however, one must bear in mind that proteolytic degradation may be a symptom, rather than a cause, of folding problems, serving to remove misfolded material (see also [54••]) and already formed aggregates whose accumulation in the absence of proteases might prove to be toxic. Thus, the possibility exists that a primary effect of protease depletion may be to actually increase toxic effects of heterologous protein expression on the cell.

Disulfide bond isomerization and proline *cis/trans* isomerization

The molecular chaperones discussed thus far are involved in increasing the yield, rather than the rate of *in vivo* folding of heterologous proteins, but *E. coli* also possesses several proteins that catalyze reactions that might otherwise limit the rate of the folding process. One such protein is DsbA, a periplasmic protein that is involved in disulfide bond formation in the non-reducing environment of the *E. coli* periplasm [60•,61]. DsbA, present in an oxidized, but strained form, probably acts mainly as an oxidant. It is reoxidized by DsbB, a membrane-spanning protein which might act by coupling its own disulfide bond formation to the electron transport chain [60•]. DsbC, a further component of the disulfide bond formation machinery, may function by complementing the action of DsbA, and also by catalyzing isomerization, rather than net formation, of disulfide bonds by virtue of both its oxidized and its reduced forms being present at the oxidizing redox potential of the periplasm [60•,62]. Another membrane-bound protein with a similar active site has been discovered to be involved in periplasmic cytochrome *c* biogenesis [63•], where it may serve to facilitate the coupling of heme to cysteines. It has also turned up in the screening of Raina and colleagues [62], however, who have named it DsbD. Indeed, it has recently been shown to be required for human placental alkaline phosphatase production in *E. coli* [64], even though neither its overexpression nor that of DsbA [65] had any effect.

Several studies have investigated the effects of overexpression of these folding catalysts on the expression of heterologous proteins in *E. coli*. In the case of antibody fragment production in the *E. coli* periplasm, DsbA was found to be necessary for the correct formation of disulfide bonds [18]. This requirement for DsbA has also been demonstrated for TCR fragments [57•] (albeit in synergy with an unknown factor which was induced by σ^{32}), and bovine pancreatic trypsin inhibitor (BPTI) [66•]. Nevertheless, overexpression of DsbA increased the amount neither of correctly folded antibody fragments [18] nor of human placental alkaline phosphatase [64]. In contrast, its overproduction in combination with a number of other factors was found to result in increased yields of TCR fragments [57•] and of a protease inhibitor [67].

In general, large proteins with few disulfides are probably less likely to be limited by disulfide formation than are small disulfide-rich proteins, for which rapid rearrangement of disulfide bonds may be essential for attaining the native structure.

Conversely, some proteins have been found to require the absence of DsbA and DsbB for their production in functional form; expression of a metallo-β-lactamase from *Bacteroides fragilis* in wild-type *E. coli*, in the presence of DsbA and DsbB, has been shown to yield an enzyme with aberrant disulfide bonds that is proteolytically unstable [68•]. In the absence of the two Dsb proteins, however, the β-lactamase assumed a conformation that contained no disulfide bonds and was proteolytically stable.

The other commonly described catalyst of protein folding is peptidyl prolyl *cis/trans* isomerase (PPIase), which catalyzes the isomerization of X-Pro bonds (three-letter amino acid code, where X is any residue), a potentially rate-limiting step in the folding of many proteins. An investigation *in vitro* of the role of three PPIases in the folding of carbonic anhydrase II, however, found no increase in the folding yield [69]. This indicates that the sole role of these isomerases in protein folding must be to accelerate isomerization and, currently, no basis exists for their suggested role as chaperones (i.e. preventing aggregation [69]). Indeed, no dramatic effect was observed upon disruption of the periplasmic *rotA* gene, which encodes the *E. coli* rotamase enzyme (M Kleerebezem *et al.*, unpublished data).

E. coli probably has at least eight PPIases (SWISS-PROT accession numbers: P20752, P23869, P39159, P21202, P22563, P39311, P30856 and P22257). This is mostly deduced from sequence homology because not all have been tested for activity yet, but more enzymes may also be discovered. PPIases belong to three different protein families, with two periplasmic and six cytoplasmic proteins (reviewed in [70•]; G Fischer, personal communication). We are not aware of systematic studies on the overexpression of all of these enzymes to test any hypothetical specificity on the folding of recombinant proteins *in vivo*. Single enzymes have been overexpressed, however (see below).

In a study with secreted recombinant antibody fragments, results similar to those found with DsbA were obtained [18]. Overexpression of the periplasmic *E. coli* rotamase enzyme had no effect on the yield of active

soluble Fab fragments and only a minor effect on single-chain Fv fragments. A similar study also found that overproduction of either PPIase or DsbA had no effect on the yield in *E. coli* of periplasmic soluble human placental alkaline phosphatase protein from which the amino-terminal transmembrane helix was deleted [64]. This was undoubtedly a reasonable experiment to try in the case of an enzyme containing four disulfide bonds and 27 proline residues, but it shows that cellular protein folding is highly complex and that non-chemical steps are frequently limiting the folding yield.

Finally, if one considers the pathway of protein folding as a series of essentially sequential, potentially limiting, steps, it is obvious that manipulation of any stage of the pathway, such as disulfide bond or proline isomerization, can lead to improved folding only when other steps do not lead to the accumulation of aggregation-prone intermediates. Therefore, should other folding bottlenecks be identified and/or overcome (by catalyzing the relevant steps or plugging deleterious pathways), it is possible that folding modulators, such as Dsb proteins or PPIases, whose overproduction had previously been found to have no effect, might be beneficially reinvestigated.

Conclusions

The effects of folding catalyst overexpression reviewed here quite clearly demonstrate the potential of this approach in partially solving at least some of the problems in the folding of several heterologous proteins in *E. coli*. For each success story, however, there are many more (often unpublished) failures at present, and it is this unpredictability which dictates that, for now, the generation of a better understanding of *E. coli*'s protein folding mechanisms remains a priority. The hope is obviously that a clearer picture of the *in vivo* process will point the way to a more rational approach to improving the folding of heterologous proteins in *E. coli*. Furthermore, it will be interesting to note whether overexpression of folding modulators might become an even more fruitful approach in the future when current folding bottlenecks in some individual proteins have been eliminated by rational engineering or evolutionary approaches to changing the protein sequence. In the meantime, sufficient success has been achieved to date with individual proteins to encourage any researcher to use a battery of folding modulators to attempt to solve folding problems.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Derman AI, Prinz WA, Belin D, Beckwith J: **Mutations that allow disulfide bond formation in the cytoplasm of *Escherichia coli*.** *Science* 1993, **262**:1744–1747.

2. Proba K, Ge L, Plückthun A: **Functional antibody single-chain fragments from the cytoplasm of *Escherichia coli*: influence of thioredoxin reductase (TrxB).** *Gene* 1995, **159**:203–207.

3. Withka JM, Wyss DF, Wagner G, Arulanandam AR, Reinherz EL, Recny MA: **Structure of the glycosylated adhesion domain of human T lymphocyte glycoprotein CD2.** *Structure* 1993, **1**:69–81.

4. Bergeron JJ, Brenner MB, Thomas DY, Williams DB: **Calnexin: a membrane-bound chaperone of the endoplasmic reticulum.** *Trends Biochem Sci* 1994, **19**:124–128.

5. • Hockney RC: **Recent developments in heterologous protein production in *Escherichia coli*.** *Trends Biotechnol* 1994, **12**:456–463.
A useful review of recent progress in the field, stressing various factors that contribute to heterologous protein production.

6. Mitraki A, King J: **Amino acid substitutions influencing intracellular protein folding pathways.** *FEBS Lett* 1992, **307**:20–25.

7. Chrunyk BA, Wetzel R: **Breakdown in the relationship between thermal and thermodynamic stability in an interleukin-1β point mutant in a surface loop.** *Protein Eng* 1993, **6**:733–738.

8. Mitraki A, Danner M, King J, Seckler R: **Temperature-sensitive mutations and second-site suppressor mutations affect folding of the P22 tailspike protein *in vitro*.** *J Biol Chem* 1993, **268**:20071–20075.

9. • Dale GE, Broger C, Langen H, D'Arcy A, Stüber D: **Improving protein solubility through rationally designed amino acid replacements: solubilization of the trimethoprim-resistant type S1 dihydrofolate reductase.** *Protein Eng* 1994, **7**:933–939.
Manipulation of the dihydrofolate reductase amino terminus leads to a 20-fold increase in expression by elimination of an internal translation start. Further directed mutagenesis yields soluble protein with a 264-fold increase in activity.

10. Goldenberg DP, Creighton TE: **Phage tailspike protein. A fishy tail of protein folding.** *Curr Biol* 1994, **4**:1026–1029.

11. Wetzel R, Chrunyk BA: **Inclusion body formation by interleukin-1β depends on the thermal sensitivity of a folding intermediate.** *FEBS Lett* 1994, **350**:245–248.

12. • Knappik A, Plückthun A: **Engineered turns of a recombinant antibody improve its *in vivo* folding.** *Protein Eng* 1995, **8**:81–89.
This study shows that periplasmic protein folding can limit expression and that rational engineering can be used to improve it. It also shows how single point mutations in the recombinant protein can generate phenotypic consequences such as leakiness of the outer membrane.

13. Goloubinoff P, Gatenby A, Lorimer G: **GroE heat shock proteins promote assembly of foreign prokaryotic ribulose bisphosphate carboxylase oligomers in *E. coli*.** *Nature* 1989, **337**:44–47.

14. • Caspers P, Stieger M, Burn P: **Overproduction of bacterial chaperones improves the solubility of recombinant protein tyrosine kinases in *Escherichia coli*.** *Cell Mol Biol* 1994, **40**:635–644.
Simultaneous overproduction of Hsp70 (DnaK, DnaJ and GrpE) or Hsp60 (GroEL and GroES) families in engineered *E. coli* strains leads to a (limited) increase in solubility of several kinases normally poorly soluble when overproduced. Co-expression of DnaK, DnaJ and GrpE leads to a significant reduction in total expression.

15. • Amrein KE, Takacs B, Steiger M, Molnos J, Flint NA, Burn P: **Purification and characterization of recombinant human p50csk protein-tyrosine kinase from an *Escherichia coli* expression system overproducing the bacterial chaperones GroES and GroEL.** *Proc Natl Acad Sci USA* 1995, **92**:1048–1052.
Co-overproduction of GroEL/ES leads to increased solubility of tyrosine kinase expressed in *E. coli*.

16. • Ashiuchi M, Yoshimura T, Kitamura T, Kawata Y, Nagai J, Gorlatov S, Esaki N, Soda K: ***In vivo* effect of GroESL on the folding of glutamate racemase of *Escherichia coli*.** *J Biochem* 1995, **117**:495–498.
Co-overproduction of GroEL/ES changes glutamate racemase from mostly insoluble protein in inclusion bodies to mostly soluble active enzyme.

This results in a net two- to threefold increase in production of the enzyme.

17. Dale GE, Schönfeld H-J, Langen H, Stieger M: **Increased
• solubility of trimethoprim-resistant type S1 DHFR from *Staphylococcus aureus* in *Escherichia coli* cells overproducing the chaperonins GroEL and GroES.** *Protein Eng* 1994, **7**:925–931.
Higher GroEL/ES levels in the cell are shown to increase levels of soluble and functional dihydrofolate reductase. Increased yields of refolding of the reductase are also demonstrated *in vitro*.

18. Knappik A, Krebber C, Plückthun A: **The effect of folding catalysts on the *in vivo* folding process of different antibody fragments in *Escherichia coli*.** *Biotechnology* 1993, **11**:77–83.

19. Fong G, Bridger WA: **Folding and assembly of the *Escherichia coli* succinyl-CoA synthetase heterotetramer without participation of molecular chaperones.** *Biochemistry* 1992, **31**:5661–5664.

20. Lee SC, Olins PO: **Effect of overproduction of heat shock chaperones GroESL and DnaK on human procollagenase production in *Escherichia coli*.** *J Biol Chem* 1992, **267**:2849–2852.

21. Dueñas M, Vázquez J, Ayala M, Söderlind E, Ohlin M, Pérez L, Borrebaeck CAK, Gavilondo JV: **Intra- and extracellular expression of an scFv antibody fragment in *E. coli*: effect of bacterial strains and pathway engineering using GroEL/ES chaperonins.** *Biotechniques* 1994, **16**:476–483.

22. Cloney LP, Bekkaoui DR, Hemmingsen SM: **Co-expression of plastid chaperonin genes and a synthetic plant Rubisco operon in *Escherichia coli*.** *Plant Mol Biol* 1993, **23**:1285–1290.

23. Kanemori M, Mori H, Yura T: **Effects of reduced levels of GroE chaperones on protein metabolism: enhanced synthesis of heat shock proteins during steady-state growth of *Escherichia coli*.** *J Bacteriol* 1994, **176**:4235–4242.

24. Blum P, Ory J, Bauernfeind J, Krska J: **Physiological consequences of DnaK and DnaJ overproduction in *Escherichia coli*.** *J Bacteriol* 1992, **174**:7436–7444.

25. Blum P, Velligan M, Liu N, Martin A: **DnaK-mediated alterations in human growth hormone protein inclusion bodies.** *Biotechnology* 1992, **10**:301–304.

26. Phillips GJ, Silhavy TJ: **Heat-shock proteins DnaK and GroEL facilitate export of *lacZ* hybrid proteins in *E. coli*.** *Nature* 1990, **344**:882–884.

27. Fraipont C, Adam M, Nguyen-Disteche M, Keck W, Van
• Beeumen J, Ayala JA, Granier B, Hara H, Ghuysen JM: **Engineering and overexpression of periplasmic forms of the penicillin-binding protein 3 of *Escherichia coli*.** *Biochem J* 1994, **15**:189–195.
These authors show that increases in the SecB content of *E. coli* lead to improved translocation of secreted forms of PBP-3, in which the amino-terminal transmembrane region is replaced by a signal sequence.

28. Pérez-Pérez J, Martinéz-Caja C, Barbero JL, Gutierrez J: **DnaK/DnaJ supplementation improves the periplasmic production of human granulocyte-colony stimulating factor in *Escherichia coli*.** *Biochem Biophys Res Comm* 1995, **210**:524–529.

29. Jones PG, Inouye M: **The cold shock response — a hot topic.** *Mol Microbiol* 1994, **11**:811–818.

30. Nunoshiba T, Hidalgo E, Amabile Cuevas CF, Demple B: **Two-stage control of an oxidative stress regulon: the *Escherichia coli* SoxR protein triggers redox-inducible expression of the *soxS* regulatory gene.** *J Bacteriol* 1992, **174**:6054–6060.

31. Yamashino T, Kakeda M, Ueguchi C, Mizuno T: **An analogue of the DnaJ molecular chaperone whose expression is controlled by sigma s during the stationary phase and phosphate starvation in *Escherichia coli*.** *Mol Microbiol* 1994, **13**:475–483.

32. Welch TJ, Farewell A, Neidhardt FC, Bartlett DH: **Stress response of *Escherichia coli* to elevated hydrostatic pressure.** *J Bacteriol* 1993, **175**:7170–7177.

33. Gross M, Kosmowsky IJ, Lorenz R, Molitoris HP, Jaenicke R: **Response of bacteria and fungi to high-pressure stress as investigated by two-dimensional polyacrylamide gel electrophoresis.** *Electrophoresis* 1994, **15**:1559–1565.

34. Weiner L, Model P: **Role of an *Escherichia coli* stress-response operon in stationary-phase survival.** *Proc Natl Acad Sci USA* 1994, **91**:2191–2195.

35. Kleerebezem M, Tommassen J: **Expression of the *pspA* gene stimulates efficient protein export in *Escherichia coli*.** *Mol Microbiol* 1993, **7**:947–956.

36. Allen SP, Polazzi JO, Gierse JK, Easton AM: **Two novel heat shock genes encoding proteins produced in response to heterologous protein expression in *Escherichia coli*.** *J Bacteriol* 1992, **174**:6938–6947.

37. Nystrom T, Neidhardt FC: **Expression and role of the universal stress protein, UspA, of *Escherichia coli* during growth arrest.** *Mol Microbiol* 1994, **11**:537–544.

38. Chuang SE, Burland V, Plunkett G III, Daniels DL, Blattner FR: **Sequence analysis of four new heat-shock genes constituting the *hslTS/ibpAB* and *hslUV* operons in *Escherichia coli*.** *Gene* 1993, **134**:1–6.

39. Nadeau K, Das A, Walsh CT: **Hsp90 chaperonins possess ATPase activity and bind heat shock transcription factors and peptidyl prolyl isomerases.** *J Biol Chem* 1993, **268**:1479–1487.

40. Seaton BL, Vickery LE: **A gene encoding a DnaK/Hsp70 homolog in *Escherichia coli*.** *Proc Natl Acad Sci* 1994, **91**:2066–2070.

41. Ueguchi C, Shiozawa T, Kekeda M, Yamada H, Mizuno, T: **A study of the double mutation of DnaJ and CbpA, whose gene products function as molecular chaperones in *Escherichia coli*.** *J Bacteriol* 1995, **177**:3894–3896.

42. Aamand JL, Hobson AH, Buckley CM, Jorgensen ST, Diderichsen B, McConnell DJ: **Chaperone-mediated activation *in vivo* of a *Pseudomonas cepacia* lipase.** *Mol Gen Genet* 1994, **245**:556–564.

43. Hultgren SJ, Jacob-Dubuisson F, Jones CH, Branden CI: **PapD and superfamily of periplasmic immunoglobulin-like pilus chaperones.** *Adv Protein Chem* 1993, **44**:99–123.

44. Blight MA, Chervaux C, Holland IB: **Protein secretion pathways
• in *Escherichia coli*.** *Curr Opin Biotechnol* 1994, **5**:468–474.
A detailed review of secretion to the medium of heterologous proteins produced in *E. coli*.

45. Collier DN: **Expression of *Escherichia coli secB* in *Bacillus subtilis* facilitates secretion of the *secB*-dependent maltose-binding protein of *E. coli*.** *J Bacteriol* 1994, **176**:4937–4940.

46. Pogliano JA, Beckwith J: **SecD and SecF facilitate protein
• export in *Escherichia coli*.** *EMBO J* 1994, **13**:554–561.
A careful study of the function of SecD and SecF, both by overexpression and deletion of the operon. A null mutant in either protein is barely viable, whereas overexpression improves export of proteins with mutant signal sequences.

47. Arkowitz RA, Wickner W: **SecD and SecF are required for the proton electrochemical gradient stimulation of preprotein translocation.** *EMBO J* 1994, **13**:954–963.

48. Matsuyama S, Fujita Y, Mizushima S: **SecD is involved in the release of translocated secretory proteins from the cytoplasmic membrane of *Escherichia coli*.** *EMBO J* 1992, **12**:265–270.

49. Wülfing C, Plückthun A: **Protein folding in the periplasm of
• *Escherichia coli*.** *Mol Microbiol* 1994, **12**:685–692.
A critical review of current knowledge of the periplasmic folding process.

50. Andersson H, Von Heijne G: **Positively charged residues influence the degree of SecA dependence in protein translocation across the *E. coli* inner membrane.** *FEBS Lett* 1994, **347**:169–172.

51. Peters EA, Schatz PJ, Johnson SS, Dower WJ: **Membrane
• insertion defects caused by positive charges in the early mature region of protein pIII of filamentous phage fd can be corrected by *prlA* suppressors.** *J Bacteriol* 1994, **176**:4296–4305.

Suppressor *E. coli* strains in the *prlA* (*secY*) gene are found to overcome growth restrictions caused by positive charges introduced next to the signal sequence cleavage site in pIII by displayed peptides. It is proposed that these strains may be useful in increasing the diversity of phage peptide libraries.

52. •• Pérez-Pérez J, Márquez G, Barbero J-L, Gutiérrez J: **Increasing the efficiency of protein export in *Escherichia coli*.** *Biotechnology* 1994, **12**:178–180.

In this paper, the authors achieve higher periplasmic yields of IL by manipulating the translocation machinery, specifically by overexpression of SecE and the PrlA4 mutant form of SecY. This paper has broadened the range of factors to consider for overexpression.

53. Cheah K-C, Harrison S, King R, Crocker L, Wells JRE, Robins A: **Secretion of eukaryotic growth hormones in *Escherichia coli* is influenced by the sequence of the mature proteins.** *Gene* 1994, **138**:9–15.

54. •• Minas W, Bailey J: **Co-overexpression of PrlF increases cell viability and enzyme yields in recombinant *Escherichia coli* expressing *Bacillus stearothermophilus* α-amylase.** *Biotechnol Prog* 1995, **11**:403–411.

The classic problems of low yield and slowed host cell growth upon overexpression of a heterologous protein are overcome by increasing the level of PrlF in the cells. PrlF may help the Lon protease to degrade untranslocated precursor.

55. Snyder WB, Silhavy TJ: **Enhanced export of beta-galactosidase fusion proteins in *prlF* mutants is Lon dependent.** *J Bacteriol* 1992, **174**:5661–5668.

56. Dobberstein B: **Protein transport. On the beaten pathway.** *Nature* 1994, **367**:599–600.

57. • Wülfing C, Plückthun A: **Correctly folded T-cell receptor fragments in the periplasm of *Escherichia coli*.** *J Mol Biol* 1994, **242**:655–669.

Improved periplasmic folding of TCRs in *E. coli* is achieved by induction of the heat-shock response at low temperature and concomitant overexpression of DsbA. This strategy should induce a whole variety of chaperones, including yet undiscovered periplasmic ones. The low temperature improves the intrinsic tendencies of the folding intermediates to go to the native state.

58. Wülfing C, Plückthun A: **Soluble T-cell receptor fragments—guidance of folding and assembly.** *Immunologist* 1995, **3**:59–66.

59. • Meerman HJ, Georgoiu G: **Construction and characterization of a set of *E. coli* strains deficient in all known loci affecting the proteolytic stability of secreted recombinant proteins.** *Biotechnology* 1994, **12**:1107–1110.

A useful study of the consequences of protease deficiencies on secreted proteins. The genes for proteases *degP*, *ptr*, *trp (prc)* and *ompT* are disrupted in various combinations. Increased production yields of labile fusion proteins are observed.

60. • Bardwell JCA: **Building bridges: disulphide bond formation in the cell.** *Mol Microbiol* 1994, **14**:199–205.

A comprehensive review of the *E. coli* disulfide bond formation machinery.

61. Bardwell JCA, Lee J-O, Jander G, Martin N, Belin D, Beckwith J: **A pathway for disulfide bond formation *in vivo*.** *Proc Natl Acad Sci USA* 1993, **90**:1038–1042.

62. Missiakas D, Georgopoulos C, Raina S: **The *Escherichia coli dsbC* (*xprA*) gene encodes a periplasmic protein involved in disulfide bond formation.** *EMBO J* 1994, **13**:2013–2020.

63. • Crooke H, Cole J: **The biogenesis of *c*-type cytochromes in *Escherichia coli* requires a membrane-bound protein, DipZ, with a protein disulfide isomerase-like domain.** *Mol Microbiol* 1995, **15**:1139–1150.

A description of the identification of DipZ (DsbD) and characterization of its involvement in the covalent attachment of heme to the appropriate cysteine residues of cytochrome *c*. This protein may protect cytochrome cysteines from illegitimate oxidation.

64. Beck R, Crooke H, Jarsch M, Cole J, Burtscher H: **Mutation in *dipZ* leads to reduced production of active human placental alkaline phosphatase in *Escherichia coli*.** *FEMS Microbiol Lett* 1994, **124**:209–214.

65. Beck R, Burtscher H: **Expression of human placental alkaline phosphatase in *Escherichia coli*.** *Protein Expr Purif* 1994, **5**:192–197.

66. • Ostermeier M, Georgoiu G: **The folding of bovine pancreatic trypsin inhibitor in the *Escherichia coli* periplasm.** *J Biol Chem* 1994, **269**:21072–21077.

A study of periplasmic folding of BPTI that demonstrates an absolute dependence of folding on DsbA and the existence of two intermediates. The DsbA-catalyzed reaction is shown to be faster than the uncatalyzed reaction. Reduced/oxidized glutathione additions to the growing culture have no effect.

67. Wunderlich M, Glockshuber R: ***In vivo* control of redox potential during protein folding catalyzed by bacterial protein disulfide-isomerase (DsbA).** *J Biol Chem* 1993, **268**:24547–24550.

68. • Alksne LE, Keeney D, Rasmussen BA: **A mutation in either *dsbA* or *dsbB*, a gene encoding a component of a periplasmic disulfide bond-catalyzing system, is required for high-level expression of the *Bacteroides fragilis* metallo-β-lactamase, CcrA, in *Escherichia coli*.** *J Bacteriol* 1995, **177**:462–464.

An important study in that it describes the production in *E. coli* of heterologous proteins that require the absence of DsbA and DsbB for functional expression.

69. Kern G, Kern D, Schmid FX, Fischer G: **Reassessment of the putative chaperone function of prolyl-*cis/trans*-isomerases.** *FEBS Lett* 1994, **348**:145–148.

70. • Fischer G: **Peptidyl-prolyl *cis/trans* isomerases and their effectors.** *Angew Chemie Int* 1994, **33**:1415–1436.

A comprehensive review of PPIases, their presumed biological roles, and their inhibitors.

71. Anfinsen CB: **Principles that govern the folding of protein chains.** *Science* 1973, **181**:223–230.

72. Gilbert HF: **Protein chaperones and protein folding.** *Curr Opin Biotechnol* 1994, **5**:534–539.

73. Martin J, Hartl FU: **Molecular chaperones in cellular protein folding.** *Bioessays* 1994, **16**:689–692.

74. Braig K, Otiwinowski Z, Hegde R, Boisvert DC, Joachimiak A, Horwich AL, Sigler PB: **The crystal structure of the bacterial chaperonin GroEL at 2.8 Å.** *Nature* 1994, **371**:578–586.

75. Randall LL, Hardy SJS: **High selectivity with low specificity: how SecB has solved the paradox of chaperone binding.** *Trends Biochem Sci* 1995, **20**:65–69.

76. Schekman R: **Translocation gets a push.** *Cell* 1994, **78**:911–913.

77. Crowley KS, Liao S, Worrell VE, Reinhart GD, Johnson AE: **Secretory proteins move through the endoplasmic reticulum membrane via an aqueous, gated pore.** *Cell* 1994, **78**:461–471.

JG Wall and A Plückthun, Biochemisches Institut der Universität Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland.
A Plückthun E-mail: plueckthun@biocfebs.unizh.ch