Molecular Microbiology (1994) **12**(5), 685–692

MicroReview

Protein folding in the periplasm of *Escherichia coli*

Christoph Wülfing and Andreas Plückthun*
Biochemisches Institut, Universität Zürich, Winterthurer Straße 190, CH-8057 Zürich, Switzerland.

Summary

With the discovery of molecular chaperones and the development of heterologous gene expression techniques, protein folding in bacteria has come into focus as a potentially limiting factor in expression and as a topic of interest in its own right. Many proteins of importance in biotechnology contain disulphide bonds, which form in the *Escherichia coli* periplasm, but most work on protein folding in the periplasm of *E. coli* is very recent and is often speculative. This *MicroReview* gives a short overview of the possible fates of a periplasmic protein from the moment it is translocated, as well as of the *E. coli* proteins involved in this process. After an introduction to the specific physiological situation in the periplasm of *E. coli*, we discuss the proteins that might help other proteins to obtain their correctly folded conformation — disulphide isomerase, rotamase, parts of the translocation apparatus and putative periplasmic chaperones — and briefly cover the guided assembly of multi-subunit structures. Finally, our *MicroReview* turns to the fate of misfolded proteins: degradation by periplasmic proteases and aggregation phenomena.

Introduction

The periplasm of *Escherichia coli* is of particular interest in the heterologous expression of recombinant proteins: its non-reducing environment allows disulphide bridges to be formed there. Periplasmic secretion has been used for the functional expression of a wide variety of recombinant proteins from different sources, such as antibody fragments (Plückthun, 1992). Despite the considerable importance of functional expression in the periplasm, periplasmic protein folding has not been studied in any great detail to date. Apart from the catalysts of disulphide bond formation and the suspected catalysts of the prolyl–peptidyl *cis–trans* isomerization no periplasmic proteins involved in protein folding are known. Furthermore, the physiological conditions in the periplasm differ from those in other folding compartments, making folding in the periplasm a topic in its own right. This *MicroReview* follows the pathway of a periplasmic protein from the moment of translocation into the periplasm, and also describes the physiological parameters relevant for folding. The folding process is described, along with its failed outcomes such as aggregation and degradation of the secreted protein. The review concentrates on proteins that are thought to be directly involved in these processes.

Received 23 December, 1993; revised 22 February 1994; accepted 28 February, 1994. *For correspondence. Tel. (1) 257 5570; Fax (1) 257 5712.

Specific properties of the periplasm

The periplasm is the region between the inner and the outer membrane of Gram-negative bacteria. Estimates of its size vary considerably (van Wielink and Duine, 1990), but combining data from electron microscopy, known amounts of proteins present in the periplasm and electron transfer rates, van Wielink and Duine argue that the periplasm constitutes about 30% of the total cell volume of *E. coli,* having a width of about 50 nm.

It has been argued that the periplasm has a gel-like structure. Hobot *et al.* (1984) have, based on electron microscopy data with special sample preparation procedures, put forward a model that the periplasm is essentially filled with a peptidoglycan matrix with large pores. This matrix, it is proposed, is progressively more tightly cross-linked towards the outer membrane and has a high content of bound water. Using photobleaching recovery of several fluorescently labelled proteins, Brass *et al.* (1986) have shown that the lateral diffusion coefficients in the periplasm are smaller by two orders of magnitude than the diffusion coefficients of proteins in the cytoplasm. From these data, Brass *et al.* argue that periplasmic proteins (e.g. the metabolite-binding proteins) exist in a partially ordered fashion with considerably impaired mobility. Therefore, even though data concerning the size of the periplasm in the cited studies are conflicting to some extent, it seems reasonable to assume that a protein in the periplasm is moving much more slowly than in the cytoplasm, as a result of being in permanent contact either with other proteins of the periplasm or with polysaccharides.

Energetics of the periplasm

Another important point still open to discussion is the energetics of the periplasm. From the mere existence of periplasmic phosphatases and nucleotidases it was argued that the existence of ATP in the periplasm is highly improbable (Rosen, 1987). It is also improbable that ATP could be excreted into the periplasm, since the inner membrane is thought to be impermeable to whole nucleotides. In classic uptake experiments, investigating nucleotide triphosphate uptake with CMP, simultaneously ^{14}C- and ^{32}P-labelled, the carbon label, but not the phosphate label (in presence of cold P_i), was taken up and used biosynthetically (Lichtenstein *et al.*, 1960). Thus, there does not seem to be a translocator for whole nucleotides; rather they have to be converted to nucleosides by phosphatases before they are taken up. This implies that ATP transport to the periplasm would require a specific energy-driven translocation apparatus. To our knowledge, there is so far no evidence for the existence of such a translocator.

Since one of the major energy-consuming processes in the periplasm is the active transport of metabolites, the question of energetics has frequently been discussed in this context. The prototypical transport systems, such as the maltose-transport proteins, are dependent on intracellular ATP (reviewed, e.g., in Shuman, 1987; Nikaido and Saier, 1992). One membrane-associated component, thought to be located on the cytoplasmic side of the inner membrane, possesses a nucleotide-binding motif (Higgins *et al.*, 1985). Another set of experiments with the histidine-transport system, believed to be another example of the same family, suggests, however, that the ATP binding, membrane-associated component HisP (a protein without a signal sequence) is accessible to the periplasm (Baichwal *et al.*, 1993). Even so, the ATP-hydrolysis site is believed to be on the cytoplasmic side. It is generally assumed that the energy required by these transport systems for transport is provided by cytoplasmic ATP hydrolysis, indirectly arguing against the presence of a periplasmic source of energy.

Contrary, yet indirect, evidence in favour of a periplasmic high-energy phosphate donor has come from studies of the arginine-transport systems. The periplasmic arginine–ornithine-binding protein was found to occur either in a phosphorylated or an unphosphorylated state *in vivo*, with binding constants for arginine differing between these states by a factor of 50 (Celis, 1984). Similarly, a gene coding for a periplasmic kinase/ATPase which is able to phosphorylate the binding protein *in vivo* has been cloned (Urban and Celis, 1990), while an *E. coli* strain producing a mutant kinase with decreased affinity for ATP has been shown to have impaired arginine transport (Celis, 1990). Nevertheless, even with these data in hand, the authors do not dare to suggest the existence of ATP in the periplasm. In summary, therefore, every energy-consuming step in the periplasm has to cope with the specific restraints of energetics in the periplasm briefly outlined above.

Late steps in protein translocation and early steps in protein folding

All proteins are synthesized in the cytoplasm. For post-translational transport they have to cross the cytoplasm to reach the inner membrane. Cytoplasmic chaperones are known to influence this process and chaperones specialized for transport (such as SecB) are assumed to exist. An effect of the cytoplasmic chaperone GroE has been observed, for example, in the assembly of filamentous phages (Söderlind *et al.*, 1993). Proteins assisting in the transport steps are outside the scope of this *MicroReview*, however, and there is no evidence to suggest, nor is it reasonable to assume, that cytoplasmic chaperones have a direct effect on the periplasmic folding process.

For this *MicroReview*, our consideration of the folding process starts where proteins cross the inner membrane in order to reach the periplasm. The usual way to achieve this is by the use of a signal sequence and the Sec system of proteins. Transport to the periplasm requires that the protein to be transported is not already completely folded in the cytoplasm (Randall *et al.*, 1990; Hardy and Randall, 1993; Derman *et al.*, 1993), and the signal sequence has been shown to slow down folding (Park *et al.*, 1988; Laminet and Plückthun, 1989). Slow-folding variants of a protein apparently can be transported without a signal sequence, albeit less efficiently (Derman *et al.*, 1993, Bowden *et al.*, 1992). There is some discussion about the port of entry to the periplasm, as briefly discussed by Bassford *et al.* (1991): even though homologues to the RNA species and one protein subunit of the eukaryotic signal recognition particle as well as the docking protein have been identified in *E. coli* (Poritz *et al.*, 1988; Römisch *et al.*, 1989; Bernstein *et al.*, 1989), Bassford *et al.* (1991) summarize convincing evidence that the primary pathway of protein transport is the Sec system. However, this problem is far from being solved (Dobberstein, 1994). The energy of the translocation process is provided by the membrane potential and by ATP (Johnson *et al.*, 1992; Jacq and Holland, 1993).

At least one component of the translocation machinery, SecD, appears to be involved in a post-translocational step (Bieker-Brady and Silhavy, 1992). SecD is a transmembrane protein with a large periplasmic domain; Matsuyama *et al.* (1993) have shown that blocking SecD (exposed on the surface of sphaeroplasts) with antibodies resulted in retarded release of maltose-binding protein (MBP).

Furthermore, the mature protein was misfolded, as shown by trypsin sensitivity, and precursor molecules accumulated. SecD was thus proposed to be involved in the release of translocated protein, although whether release precedes folding or vice versa is a crucial question for interpreting its function. Since alkaline phosphatase, which lacks the correctly formed disulphide bonds, cannot be released from the translocation apparatus (Kamitani *et al.*, 1992), it appears that at least part of the folding in the periplasm takes place while the protein is still connected to the translocation machinery. In this case, SecD might function as a chaperone and prevent side-reactions. Since the final folding step of MBP, as judged by trypsin sensitivity, is dependent on the membrane potential, a study of the activity of SecD as a function of the membrane potential might help clarify its exact role. While SecF may also have a substantial periplasmic domain, it is present in smaller amounts and its role in translocation remains to be elucidated (Gardel *et al.*, 1990).

Another protein that was suggested to be involved in a late translocation step across the plasma membrane is SKP (Thome and Müller, 1991). SKP is a highly positively charged 15 kDa protein, which, after cell rupture, has been found in association with ribosomes and DNA (Holck *et al.*, 1987; Thome *et al.*, 1990). Thome and Müller (1991) have shown that SKP, despite forming these associations after cell rupture, is a periplasmic protein, and its only demonstrated activity to date is that of substituting for the SecA function during *in vitro* protein translocation assays. Since SecA is suggested to be an inhibitor of reverse translocation (i.e. the return to the cytoplasm of partially translocated periplasmic proteins) (Schiebel *et al.*, 1991), Thome and Müller (1991) suggest that SKP might act in a similar manner in a late translocation step on the periplasmic side. Its exact function remains unknown, however.

Putative folding catalysts

We shall call proteins that catalyse a chemically defined step in protein folding, 'folding catalysts' as opposed to 'molecular chaperones'. The latter presumably work by controlling protein–protein interactions, thereby preventing side-reactions, and usually without actually catalysing protein folding.

The periplasm of *E. coli* is presumed to be non-reducing, allowing the formation of disulphide bonds. Bardwell *et al.* (1991) identified a protein involved in disulphide bond formation of *E. coli*, which they called DsbA, and the same protein was discovered independently by Kamitani *et al.* (1992). The DsbA protein has a possible high-energy disulphide bond, similar to disulphide isomerase (reviewed by Noiva and Lennarz, 1992; Bulleid, 1993). Its crystal structure has been solved recently (Martin *et al.*, 1993) and the redox potential has been determined as -0.089 V at pH 7.0 (Wunderlich and Glockshuber, 1993). It is thus a significantly stronger oxidant than the cytoplasmic thioredoxin and it more closely resembles the eukaryotic disulphide isomerases. The same conclusion was reached by Zapun *et al.* (1993), who independently found that unfolding DsbA stabilizes the reactive disulphide bond by about 18.9 kJ mol^{-1}. They interpret this result to indicate that the disulphide bond destabilizes the folded form of DsbA because of its high energy content.

Furthermore, Bardwell *et al.* (1993) and, independently, Missiakas *et al.* (1993) cloned the gene of a second protein involved in disulphide bond formation, which they called *dsbB*. DsbB is an integral membrane protein spanning the inner membrane. It seems to be involved in the reoxidation of DsbA and thus may form part of a chain that links an electron transfer step to the formation of disulphide bonds in the periplasm (Bardwell *et al.*, 1993). This raises the interesting possibility that molecular oxygen may be involved in the disulphide bond formation only very indirectly or even not at all. A third gene for disulphide bond formation, *dsbC*, was recently identified by Missiakas, Georgopoulos and Raina (C. Georgopoulos, personal communication). It encodes a 26 kDa periplasmic protein that can functionally substitute for DsbA.

Bardwell *et al.* (1991; 1993), as well as Ito and coworkers (Kamitani *et al.*, 1992; Akiyama *et al.*, 1992; Akiyama and Ito, 1993) have shown that DsbA is required for the correct formation of disulphide bonds of periplasmic *E. coli* proteins *in vivo* and *in vitro*. The same was found for the production of recombinant eukaryotic proteins in the periplasm of *E. coli*, e.g. different serine proteases (Bardwell *et al.*, 1993), antibody fragments (Knappik *et al.*, 1993) and fragments of a T-cell receptor (C. Wülfing and A. Plückthun, submitted). The fraction of the recombinant molecules which becomes correctly folded can be smaller than in the case of natural periplasmic proteins, and is very sequence dependent (A. Knappik and A. Plückthun, in preparation). Overproduction of DsbA did not help to increase the proportion of correctly folded periplasmic antibody fragments (Knappik *et al.*, 1993), indicating that other steps limit their folding in the periplasm (A. Knappik and A. Plückthun, in preparation). With other proteins as substrates, however, overproduction of DsbA was found to increase the yield of a protein with multiple disulphide bonds (C. Wülfing and A. Plückthun, submitted; R. Glockshuber, personal communication). It appears that improved catalysis of disulphide formation is only effective for increasing expression yields under a narrow set of conditions.

The formation of intermolecular disulphide bonds in *E. coli* is a rather complex matter. When the antibody hinge region, fused to β-lactamase (de Sutter *et al.*, 1992) or to

an antibody Fab fragment (Carter *et al.*, 1992), was used as a model, large proportions of neither β-lactamase dimers nor $(Fab)_2$ were observed. Rather, de Sutter *et al.* observed some intramolecular disulphide bond formation and non-specific crosslinking to other periplasmic proteins. It appears that the non-covalent interactions must guide protein domains to the right orientation and specific dimer formation, and that disulphide bonds form as a consequence of non-covalent association.

A comparison of the process of periplasmic folding, including the formation of disulphide bonds, with that which takes place in the endoplasmic reticulum (ER) of eukaryotic cells reveals differences that might help explain why folding of disulphide-containing proteins in the periplasm of *E. coli* can be relatively inefficient. Braakman *et al.* (1992) found that the correct formation of the disulphide bonds of the influenza-virus haemagglutinin (HA0) in the ER requires not only an oxidant, but also ATP. In sodium azide-treated and ATP-depleted cells, HA0 was oxidized, but only formed aggregates. Part of the explanation is probably that the ER contains a chaperone of the hsp70 family, the so-called antibody heavy chain binding protein (BiP), which is probably involved in an ATP-requiring step, as it possesses ATPase activity (Munro and Pelham, 1986). BiP is clearly required for the secretion of antibodies from the cell, and perhaps of many other proteins. It is certainly possible that there are other, undiscovered, chaperones in the ER which also require ATP. The crucial difference between folding in the periplasm as compared with folding in the ER may be the participation of chaperones in energy-requiring steps in the ER.

Another well-studied step along the folding pathway of proteins that has a considerable activation energy and might conceivably need catalysis *in vivo* is the *cis–trans* isomerization of peptide bonds (Fischer *et al.*, 1984). Proteins that catalyse peptidyl–prolyl *cis–trans* isomerization *in vitro* have been identified. Liu and Walsh (1990) cloned the gene for an *E. coli* periplasmic cyclophilin homologue, *rot* (for rotamase), whose product has a second-order rate constant close to the diffusion limit for a tetrapeptide substrate. Furthermore, Compton *et al.* (1992) showed that the *E. coli* periplasmic rotamase is able to catalyse the *in vitro* refolding of thermally denatured type III collagen. These data suggest that an involvement of rotamase in periplasmic protein folding is at least conceivable. Knappik *et al.* (1993) overproduced the periplasmic rotamase in order to test for an increase of the expression yields of functional antibody fragments in the periplasm of *E. coli*. Only a very small effect was seen, and only for one of the fragments tested, suggesting that the isomerization of peptide bonds does not limit antibody folding in the periplasm. It might be noted that the question of the physiological function of the rotamase is clearly still open. Assuming that the rotamases are involved in protein folding, it is a mystery how a presumably crucial function can be as dispensable as it seems, since all knock-out mutations of rotamases described so far are perfectly viable under most conditions tested (see, e.g., Davis *et al.*, 1992; McLaughlin *et al.*, 1992). In contrast, knock-out mutants in the cytoplasmic chaperone genes *groEL* and *groES* are lethal under all conditions tested (Fayet *et al.*, 1989).

General periplasmic chaperones

Throughout this review, we use the term 'general chaperone' to describe chaperones with very broad substrate specificity. Cytoplasmic examples of this type would be GroE and DnaK. To our knowledge, such chaperones have not yet been identified for the periplasm.

One possible candidate for being a general periplasmic chaperone is ClpB (Squires *et al.*, 1991). ClpB is a member of a large family of homologous proteins which exist in all kingdoms, i.e. the Clp family. This family consists of three subgroups, ClpA, B and C, and the properties of the whole family have been reviewed recently (Squires and Squires, 1992). The best-studied member of the ClpB subfamily is the yeast heat-shock protein hsp104 (Parsell *et al.*, 1991). Antibodies which recognize yeast hsp104 cross-react with human hsp110 and the two forms of *E. coli* ClpB, emphasizing the degree of conservation. The members of this subfamily have been suggested to be chaperones. Evidence for this idea comes from studies where the effect of the lack of yeast hsp104 could be partially overcome by overproduction of hsp70 (Parsell *et al.*, 1991).

Like the yeast hsp104 gene, the *E. coli clpB* gene is also subject to heat-shock control. It is transcribed using the heat-shock sigma factor, σ^{32} (Kitagawa *et al.*, 1991). From the *E. coli clpB* gene two proteins are made, a full-length hsp104 homologue and a truncated form, lacking the first 147 amino acids. The full-length protein is a tetrameric ATPase (Woo *et al.*, 1992). Since the *E. coli* member of the ClpA family is a (non-catalytic) subunit of the cytoplasmic Clp-protease (Squires and Squires, 1991), it might be speculated that ClpB fulfils a similar function. However, it was found that ClpB could not substitute for ClpA (Woo *et al.*, 1992) in activating the protease. It therefore seems plausible that ClpB might share the ability to bind to non-native proteins, but, rather than leading them to degradation, it leads them to a folded state analogous to the yeast hsp104.

The first 147 amino acids of the full-length ClpB are homologous to the ClpC leader sequence, which includes a tandem repeat of 32 amino acids each, and has been suggested to have a potential signal-sequence function, although the similarity to 'classical' signal sequences is only remote. Squires and Squires (1991) suggest, as yet

without experimental evidence, that the two *E. coli* isoforms of ClpB that have been found to exist (one with and one without the putative signal sequence) might be located in different cell compartments. A periplasmic localization for full-length ClpB seems possible, but has yet to be proven. In conclusion, whereas the function of ClpB as a general chaperone seems to be a quite plausible hypothesis, a periplasmic localization can only be suggested. With regard to the question of folding in the periplasm, localization of the large isoform of ClpB is an important issue, especially when taking into account the ClpB ATPase activity.

Indirect evidence for the existence of a general periplasmic chaperone has recently been obtained in our laboratory. In order to produce functional fragments of a T-cell receptor (TCR) in the periplasm of *E. coli*, we used modest overproduction of DsbA and the *E. coli* heat-shock proteins at low temperature. The latter was achieved by overexpression of *rpoH*, coding for the heat-shock sigma factor σ^{32} (Grossmann *et al.*, 1984). This increased the folding yield of the TCR fragments in the periplasm by about two orders of magnitude (C. Wülfing and A. Plückthun, submitted). Since the folding yield of a recombinant protein has been increased, it seems reasonable to assume that, rather than a specific sequence, general features of the folding intermediates are recognized by the putative chaperone. Therefore, in our opinion, this effect argues for the existence of a general heat-shock inducible periplasmic chaperone.

A further hint of the existence of a general periplasmic chaperone comes from studies of protein secretion in *Bacillus subtilis*. Both Gram-positive *B. subtilis* and Gram-negative *E. coli* possess a general secretion apparatus which makes use of signal sequences, i.e. the homologous *sec* systems. The SecA, and the SecY proteins of *B. subtilis* have been shown to be able to substitute, at least partially, for their *E. coli* homologues (Nakumara *et al.*, 1990; Overhoff *et al.*, 1991). In addition, in *B. subtilis* a gene for a further protein, PrsA, which has been shown to be involved in secretion, has been cloned (Kontinen and Sarvas, 1988) and shown to have a homologue in *Lactococcus lactis,* PrtM, that is involved in exoproteinase secretion (Haandrikman *et al.,* 1989; Vos *et al.,* 1989). Recently, Kontinen and Sarvas (1993) have shown that PrsA is located on the extracellular surface of the *B. subtilis* cytoplasmic membrane, probably anchored in the membrane only by an *N*-terminal lipid. PrsA is essential for *B. subtilis* viability, it is the most abundant protein on the extracellular side of the cytoplasmic membrane (Kontinen and Sarvas, 1993), and it has been shown to be involved in the correct folding of subtilisin, α-amylase and different subtilisin–alkaline phosphatase hybrids (Jacobs *et al.*, 1993). These data suggest that PrsA may be a general extracellular chaperone in a central step in protein secretion in *B. subtilis*. Taking into account the homology between the *B. subtilis* and the *E. coli* secretion apparatus, it would not be surprising to find a PrsA homologue in *E. coli.*

Specific periplasmic chaperones

Summarizing the above data, we must conclude that the *E. coli* periplasm does not exactly appear to teem with chaperones. Interestingly, several classes of multi-protein assemblies, adhesive pili or fimbriae, have their own specific chaperones. Since this field has recently been reviewed comprehensively (Hultgren *et al.*, 1993), it will be covered here only very briefly. The best-characterized chaperone in pilus assembly is the PapD protein. Its crystal structure, with and without substrate peptides (Kuehn *et al.*, 1993; Holmgren and Bränden, 1989), revealed the presence of a deep substrate-binding cleft. Since no substrates other than the subunits of the P-pilus to be assembled are known to interact with PapD, it is thought to be a specific chaperone. Similar proteins have been described in the biosynthesis of other fimbriae and these constitute a large family of homologous proteins (Bertin *et al.*, 1993).

Further evidence for the need for specific chaperoning — perhaps because of the lack of a suitable general one — comes from a series of studies on the expression of the secreted proteases of Gram-positive bacteria in the periplasm of *E. coli*. Subtilisin (from *B. subtilis*) and α-lytic protease (from *Lysobacter enzymogenes*) are made as pre-proenzymes, and the pro-piece, even when expressed *in trans,* acts in improving the folding efficiency of the mature enzyme (reviewed in Inouye, 1991). Since these proteins are secreted in Gram-positive bacteria, they cannot rely on secreted chaperones, as these would diffuse away far too rapidly. When expressing them in the periplasm of *E. coli*, the failure to observe folding of the mature part in the absence of the pro-piece argues that there is no activity present in the periplasm of *E. coli,* at least under the conditions tested, which can efficiently substitute for this function.

The fate of misfolded proteins

The principal physiological situation in which misfolded proteins accumulate to a large extent, even in the absence of overproduction, is during heat shock (Kucharczyk *et al.*, 1991). However, *E. coli* can remove these aggregates within only 10 min after the return to 37°C. This efficient clean-up is probably accomplished by heat-shock proteins, since it is dependent on the expression of the gene for the heat shock regulating sigma-factor *rpoH.* Several heat-shock proteins have been shown to co-localize with the aggregates (Kucharczyk *et al.*, 1991), of which the

only periplasmic protein identified was the protease DegP or HtrA. Its gene has been cloned independently by Strauch *et al.* (1989) and Lipinska *et al.* (1990). DegP is an ATP-independent periplasmic protease which is essential for survival at elevated temperatures (Lipinska *et al.*, 1990), yet its transcription is controlled not by the heat-shock-specific sigma factor σ^{32} (encoded by *rpoH*), but by the sigma factor σ^{E}, also called σ^{24} (Erickson and Gross, 1989). The only other gene currently known to be transcribed by σ^{E} is *rpoH* (Wang and Kaguni, 1989).

Examples of recombinant proteins in whose degradation DegP is involved include a slow-folding mutant of LamB (Misra *et al.*, 1991), truncated derivatives of diphtheria toxin (Zandovsky *et al.*, 1992), and *E. coli* proteins that are defective in some processing step, such as the acetylated precursor of the colicin A lysis protein which accumulates in the presence of globomycin (Cavard *et al.*, 1989). It is possible that use of *degP*$^{-}$ strains might help to increase the yields of inefficiently folding proteins, although it would remain to be shown whether any additional amount of protein obtained would fold correctly. Since DegP seems to be primarily involved in the degradation of misfolded, aggregated protein, the use of *degP*$^{-}$ strains cannot solve the folding problem, which is probably the underlying reason for the protease susceptibility. Georgopoulos and co-workers adopted the strategy of looking for extragenic suppressors of *degP*. They discovered a cytoplasmic protein, probably involved in transport and encoded by a gene called *sohA* or *prlF* (Baird and Geogopoulos, 1990; Kiino *et al.*, 1990) and another periplasmic protease, encoded by *sohB* (Baird *et al.*, 1991). Their approach promises to identify a general periplasmic chaperone, if it exists, since efficient folding would prevent the accumulation of misfolded protein and thus the need for degradation of the precipitated protein.

Degradation of misfolded proteins in the periplasm is almost certainly mediated by a variety of proteases, many of which probably remain to be discovered. We ignore here the specialized proteases involved in cleaving signal sequences of 'normal' proteins (*lep*) and lipoproteins (*lsp*) as well as protease IV (*sppA*), which is presumed to degrade the signal peptide, even though it cannot be rigorously excluded that some of these carry out side-reactions on misfolded proteins in the periplasm, at least after cell rupture (Pacaud, 1982; Palmer and St John, 1987). No specialized target has been suggested for protease I (*apeA*) (Ichihara *et al.*, 1993) or protease III (*ptr*) (Finch *et al.*, 1986; Claverie-Martin *et al.*, 1987), both of which may be involved in degrading misfolded proteins, and very little is known about protease V (Pacaud, 1982), a membrane-bound enzyme. Whether their physiological role is to degrade abnormal proteins or external peptides for providing nutrients remains to be elucidated. Another protease of unknown function is OmpT (Sugimura and Nishihara, 1988), which is known to cleave after basic residues. Strains that are deficient in OmpT and DegP, or triple mutants deficient in OmpT, DegP and Ptr were shown to yield increased amounts of recombinant material, but only in minimal medium (Baneyx and Georgiou, 1990; 1991; for a short discussion of the usefulness of protease-deficient strains see the DegP paragraph above). OmpT is known to be located in the outer membrane, although it remains unclear as to which face of the membrane the active site is exposed. More recently, a new protease, OmpP, has been discovered in *E. coli* K-12 (Kaufmann *et al.*, 1994), which is homologous to OmpT, and the bulk of the protein appears to be exposed to the outside of the cell. This would suggest that OmpP and, by analogy, OmpT start their degradative action only after cell rupture. In conclusion, much work remains to be done before we can understand and control the degradation of misfolded proteins in the periplasm.

Another possible fate of misfolded proteins, and perhaps also of overproduced, correctly folded proteins, is aggregation into periplasmic inclusion bodies. It is technically very difficult to rigorously differentiate between 'membrane-associated' and 'periplasmically precipitated' forms by cell-fractionation experiments. While a differentiation, sucrose density-gradient centrifugation (Kucharczyk *et al.*, 1991) is available in principle, it has been applied only rarely to recombinant proteins. The underlying causes and the implications for improvement would be rather different for membrane-bound or precipitated recombinant proteins. In one case the problem would lie with the secretion process itself, and in the other in the folding reaction occurring after secretion. For some proteins, however, electron microscopy could be used to visualize periplasmic inclusion bodies directly, which is only consistent with the aggregation of intermediates of periplasmic folding. Using *E. coli*'s own β-lactamase as an example, Bowden *et al.* (1991) showed that cytoplasmic and periplasmic inclusion bodies of the same protein are morphologically different, can be solubilized under different conditions and appear to have different contaminants (Valax and Georgiou, 1993).

Summary

The mechanism of protein folding in the periplasm is essentially unsolved. Nevertheless, there are several potential routes to an understanding and further investigation of this problem. On the basis of the specific physiological environment in the periplasm, namely the reduced mobility of proteins and probably the different energetical situation, folding in the periplasm can be assumed to have its own specific properties. Folding is probably associated with translocation into the periplasm and proteins such as SecD or SKP, which are conceivably

involved in late translocation steps, might therefore have important functions in folding. Another route to defining proteins involved in periplasmic folding is the search for homologues of well-established folding modulators and chaperones. This approach has proved fruitful for the identification and characterization of the periplasmic rotamase and ClpB. Although proteins of the PapD family are undoubtedly chaperones, it is questionable whether, because of their very specific function, their further characterization would help elucidate general periplasmic folding processes. Rather, searching for heat-shock-inducible periplasmic proteins might be a promising way to identify a general periplasmic chaperone, as might be the search for extragenic suppressors of mutations affecting periplasmic heat-shock proteases. The further investigation of proteolysis and aggregation in the periplasm will probably not contribute significantly to the understanding of folding, since these processes deal with already misfolded proteins. However, because of their importance in the production of recombinant proteins, they certainly deserve more attention. It remains to be shown whether the mere overproduction of suspected chaperones will improve folding or whether serious engineering of the protein to be folded in the periplasm may be required to funnel the protein to the often so elusive native state.

References

Akiyama, Y., and Ito, K. (1993) *J Biol Chem* **168:** 8146–8150.

Akiyama, Y., Kamitani, S., Kusukawa, N., and Ito, K. (1992) *J Biol Chem* **267:** 22440–22445.

Baichwal, V., Liu, D., and Ferro-Luzzi Ames, G. (1993) *Proc Natl Acad Sci USA* **90:** 620–624.

Baird, L., and Georgopoulos, C. (1990) *J Bacteriol* **172:** 1587–1594.

Baird, L., Lipinska, B., Raina, S., and Georgopoulos, C. (1991) *J Bacteriol* **173:** 5763–5770.

Baneyx, F., and Georgiou, G. (1990) *J Bacteriol* **172:** 491–494.

Baneyx, F., and Georgiou, G. (1991) *J Bacteriol* **173:** 2696–2703.

Bardwell, J.C.A., McGovern, K., and Beckwith, J. (1991) *Cell* **67:** 581–589.

Bardwell, J.C.A., Lee, J.-O., Jander, G., Martin, N., Belin, D., and Beckwith, J. (1993) *Proc Natl Acad Sci USA* **90:** 1038–1042.

Bassford, P., Beckwith, J., Ito, K., Kumamoto, C., Mizushima, S., Oliver, D., Randall, L., Silhavy, T., Tai, P.C., and Wickner, B. (1991) *Cell* **65:** 367–368.

Bernstein, H.D., Poritz, M.A., Strub, K., Hoben, P.J., Brenner, S., and Walter, P. (1989) *Nature* **340:** 482–486.

Bertin, Y., Girardeau, J.-P., Vartanian, M.D., and Martin, C. (1993) *FEMS Microbiol Lett* **108:** 59–68.

Bieker-Brady, K., and Silhavy, T.J. (1992) **11:** 3165–3174.

Bowden, G.A., Paredes, A.M., and Georgiou, G. (1991) *Biotechnology* **9:** 725–730.

Bowden, G.A., Baneyx, F., and Georgiou, G. (1992) *J Bacteriol* **174:** 3407–3410.

Braakman, I., Helenius, J., and Helenius, A. (1992) *Nature* **356:** 260–262.

Brass, J.M., Higgins, C.F., Foley, M., Rugman, P.A., Birmingham, J., and Garland, P.B. (1986) *J Bacteriol* **165:** 787–794.

Bulleid, N.J. (1993) *Adv Prot Chem* **44:** 125–150.

Carter, P., Kelley, R.F., Rodrigues, M.L., Snedecor, B., Covarrubias, M., Velligan, M.D., Wong, W.L.T., Rowland, A.M., Kotts, C.E., Carver, M.E., Maria, Y., Bourell, J.H., Shepard, H.M., and Henner, D. (1992) *Biotechnology* **10:** 163–167.

Cavard, D., Lazdunski, C., and Howard, S.P. (1989) *J Bacteriol* **171:** 6316–6322.

Celis, R.T.F. (1984) *Eur J Biochem* **145:** 403–411.

Celis, R.T.F. (1990) *J Biol Chem* **265:** 1787–1793.

Claverie-Martin, F., Diaz-Torres, M.R., and Kushner, S.R. (1987) *Gene* **54:** 185–195.

Compton, L.A., Davis, J.M., MacDonald, J.R., and Bächinger, H.P. (1992) *Eur J Biochem* **206:** 927–934.

Davis, E.S., Becker, A., Heitman, J., Hall, M.N., and Brennan, M.B. (1992) *Proc Natl Acad Sci USA* **89:** 11169–11173.

Derman, A.I., Puziss, J.W., Bassford, P.J., and Beckwith, J. (1993) *EMBO J* **12:** 879–888.

Dobberstein, B. (1994) *Nature* **367:** 599–600.

Erikson, J.W., and Gross, C.A. (1989) *Genes Dev* **3:** 1462–1471.

Fayet, O., Ziegelhoffer, T., and Georgopoulos, C. (1989) *J Bacteriol* **171:** 1379–1385.

Finch, P.W., Wilson, R.E., Brown, K., Hickson, I.D., and Emmerson, P.T. (1986) *Nucl Acids Res* **14:** 7695–7703.

Fischer, G., Bang, H., and Mech, C. (1984) *Biomed Biochim Acta* **10:** 1101–1111.

Gardel, C., Johnson, K., Jacq, A., and Beckwith, J. (1990) *EMBO J* **9:** 3209–3216.

Grossman, A.D., Erickson, J.W., and Gross, C.A. (1984) *Cell* **38:** 383–390.

Haandrikman, A.J., Kok, J., Laan, H., Soemitro, S., Ledeboer, A.M., Konigs, W.N., and Venema, G. (1989) *J Bacteriol* **171:** 2789–2794.

Hardy, S.J., and Randall, L.L. (1993) *Phil Trans Roy Soc Lond B* **339:** 343–354.

Higgins, C.F., Hiles, I.D., Whalley, K., and Jamieson, D.J. (1985) *EMBO J* **4:** 1033–1040.

Hobot, J.A., Carleman, E., Villiger, W., and Kellenberger, E. (1984) *J Bacteriol* **160:** 143–152.

Holck, A., Lossius, I., Aasland, R., Haarr, L., and Kleppe, K. (1987) *Biochim Biophys Acta* **908:** 188–199.

Holmgren, A., and Bränden, C.-I. (1989) *Nature* **342:** 248–251.

Hultgren, S.J., Abraham, S., Caparon, M., Falk, P., St Geme, III J.W., and Normark, S. (1993) *Cell* **73:** 887–901.

Ichihara, S., Matsubara, Y., Kato, C., Akasaka, K., and Mizushima, S. (1993) *J Bacteriol* **175:** 1032–1037.

Inouye, M. (1991) *Enzyme* **45:** 314–321.

Jacobs, M., Andersen, J.B., Kontinen, V.P., and Sarvas, M. (1993) *Mol Microbiol* **8:** 957–966.

Jacq, A., and Holland, B. (1993) *Curr Op Struct Biol* **3:** 541–546

Johnson, K., Murphy, C.K., and Beckwith, J. (1992) *Curr Op Biotechnol* **3:** 481–485.

Kamitani, S., Akiyama, Y., and Ito, K. (1992) *EMBO J* **11:** 57–62.

Kaufmann, A., Stierhof, Y.D., and Henning, V. (1994) *J Bacteriol* **176:** 359–367.

Kiino, D.R., Gregory, J.P., and Silhavy, T.J. (1990) *J Bacteriol* **172:** 185–192.

Kitagawa, M., Wada, C., Yoshioka, S., and Yura, T. (1991) *J Bacteriol* **173:** 4247–4253.

Knappik, A., Krebber, C., and Plückthun, A. (1993) *Biotechnology* **11:** 77–83.

Kontinen, V.P., and Sarvas, M. (1988) *J Gen Microbiol* **134:** 2333–2344.

Kontinen, V.P., and Sarvas, M. (1993) *Mol Microbiol* **8:** 727–737.

Kucharczyk, K., Laskowska, E., and Taylor, A. (1991) *Mol Microbiol* **5:** 2935–2945.

Kuehn, M.J., Ogg, D.J., Kihlberg, J., Slonim, L.N., Flemmer, K., Bergfors, T., and Hultgren, S.J. (1993) *Science* **262:** 1234–1241.

Laminet, A.A., and Plückthun, A. (1989) *EMBO J* **8:** 1469–1477.

Lichtenstein, J., Barner, H.D., and Cohen, S.S. (1960) *J Biol Chem* **235:** 457–465.

Lipinska, B., Zylicz, M., and Georgopoulos, C. (1990) *J Bacteriol* **172:** 1791–1797.

Liu, J., and Walsh, C.T. (1990) *Proc Natl Acad Sci USA* **87:** 4028–4032.

McLaughlin, M.M., Bergsma, D.J., Levy, M.A., and Livi, G.P. (1992) *Gene* **111:** 85–92.

Martin, J.L., Bardwell, J.C.A., and Kuriyan, J. (1993) *Nature* **365:** 464–468.

Matsuyama, S.-I., Fujita, Y., and Mizushima, S. (1993) *EMBO J* **12:** 265–270.

Misra, R., Peterson, A., Ferenci, T., and Silhavy, T.J. (1991) *J Biol Chem* **266:** 13592–13597.

Missiakas, D., Georgopoulos, C., and Raina, S. (1993) *Proc Natl Acad Sci USA* **90:** 7084–7088.

Munro, S., and Pelham, H.R.B. (1986) *Cell* **46:** 291–300.

Nakamura, K., Takamatsu, H., Akiyama, Y., and Ito, K. (1990) *FEBS Lett* **273:** 75–78.

Nikaido, H., and Saier, M.H. (1992) *Science* **258:** 936–942.

Noiva, R., and Lennarz, W.J. (1992) *J Biol Chem* **267:** 3553–3556.

Overhoff, B., Klein, M., Spies, M., and Freundl, R. (1991) *Mol Gen Genet* **228:** 417–423.

Pacaud, M. (1982) *J Biol Chem* **257:** 4333–4338.

Palmer, S.M., and St. John, C. (1987) *J Bacteriol* **169:** 1474–1479.

Park, S., Liu, G., Topping, T.B., Cover, W.H., and Randall, L.L. (1988) *Science* **239:** 1033–1035.

Parsell, P.A., Sanchez, Y., Stitzel, J.D., and Lindquist, S. (1991) *Nature* **353:** 270–273.

Plückthun, A. (1992) *Immunol Rev* **130:** 151–188.

Poritz, M.A., Strub, K., and Walter, P. (1988) *Cell* **55:** 4–6.

Randall, L.L., Topping, T.B., and Hardy, S.J.S. (1990) *Science* **248:** 860–863.

Römisch, K., Webb, J., Herz, J., Prehn, S., Frank, R., Vingron, M., and Dobberstein, B. (1989) *Nature* **340:** 478–482.

Rosen, B.P. (1987) In Escherichia coli *and* Salmonella typhimurium: *Cellular and Molecular Biology.* Neidhardt, F.C. (ed.). Washington, D.C.: American Society for Microbiology, pp. 760–767.

Schiebel, E., Driessen, A.J.M., Hartl, F.-U., and Wickner, W. (1991) *Cell* **64:** 927–939.

Shuman, H.A. (1987) *Annu Rev Genet* **21:** 155–177.

Söderlind, E., Simonsson Lagerkvist, A.C., Dueñas, M., Malmborg, A.-C., Ayala, M., Danielsson, L., and Borrebaeck, C.A.K. (1993) *Biotechnology* **11:** 503–507.

Squires, C., and Squires, C.L. (1992) *J Bacteriol* **174:** 1081–1085.

Squires, C.L., Pedersen, S., Ross, B.M., and Squires, C. (1991) *J Bacteriol* **173:** 4254–4262.

Strauch, K.L., Johnson, K., and Beckwith, J. (1989) *J Bacteriol* **171:** 2689–2696.

Sugimura, K., and Nishihara, T. (1988) *J Bacteriol* **170:** 5625–5632.

de Sutter, K., Remaut, E., and Fiers, W. (1992) *Mol Microbiol* **6:** 2201–2208.

Thome, B.M., and Müller, M. (1991) *Mol Microbiol* **5:** 2815–2821.

Thome, B.M., Hoffschulte, H.K., Schiltz, E., and Müller, M. (1990) *FEBS Lett* **269:** 113–116.

Urban, C., and Celis, R.T.F. (1990) *J Biol Chem* **265:** 1783–1786.

Valax, P., and Georgiou, G. (1993) *Biotechnol Prog* **9:** 539–547.

Vos, P., van Asseidonk, M., Van Jeveren, F., Slezen, R., Simons, G., and de Vos, W.M. (1989) *J Bacteriol* **171:** 2795–2802.

Wang, Q., and Kaguni, J.M. (1989) *J Bacteriol* **171:** 4248–4253.

van Wielink, J.E., and Duine, J.A. (1990) *Trends Biochem* **15:** 136–137.

Woo, K.M., Kim, K.I., Goldberg, A.L., Ha, D.B., and Chung, C.H. (1992) *J Biol Chem* **267:** 20429–20434.

Wunderlich, M., and Glockshuber, R. (1993) *Prot Sci* **2:** 717–726.

Zandovsky, A.G., Kulaeva, O.I., and Yankovsky, N.K. (1992) *Gene* **116:** 81–86.

Zapun, A., Bardwell, J.C.A., and Creighton, T.E. (1993) *Biochemistry* **32:** 5083–5092.