33 Halling, P.J. *et al.* (1994) Relationsship between water activity and catalytic activity of lipases in organic media. *Eur. J. Biochem.* 222, 461–466
34 Schmid, U. *et al.* (1999) Highly selective synthesis of 1,3-oleoyl-2-palmitoylglycerol by lipase catalysis. *Biotechnol. Bioeng.* 64, 678–684
35 Adlercreutz, D. *et al.* (2002) Synthesis of phosphatidylcholine with defined fatty acid in the sn-1 position by lipase-catalyzed esterification and transesterification reaction. *Biotechnol. Bioeng.* 78, 403–411
36 Kazlauskas, R.J. (2000) Molecular modeling and biocatalysis: Explanations, predictions, limitations, and opportunities. *Curr. Opin. Chem. Biol.* 4, 81–88
37 Harris, J.L. and Craik, C.S. (1998) Engineering enzyme specificity. *Curr. Opin. Chem. Biol.* 2, 127–132
38 Hæffner, F. and Norin, T. (1999) Molecular modeling of lipase-catalyzed reactions. Prediction of enantioselectivities. *Chem. Pharm. Bull.* 47, 591–600
39 Cedrone, G. *et al.* (2000) Tailoring new enzyme functions by rational redesign. *Curr. Opin. Struct. Biol.* 10, 405–410
40 Rotticci, D. *et al.* (2001) Improved enantioselectivity of a lipase by rational protein engineering. *Chembiochem.* 2, 766–770
41 Magnusson, A. *et al.* (2001) Creation of an enantioselective hydrolase by engineered substrate-assisted catalysis. *J. Am. Chem. Soc.* 123, 4354–4355
42 Schulz, T. *et al.* (2000) Stereoselectivity of *Pseudomonas cepacia* lipase toward secondary alcohols: A quantitative model. *Protein Sci.* 9, 1053–1062
43 Hwang, B-Y. *et al.* (2000) Computer-aided molecular modeling of the enantioselectivity of *Pseudomonas cepacia* lipase toward γ- and δ-lactones. *J. Mol. Catal. B Enzym.* 10, 223–231
44 Henke, E. *et al.* Activity of lipases and esterases towards tertiary alcohols: New insights into structure–function relationships. *Angew. Chem. Int. Ed. Engl.* (in press)
45 Cadwell, R.C. and Joyce, G.F. (1992) Randomization of genes by PCR mutagenesis. *PCR Methods Appl.* 2, 28–33
46 Stemmer, W.P.C. (1994) Rapid evolution of a protein *in vitro* by DNA shuffling. *Nature* 370, 389–391
47 Bornscheuer, U.T. and Pohl, M. (2001) Improved biocatalysts by directed evolution and rational protein design. *Curr. Opin. Chem. Biol.* 5, 137–142
48 Schmidt-Dannert, C. (2001) Directed evolution of proteins, metabolic pathways, and viruses. *Biochemistry* 40, 13125–13136
49 Wiseman, A. *et al.* (2001) Novel biocatalytic enzymes by directed evolution. *Trends Biotechnol.* 19, 382
50 Brakmann, S. and Johnsson, K. eds (2002) *Directed molecular evolution of proteins*, Wiley-VCH
51 Wahler, D. and Reymond, J-L. (2002) The adrenaline test for enzymes. *Angew. Chem. Int. Ed. Engl.* 41, 1229–1232
52 Wahler, D. *et al.* (2001) Enzyme fingerprints by fluorogenic and chromogenic substrate arrays. *Angew. Chem. Int. Ed. Engl.* 40, 4457–4460
53 Badalassi, F. *et al.* (2000) A versatile periodate coupled fluorogenic assay for hydrolytic enzymes. *Angew. Chem. Int. Ed. Engl.* 39, 4067–4070
54 Baumann, M. *et al.* (2001) A high-throughput-screening method for the identification of active and enantioselective hydrolases. *Angew. Chem. Int. Ed. Engl.* 40, 4201–4204
55 Reetz, M.T. *et al.* (1999) A method for high throughput screening of enantioselective catalysts. *Angew. Chem. Int. Ed. Engl.* 38, 1758–1761
56 Bornscheuer, U.T. *et al.* (1998) Directed evolution of an esterase for the stereoselective resolution of a key intermediate in the synthesis of epothilones. *Biotechnol. Bioeng.* 58, 554–559
57 Henke, E. and Bornscheuer, U.T. (1999) Directed evolution of an esterase from *Pseudomonas fluorescens*. Random mutagenesis by error-prone PCR or a mutator strain and identification of mutants showing enhanced enantioselectivity by resorufin-based fluorescence assay. *Biol. Chem.* 380, 1029–1033
58 Liebeton, K. *et al.* (2000) Directed evolution of an enantioselective lipase. *Chem. Biol.* 7, 709–718
59 Reetz, M.T. *et al.* (1997) Creation of enantioselective biocatalysts for organic chemistry by *in vitro* evolution. *Angew. Chem. Int. Ed. Engl.* 36, 2830–2832
60 Reetz, M.T. *et al.* (2001) Directed evolution of an enantioselective enzyme through combinatorial multiple-cassette mutagenesis. *Angew. Chem. Int. Ed. Engl.* 40, 3589–3591
61 Kauffmann, I. and Schmidt-Dannert, C. (2001) Conversion of *Bacillus thermocatenulatus* lipase into an efficient phospholipase with increased activity towards long-chain fatty acyl substrates by directed evolution and rational design. *Protein Eng.* 14, 919–928
62 van Kampen, M.D. and Egmond, M.R. (2000) Directed evolution: From a staphylococcal lipase to a phospholipase. *Eur. J. Lipid Sci. Technol.* 102, 717–726

Preparative protein refolding

Anton P.J. Middelberg

The rapid provision of purified native protein underpins both structural biology and the development of new biopharmaceuticals. The dominance of *Escherichia coli* as a cellular biofactory depends on technology for solubilizing and refolding proteins that are expressed as insoluble inclusion bodies. Such technology must be scale invariant, easily automated, generic for a broad range of similar proteins and economical. Refolding methods relying on denaturant dilution and column-based approaches meet these criteria. Recent developments, particularly in column-based methods, promise to extend the range of proteins that can be refolded successfully. Developments in preparing denatured purified protein and in the analysis of protein refolding products promise to remove bottlenecks in the overall process. Combined, these developments promise to facilitate the rapid and automated determination of appropriate refolding conditions and to simplify scale-up.

Published online: 19 August 2002

Structural biology promises to add significant value to information obtained from the human genome project. In the race to define protein structure, and hence binding sites, the ability to rapidly convert gene sequences into soluble and crystallizable protein is a source of significant competitive advantage in both academic and industrial laboratories. *Escherichia coli* retains its dominant position as the first choice of host for reasons of speed and simplicity, although a key disadvantage is the formation of insoluble inclusion bodies for some gene sequences.

Inclusion bodies are dense aggregates of misfolded polypeptide. They are formed intracellularly because of the aggregating characteristics of the protein or the inability of the cellular processes to ensure that the expressed polypeptide is soluble and folded correctly. Formation of inclusion bodies might be considered as a dynamic equilibrium between the addition and removal of partially folded proteins from the aggregate (Fig. 1) [1], with the dominant driving force often towards the insoluble state. The likelihood of the formation of inclusion bodies will thus be determined by the interplay between the molecular processing systems of the cells and the thermodynamic drivers that cause the self-association of polypeptides. Inclusion bodies are usually located in the bacterial cytoplasm, although

 0167-7799/02/$ – see front matter PII: S0167-7799(02)02047-4

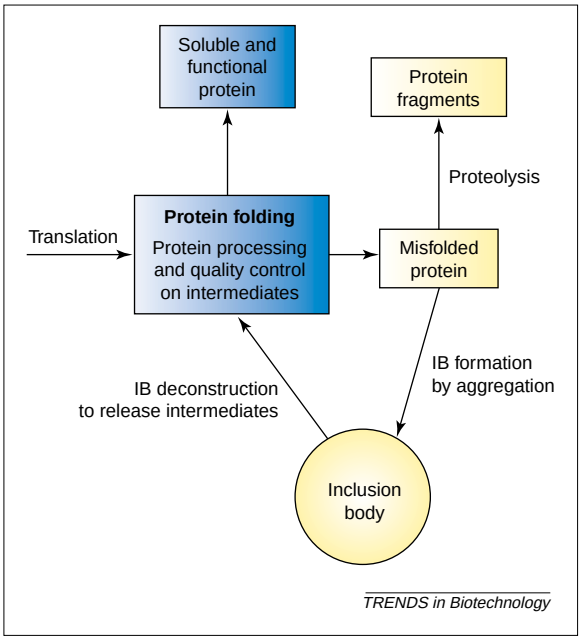

Fig. 1. Inclusion body (IB) formation as a dynamic process involving the addition and removal of incorrectly or incompletely folded proteins (adapted from [1]).

secreted proteins can form inclusion bodies in the periplasmic space. Some proteins incorporated into inclusion bodies might possess elements of native structure but this is usually considered to be minimal. The protein must therefore be released from the inclusion body and 'refolded' or 'renatured' to give its native 3D structure. There is no universal method for refolding proteins and the usual course when confronted with an inclusion body is to search for an effective refolding protocol or to examine an alternative method of expression. If an effective refolding protocol can be defined, the aim is to scale it rapidly to provide preparative quantities (10–100 mg) of protein.

Protein refolding is also important in the commercial manufacture of biopharmaceuticals. The commercial driver for new products is speed – the advantage of being first to market means that moderate refolding yields can be tolerated. Considerations are very similar to those in structural biology and preparative protein refolding will dominate during the early stages of new biopharmaceutical development. As patent protection expires, the driver will shift from speed to cost and the aim will be to optimize refolding yields, maximize volumetric productivity and minimize cost.

With speed being a driver for both structural biology and biopharmaceutical development, only a limited set of refolding strategies can be tested for any given protein. Technology for protein refolding should therefore be:

- scale invariant – to ensure that results from screening can be translated to preparative systems and manufacture without significant changes in the technology used;
- easily automated – to address the issue of speed and to enable high-throughput processing of samples;
- generic for a broad range of similar proteins – so that the technology does not have to be reinvented for each new gene sequence;
- economical – to ensure that resources are not being wasted if a large number of gene sequences or protein variants is being examined, and possibly to ensure streamlined transfer to manufacture.

This article focuses on those techniques that are most likely to meet the above criteria. First, we discuss inclusion body preparation and solubilization because this is often a bottleneck in the overall preparation of refolded protein. Then we focus on refolding techniques that seem to meet the above criteria – dilution refolding and column based refolding. Finally, we review developments in assaying for renaturation.

Preparing purified and denatured protein

Methods for preparing denatured protein involve isolation of the inclusion bodies, with some removal of contaminants, followed by solubilization using concentrated chemical denaturant (typically urea or guanidinium chloride, which are also known as chaotropes because of their ability to disrupt the structure of water) [2].

Inclusion bodies are usually released mechanically or chemically from the cell [3], followed by centrifugation or filtration. The success of this separation can be variable and depends on differences in the underlying composition, structure and polypeptide conformation of the inclusion bodies. For example, β-lactamase inclusion bodies have been shown to contain 35–95% intact product, 5–50% contaminating polypeptides, 0.5–13% phospholipids and traces of nucleic acids, depending on the expression system and growth conditions [4].

There is increasing evidence that contaminants present in preparations of inclusion bodies can significantly reduce refolding yield. In a study using lysozyme from hen egg-white, contaminants were selectively added to refolding mixtures and the effects on the rates of folding and aggregation were monitored [5]. Increases in aggregation, and hence decreases in refolding yield, were observed when plasmid DNA, lipopolysaccharide or proteins that aggregate on folding were added to the renaturation mixture. In a study on the refolding of recombinant human macrophage colony stimulating factor, renaturation yield was increased substantially when reduced and denatured protein was purified by reversed-phase chromatography before refolding [6]. Additionally, some proteins will be sensitive to proteases associated with the cell outer membrane. Babbitt *et al.* [7] examined the recovery of creatine kinase from *E. coli* and obtained a 100-fold increase in yield by detergent washing to remove components of the cell wall from the inclusion bodies. A similar beneficial effect of improved debris removal has been observed for the pilot-scale centrifugation of recombinant insulin-like growth factor II inclusion bodies [8]. In a generic sense, there is clear benefit in removing contaminants before preparative protein refolding.

Contaminant removal during centrifugation can be maximized by considering the physical processes

Anton P.J. Middelberg
Dept of Chemical Engineering, University of Cambridge, Pembroke Street, Cambridge, UK CB2 3RA.
e-mail: antonm@cheng.cam.ac.uk

involved. In density gradient separation, inclusion bodies of *in situ* density 1.3–1.4 g ml^{-1} can be separated from lighter outer membrane vesicles (1.22 g ml^{-1}) and denser ribosomes (1.5 g ml^{-1}) [9]. However, density gradient separation does not satisfy the aforementioned criteria, and is rarely used for preparative work. Consequently, the most common method of inclusion body preparation involves mechanical or chemical cell disruption followed by differential centrifugation to separate the dense inclusion bodies from the lighter cell-membrane components and soluble contaminants. Figure 2a shows the size distributions of a suspension of *E. coli* cell debris and protein inclusion bodies following mechanical breakage of the cells [10]. A significant shift in debris size is observed with repeated mechanical disruption but debris distributions overlap with the inclusion bodies. Fractionation cannot be achieved on the basis of size. Figure 2b shows the size distributions plotted against settling velocity – the denser inclusion bodies settle faster. The importance of repeated cell disruption on the efficiency of centrifugal separation is apparent because particles with an identical settling velocity will not be separable by centrifugation. The benefits of repeated mechanical disruption are clear. In general, the separation achieved by centrifugation will be imperfect and will vary with culture conditions (i.e. inclusion body properties), the efficiency of mechanical cell disruption (i.e. cell debris properties) and centrifuge geometry. The optimum can be calculated for any system but this is impractical unless optimization is required for subsequent scale-up. The normal approach is therefore to use harsh centrifugation conditions that collect most of the inclusion bodies and a significant amount of the particulate contaminant. Repeated washing with denaturant (1–4 M urea), sucrose or detergents is then performed to reduce contaminant levels, but this is at the expense of increased cost and complexity [11]. The success of this washing strategy can also be system specific, depending on the ease of inclusion body resuspension following centrifugation. However, a robust laboratory method is available and involves a combination of enzymatic and mechanical disruption on harvested cells, enzymatic digestion of DNA, treatment with Triton X-100, centrifugation to harvest the inclusion bodies and finally resuspension in Tris buffer followed by another centrifugation step [12] (Fig. 3a). This procedure is often simplified during scale-up, resulting in reduced inclusion body purity.

Following preparation of an acceptably pure inclusion body paste, inclusion body solubilization is usually conducted in concentrated denaturant (urea or guanidinium chloride, GdmCl) [11]. Reducing agent (e.g. 1–500 mM dithiothreitol, DTT) is added to reduce any misformed disulphide bonds [11]. An interesting alternative is the use of non-detergent sulfobetaines, which both solubilize and stabilize the denatured protein and which might interfere with the formation of aggregation-prone intermediates during refolding [13]. Detergent (cetyltrimethyl ammonium chloride, CTAC) has also been used to solubilize and denature inclusion

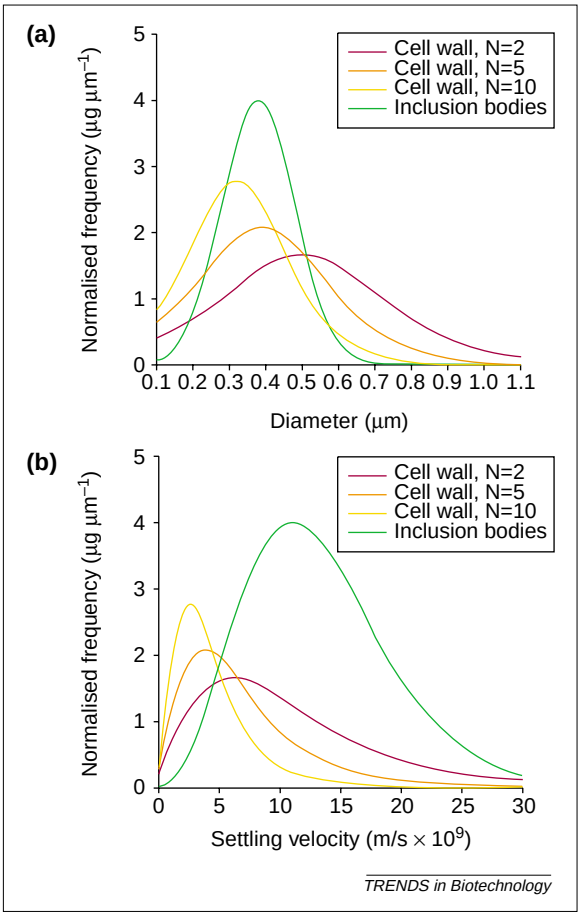

Fig. 2. (a) Size distributions of cell-wall particulate material (cell debris) and of recombinant inclusion bodies in the same suspension. N is the number of times the material was passed through a laboratory homogenizer. Frequency distributions for cell debris were derived from the cumulative size distributions reported previously (see Table 2, fermentation C1 in [10]). The inclusion bodies were assumed to have a normal size distribution of mean 0.38 μm and standard deviation 0.1 μm (see Fig. 7 in [10]). (b) The frequency data from (a) plotted against settling velocity. Velocity was calculated using the Stokes equation with a viscosity of 1.85×10^{-3} kg m^{-1}s^{-1}, cell debris density of 1085 kg m^{-3}, and inclusion body density of 1260 kg m^{-3} [10]. Note that [10] determined settling velocity distributions directly, and then transformed them to size distributions using the Stokes equations and the stated values of viscosity and density.

bodies of recombinant porcine growth hormone [14]. Protein solubilized using CTAC was shown to have some secondary structure (10–15% α-helix and 30–40% β structure), whereas protein solubilized in 6 M GdmCl or 7.5 M urea showed no α-helix content. This increase in starting secondary structure following solubilization translated into a higher refolding yield (50% versus 20%). The generic applicability of these methods still needs to be demonstrated, thus the current method of first choice remains complete solubilization and reduction.

It is apparent from the preceding discussion that the most common method for inclusion body preparation, namely cell disruption followed by centrifugation, meets none of the criteria specified earlier. This has led to a search for alternative methods of preparing solubilized and denatured protein ready for refolding. Direct chemical extraction methods employing minimal centrifugation washes offer the

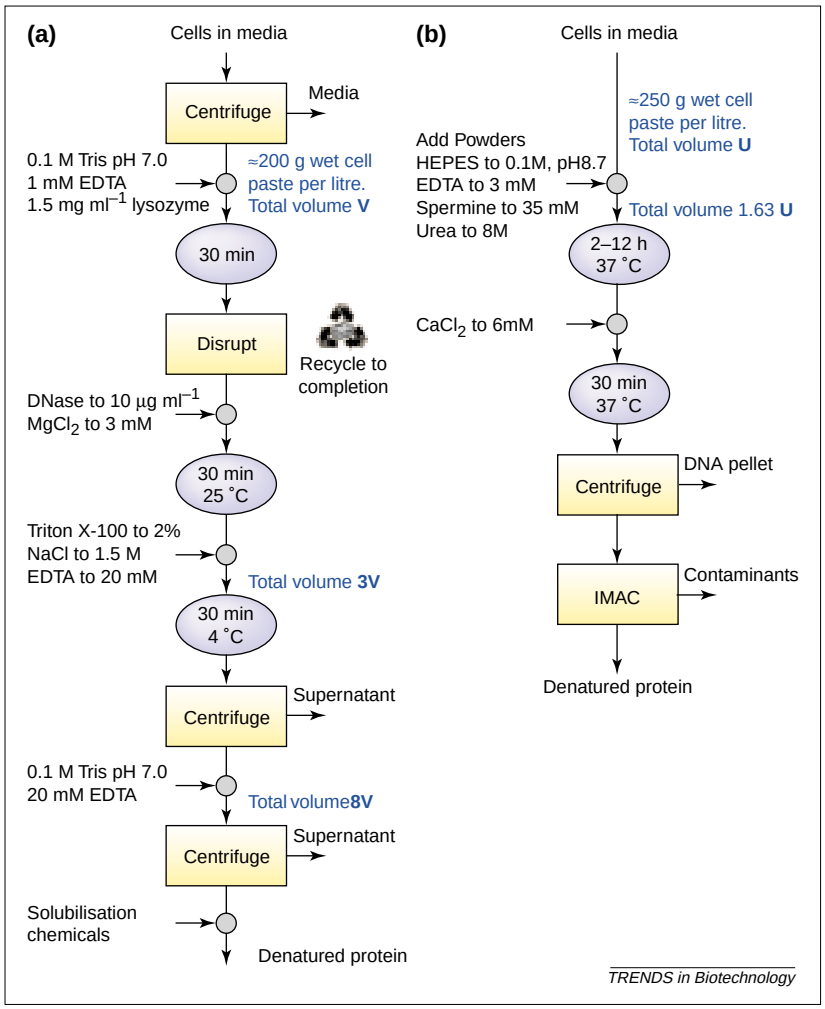

Fig. 3. Alternative methods for preparing denatured protein by solubilizing inclusion bodies. Method (a) purify then solubilize; method (b) solubilize then purify. Specific protocols are from [12] for method (a) and [17] for method (b). Method (b), as reported in [17], used an immobilized metal-affinity chromatography (IMAC) matrix with an iminodiacetic acid (IDA) or nitrolotriacetic acid (NTA) ligand. Numerous variations exist for each protocol. The time for extraction in method (b) can be minutes at low cell concentration in the presence of a reducing agent [16]. An alternative chromatographic purification method (e.g. hydrophobic interaction or ion exchange) might then be needed to prevent the reduction of metal ions during IMAC.

best promise for automated preparative work. Swartz and co-workers pioneered this technique for periplasmic inclusion bodies [15]. *E. coli* cells in fermentation broth were extracted under alkaline conditions (pH 10) in the presence of 2 M urea and 10 mM DTT. The low concentration of urea enabled simultaneous extraction and refolding, and integration with a two-phase recovery method gave 70% cumulative yield of insulin-like growth factor I (IGF-I). For cytoplasmic inclusion bodies, virtually complete extraction of a variant of IGF-I from cytoplasmic inclusion bodies has been achieved using 6 M urea, 3 mM EDTA and 20 mM DTT at pH 9.0 [16]. This approach has been extended to the direct extraction of a recombinant viral-coat protein from the cytoplasm of *E. coli*, at high cell density ($OD_{600} = 130$) and without the use of reducing agent [17]. Extraction was done directly on cells in fermentation media, overcoming the need for a cell harvest stage. DNA was removed selectively by the addition of 35 mM spermine directly into the extraction mixture, followed by low-speed centrifugation [17,18]. The extracted denatured protein was recovered at 70% overall yield and 89% purity by immobilized metal affinity chromatography, following the addition of calcium chloride to bind residual EDTA. The procedure is summarized in Figure 3b. The protein, following elution, is ready for refolding assessment or for column refolding.

The preparation of purified denatured protein is a major bottleneck in preparative protein refolding. The techniques most likely to satisfy the criteria specified earlier are summarized in Figure 3. The outcome is denatured protein ready for use in either dilution refolding or in further column-based operations.

Table 1. Common additives used in refolding buffers

Additive	Typical concentration
Urea (denaturant)	2 M
Guanidinium chloride	1 M
Arginine	0.5 M
Glycerol	0.4 M
Sucrose	0.4 M
Lauryl maltoside	0.3 mM
Polyethylene glycol (3550 MW)	0.05% w/v
Tris buffer	0.5 M
Triton X-100	10 mM
Acetamide	2 M
n-hexanol	5 mM
Salts (NaCl, Na_2SO_4, K_2SO_4)	0.5 M

Refolding by dilution

Protein refolding is initiated by a reduction in denaturant concentration and, in oxidative protein refolding, by altering the redox environment to enable disulphide bond formation. The simplest and most widely used method for reducing denaturant concentration is by dilution into an appropriate refolding buffer. Two processes occur: first-order refolding and higher order aggregation [19]. Refolding at dilute protein concentrations (<10 μg ml^{-1}) minimizes aggregation. Such low concentrations are impractical for preparative work, so yield is maximized by altering the mode of denaturant dilution or by using folding enhancers.

As the kinetic scheme for protein refolding involves competing first- and higher-order reactions, theory tells us that selectivity for the desired monomeric product is maximized by maintaining a low concentration of denatured protein [20]. Rudolph and co-workers first recognized this and patented a 'pulsed renaturation' method, whereby the denatured protein is added to the refolding buffer in pulses [21]. This strategy is useful for preparative work but impractical when searching for a suitable refolding buffer. Various additives are therefore used to enhance refolding yield and minimize aggregation. A summary of commonly used additives is provided in Table 1, and further specific details are provided elsewhere [12].

As no universal refolding buffer can be identified, it is necessary to screen a limited set of conditions for

each protein. Yasuda *et al.* [22] conducted an extensive screen of refolding additives in buffer containing various concentrations of guanidinium chloride. Water-soluble polymers [e.g. polyvinyl pyrrolidone (PVP) and poly(ethylene glycol) (PEG)] and surfactants (e.g. Triton-X 100 and Triton N57) exerted a negative influence on refolding yield, whereas urea, urea derivatives, acetone and acetoamide improved refolding yield. Cleland and Wang have previously shown that the addition of PEG to refolding buffer improved the yield of carbonic anhydrase (CAB) during refolding [23], emphasizing that additives that are good for some proteins might be bad for others.

Formal screening methods have been developed. One of the first studies used a matrix of 50 different buffers originally developed for crystallization [24]. Buffer conditions were identified that facilitated the refolding of eight of the nine proteins tested. A fractional factorial folding screen involving eight factors and two levels has been used to optimize the refolding of procathepsin S and cathepsin S [25]. An alkaline pH was beneficial for each protein but it was found that addition of arginine improved the yield of procathepsin but not of cathepsin. The final buffer for cathepsin refolding contained only Tris buffer, glycerol (vital) and a redox couple. By contrast, procathepsin refolding required arginine, detergent, sodium chloride and a redox couple. This study emphasizes that distinct but related proteins will fold differently in different buffers. A partial factorial design has also been used to optimize refolding conditions [26]. Twelve factors are tested in 16 different folding conditions. A factorial screening kit based on this method is now available commercially as the FoldIt™ kit (Hampton Research, Laguna Niguel, CA, USA).

Buffers will also need to be supplemented with appropriate redox agents if the formation of disulphide bonds is required for native structure. This is discussed in detail elsewhere [11]. Refolding from mixed disulphide has been successful for numerous proteins. The denatured protein is incubated with either oxidized glutathione or a mixture of sodium sulfite and sodium tetrathionate. Disulphide interchange is then initiated by introducing a low concentration of a reducing agent (e.g. cysteine or DTT). As an alternative, the glutathione renaturation system is simple and is widely used for preparative work. A mixture of reduced and oxidized glutathione is added to the renaturation buffer in a ratio ranging from 10:1 to 1:1, with a reduced glutathione concentration in the range 0.1–1 mM. Refolding is usually conducted at alkaline pH (pH 8–9), to promote thiolate anion formation and hence disulphide exchange. The cost of glutathione agents is substantial and, in some cases, the use of a cysteine–cystine couple provides a cost-effective alternative. For example, human retinal-binding protein has been refolded successfully in a buffer containing 3 mM cysteine and 0.3 mM cystine [27].

Dilution is a simple and approximately scale-invariant method of protein refolding. Generic screens that aid in the identification of suitable refolding conditions for some proteins have been developed. Cost effectiveness is achieved by avoiding expensive additives and thiol agents, depending on the intended scale of preparative work. Finally, refolding yields can be further improved on transfer to manufacture by altering the mode of reactor operation (e.g. by using 'pulsed renaturation').

Column refolding

Refolding using packed columns is attractive because it is easily automated using commercially available preparative chromatography systems. There are three basic approaches: (1) immobilization of the denatured protein onto a matrix and subsequent denaturant dilution to promote refolding; (2) denaturant dilution using size exclusion chromatography (SEC); and (3) immobilization of folding catalysts onto chromatographic supports so that the column behaves like a catalytic folding reactor.

Immobilization of denatured protein onto a matrix can be achieved through non-specific interactions or through specific affinity interactions. The aim is to isolate individual protein molecules spatially, thus inhibiting aggregation. Creighton [28] showed that horse cytochrome *c* could be adsorbed from urea onto an ion exchange matrix and subsequently refolded. Success is very protein specific because protein–matrix interactions can prevent refolding. Significant empirical optimization is required for each protein. Affinity interactions are generally preferred because these allow binding through specific domains, with the bulk of the protein free from the surface and hence able to refold. N- or C-terminal cationic hexa-arginine peptides have been fused to α-glucosidase to enable specific immobilization on a polyanionic support [29]. Refolding conditions require careful optimization because of non-specific protein–matrix interactions. Low concentrations of salt promoted ionic interactions with the matrix and hence reduced yield, whereas high salt promoted hydrophobic interactions that prevented renaturation. Other conditions, including pH, temperature, cosolvents and matrix material, had to be screened carefully. However, under optimized conditions, folding could be conducted at significantly higher concentrations than those achieved in dilution refolding (5 mg ml^{-1} compared with 15 μg ml^{-1}).

Nickel-chelating chromatography has also been used for affinity immobilization and refolding. Rogl *et al.* used this approach to refold membrane proteins produced as inclusion bodies in *E. coli* [30] and it has also been used for the oxidative refolding of mammalian prion proteins [31]. Nickel-chelating chromatography is attractive because cloning vectors with N- or C-terminal histidine tags are readily available. However, conditions must be optimized carefully to ensure that metal-ion reduction does not occur because of the carryover of reducing agents from solubilization. An alternative and interesting method

using a cellulose-binding domain (CBD) has also been reported [32]. In 6 M urea, the CBD retains its ability to bind specifically to a cellulose matrix, whereas the attached fusion protein is denatured and free to refold. High yield refolding of single chain antibodies was obtained (~3 times the yield obtained by dilution).

Refolding by SEC aims to inhibit aggregation by restricting the diffusion of various protein forms in the refolding mixture. The technique was developed by Werner *et al.* [33]. The refolding of hen egg-white lysozyme (HEWL) and bovine carbonic anhydrase has subsequently been achieved using Sephacryl S-100 size exclusion matrix from very high initial protein concentrations of up to 80 mg ml^{-1} [34]. High dilution factors are achieved during SEC refolding and consequently the final sample concentration is significantly reduced. For HEWL, overall refolding yield varied between 46% and 83%, depending on the starting and hence final protein concentration. Comparable yields were achievable by simple batch dilution but at high residual denaturant concentration. For example, the highest SEC yield of 83% was achieved for a feed containing 9.6 mg ml^{-1} of denatured protein, giving a refolded protein concentration of 0.18 mg ml^{-1} after dilution in the SEC column. By direct dilution into refolding buffer, a yield of 85% was achievable at a similar final protein concentration for a fourfold dilution from the denatured state. A variation is to equilibrate the SEC column with a gradient of denaturant [35]. The denatured protein then experiences a gradual decrease in urea concentration as it passes through the column. Results for HEWL suggest that the efficiency of this approach, based on the mass of protein refolded per volume of chromatographic resin, is threefold that of standard SEC refolding. These studies suggest that SEC refolding offers no significant advantage over simple dilution in terms of the yield at a given final protein concentration, at least for lysozyme. However, SEC has a key advantage over simple dilution refolding – material leaving the column has been fractionated on the basis of size. It is therefore possible to obtain refolded protein free from any contaminants present in the initial solubilization solution, and also free of aggregates and any residual unfolded protein. For preparative work these benefits can be of significant advantage.

Catalytic column refolding exploits the action of naturally occurring enzymes. Fersht and co-workers developed an oxidative refolding chromatography column that has three components immobilized on agarose: (1) GroEL minichaperone, which can prevent aggregation; (2) DsbA, which catalyses the oxidation and shuffling of disulphide bonds; and (3) peptidyl-prolyl isomerase. The matrix has been used to refold denatured and reduced scorpion toxin Cn5 [36], which has not been previously refolded in reasonable yield. A yield of 87% was achieved. Oxidative refolding chromatography has also been used to refold and assemble CD1, a major histocompatibility complex (MHC) class I-like antigen-presenting molecule [37]. Previous attempts to refold MHC class I molecules *in vitro* have required the presence of ligand. This form of column-based catalytic refolding promises to extend the range of proteins that can be refolded preparatively from inclusion bodies.

Analytical methods

The production of native protein via the inclusion body route is practical only if there is a sensitive analytical method to determine the success, or otherwise, of refolding. Obvious analytical methods exist, depending on the nature of the protein that has been refolded. Enzyme renaturation can be probed using a suitable activity assay, as commonly done for lysozyme. Therapeutic proteins can be assayed using immunoassays when a suitable antibody exists, or by using suitable bioassays. Proteins that possess internal aromatic amino acids can be probed using intrinsic fluorescence (e.g. tryptophan). Protein aggregation can be determined by measuring turbidity at 450 nm. Circular dichroism measurements give an indication of secondary structure, although interpretation of spectra can be difficult, as can automation. Limited proteolysis has been used routinely to assess the compactness of native states and, logically, can be used to compare proteins refolded in different environments. Heiring and Muller developed the 'folding screen assayed by proteolysis' method to exploit this type of analysis in a screening format [38].

Newer techniques are also being developed to complement these traditional assays and to provide information on the success or otherwise of protein refolding for those proteins that cannot be assayed using an obvious method. Online determination of molecular weight following SEC is now possible [39]. The eluate from an SEC column is probed with online light scattering and a refractive index detector to give the molecular weight of the polypeptide. This is particularly useful for detecting so-called soluble aggregates, which have been demonstrated following the expression of viral-coat protein fused to maltose-binding protein [40]. An interesting development is online capillary isoelectric focusing electrospray ionization mass spectrometry [41]. This technique detects refolding intermediates containing different numbers of disulphide bonds, and even mixed disulphides, following blockage of free thiols with iodoacetate. It can also determine conformational heterogeneity among groups of refolding intermediates. Real-time nuclear magnetic resonance (NMR) could also be developed as a viable analytical technique [42]. Finally, sophisticated variants of normal spectroscopy, such as multispectroscopic monitoring (far and near UV circular dichroism, fluorescence and UV spectrometry), coupled with detailed analysis of spectra, can give information on different protein conformations in solution [43].

New analytical methods will continue to be developed as the drive towards automation in the

proteomics field continues. At present, a combination of the methods outlined above can provide information on the success of refolding and allow a comparison of different refolding environments or methods.

Conclusions

The aim of preparative protein refolding is to obtain renatured protein from inclusion bodies quickly and using procedures that are minimally complex and easily automated. The first step involves preparing denatured and reduced protein of adequate purity for subsequent refolding (suitable methods are outlined in Fig. 3). Subsequent refolding can be achieved either by direct dilution or by column-based operations. Dilution is simple and enables the easy screening of additives and redox agents. Column-based methods can improve yield or integrate the separation and renaturation steps. Newer column-based methods, such as oxidative column chromatography, promise to extend the range of proteins that can be refolded following expression in *E. coli*. Suitable analytical methods are available to probe the products of refolding and new methods based on developments in proteomics are emerging. Finally, it is clear that off-the-shelf technology exists to allow automation of liquid-handling (dilution) and column-based refolding operations. The basis exists for highly automated preparative refolding work, especially for those proteins amenable to direct chemical extraction from the cellular cytoplasm. These technologies and developments promise to satisfy the criteria, outlined above, of scale invariance, ease of automation, applicability to a broad range of proteins and cost-effectiveness.

References

1 Carrió, M.M. and Villaverde, A. (2002) Construction and deconstruction of bacterial inclusion bodies. *J. Biotechnol.* 96, 3–12

2 Rudolph, R. and Lilie, H. (1996) *In vitro* folding of inclusion body proteins. *FASEB J.* 10, 49–56

3 Middelberg, A.P.J. (1995) Process-scale disruption of microorganisms. *Biotechnol. Adv.* 13, 491–551

4 Valax, P. and Georgiou, G. (1993) Molecular characterization of beta-lactamase inclusion-bodies produced in *Escherichia coli*. 1. Composition. *Biotechnol. Prog.* 9, 539–547

5 Maachupalli-Reddy, J. *et al.* (1997) Effect of inclusion body contaminants on the oxidative renaturation of hen egg white lysozyme. *Biotechnol. Prog.* 13, 144–150

6 Tran-Moseman, A. *et al.* (1999) Renaturation of *Escherichia coli*-derived recombinant human macrophage colony-stimulating factor. *Protein Expr. Purif.* 16, 181–189

7 Babbitt, P.C. *et al.* (1990) Removal of a proteolytic activity associated with aggregates from expression of creatine-kinase in *Escherichia coli* leads to improved recovery of active enzyme. *Biotechnology* 8, 945–949

8 Wong, H.H. *et al.* (1996) Centrifugal recovery and dissolution of recombinant Gly-IGF-II inclusion bodies: the impact of feedrate and re-centrifugation on protein yield. *Bioseparation* 6, 185–192

9 Geourgiou, G. and Valax, P. (1999) Isolating inclusion bodies from bacteria. *Methods Enzymol.* 309, 48–58

10 Wong, H.H. *et al.* (1997) Cumulative sedimentation analysis of *Escherichia coli* debris size. *Biotechnol. Bioeng.* 55, 556–564

11 Fischer, B. *et al.* (1993) Isolation, renaturation, and formation of disulfide bonds of eukaryotic proteins expressed in *Escherichia coli* as inclusion bodies. *Biotechnol. Bioeng.* 41, 3–13

12 Clark, E.D. *et al.* (1999) Inhibition of aggregation side reactions during *in vitro* protein folding. *Methods Enzymol.* 309, 217–236

13 Vuillard, L. *et al.* (1998) Interactions of non detergent sulfobetaines with early folding intermediates facilitate *in vitro* protein renaturation. *Eur. J. Biochem.* 256, 128–135

14 Puri, N.K. and Cardamone, M. (1992) A relationship between the starting secondary structure of recombinant porcine growth hormone solubilized from inclusion bodies and the yield of native (monomeric) protein after *in vitro* refolding. *FEBS Lett.* 305, 177–180

15 Hart, R.A. *et al.* (1994) Large-scale, *in situ* isolation of periplasmic IGF-I from *E. coli*. *Biotechnology* 12, 1113–1117

16 Falconer, R.J. *et al.* (1998) Chemical treatment of *Escherichia coli*. 2. Direct extraction of recombinant protein from cytoplasmic inclusion bodies in intact cells. *Biotechnol. Bioeng.* 57, 381–386

17 Choe, W-S. *et al.* (2002) Comparison of histidine-tag capture chemistries for purification following chemical extraction. *J. Chromatogr. A* 953, 111–121

18 Choe, W-S. and Middelberg, A.P.J. (2001) Selective precipitation of DNA by spermine during the chemical extraction of insoluble cytoplasmic protein. *Biotechnol. Prog.* 17, 1107–1113

19 Kiefhaber, T. *et al.* (1991) Protein aggregation *in vitro* and *in vivo*: A quantitative model of the kinetic competition between folding and aggregation. *Biotechnology* 9, 825–829

20 Kotlarski, N. *et al.* (1997) Optimal reactor selection for refolding monomeric protein. *AIChE J.* 43, 2123–2132

21 Rudolph, R. and Fischer, S. (1990) Process for obtaining renatured proteins. US Patent 4,933,434

22 Yasuda, M. *et al.* (1998) Effect of additives on refolding of a denatured protein. *Biotechnol. Prog.* 14, 601–606

23 Cleland, J.L. and Wang, D.I.C. (1993) Cosolvent effects on refolding and aggregation. *ACS Symposium Ser.*, 516, 151–166

24 Hofmann, A. *et al.* (1995) A sparse-matrix screen to establish initial conditions for protein renaturation. *Anal. Biochem.* 230, 8–15

25 Tobbell, D.A. *et al.* (2002) Identification of *in vitro* folding conditions for procathepsin S and cathepsin S using fractional factorial screens. *Protein Expres. Purif.* 24, 242–254

26 Armstrong, N. *et al.* (1999) A new protein folding screen: Application to the ligand binding domains of a glutamate and kainite receptor and to lysozyme and carbonic anhydrase. *Protein Sci.* 8, 1475–1483

27 Xie, Y.S. *et al.* (1998) Recombinant human retinol-binding protein refolding, native disulfide formation, and characterisation. *Protein Expres. Purif.* 14, 31–37

28 Creighton, T.E. (1986) Folding of proteins adsorbed reversibly to ion-exchange resins. *Protein Struct. Fold. Des.* 249–257

29 Stempfer, G. *et al.* (1996) Improved refolding of an immobilised fusion protein. *Nat. Biotechnol.* 14, 329–334

30 Rogl, H. *et al.* (1998) Refolding of *Escherichia coli* produced membrane protein inclusion bodies immobilised by nickel chelating chromatography. *FEBS Lett.* 432, 21–26

31 Zahn, R. *et al.* (1997) Human prion proteins expressed in *Escherichia coli* and purified by high affinity column refolding. *FEBS Lett.* 417, 400–404

32 Berdichevsky, Y. *et al.* (1999) Matrix-assisted refolding of single-chain Fv-cellulose binding domain fusion proteins. *Protein Expres. Purif.* 17, 249–259

33 Werner, M.H. *et al.* (1994) Refolding proteins by gel filtration chromatography. *FEBS Lett.* 345, 125–130

34 Batas, B. and Chaudhuri, J.B. (1996) Protein refolding at high concentration using size-exclusion chromatography. *Biotechnol. Bioeng.* 50, 16–23

35 Gu, Z. *et al.* (2001) Urea gradient size-exclusion chromatography enhanced the yield of lysozyme refolding. *J. Chromatogr. A* 918, 311–318

36 Altamirano, M.M. *et al.* (1999) Oxidative refolding chromatography: folding of the scorpion toxin Cn5. *Nat. Biotechnol.* 17, 187–191

37 Altamirano, M.M. *et al.* (2001) Ligand-independent assembly of recombinant human CD1 by using oxidative refolding chromatography. *Proc. Natl. Acad. Sci. U. S. A.* 98, 3288–3293

38 Heiring, C. and Muller, Y.A. (2001) Folding screening assayed by proteolysis: application to various cystine deletion mutants of vascular endothelial growth factor. *Protein Eng.* 14, 183–188

39 Wen, J. *et al.* (1996) Size-exclusion chromatography with on-line light-scattering, absorbance, and refractive index detectors for studying proteins and their interactions. *Anal. Biochem.* 240, 155–166

40 Nominé, Y. *et al.* (2001) A strategy for optimizing the monodispersity fusion proteins: Application to purification of recombinant HPV E6 oncoprotein. *Protein Eng.* 14, 297–305

41 Jensen, P.K. *et al.* (1998) Monitoring protein refolding induced by disulfide formation using capillary isoelectric focusing electrospray ionization mass spectrometry. *Anal. Chem.* 70, 2044–2049

42 Balbach, J. *et al.* (1995) Following protein-folding in real-time using NMR-spectroscopy. *Nat. Struct. Biol.* 2, 865–870

43 Navea, S. *et al.* (2001) Three-way data analysis applied to multi-spectroscopic monitoring of protein folding. *Anal. Chim. Acta* 446, 187–197