Genetic screens and directed evolution for protein solubility

Geoffrey S Waldo

Overexpressed proteins are often insoluble, and can be recalcitrant to conventional solubilization techniques such as refolding. Directed evolution methods, in which protein diversity libraries are screened for soluble variants, offer an alternative route to obtaining soluble proteins. Recently, several new protein solubility screens have been developed that do not require structural or functional information about the target protein. Soluble protein can be detected *in vivo* and *in vitro* by fusion reporter tags. Protein misfolding can be measured *in vivo* using the bacterial response to protein misfolding. Finally, soluble protein can be monitored by immunological detection. Efficient, well-established strategies for generating and recombining genetic diversity, driven by new screening and selection methods, can furnish correctly folded, soluble protein.

Addresses
BN-2, MS M888, Los Alamos National Laboratories, Los Alamos, NM 87545, USA

Current Opinion in Chemical Biology 2003, **7**:33–38

This review comes from a themed issue on
Proteomics and genomics
Edited by Matthew Bogyo and James Hurley

1367-5931/03/$ – see front matter
© 2003 Elsevier Science Ltd. All rights reserved.

DOI 10.1016/S1367-5931(02)00017-0

Abbreviations
CAT chloramphenicol acetyltransferase
GFP green fluorescent protein

Introduction

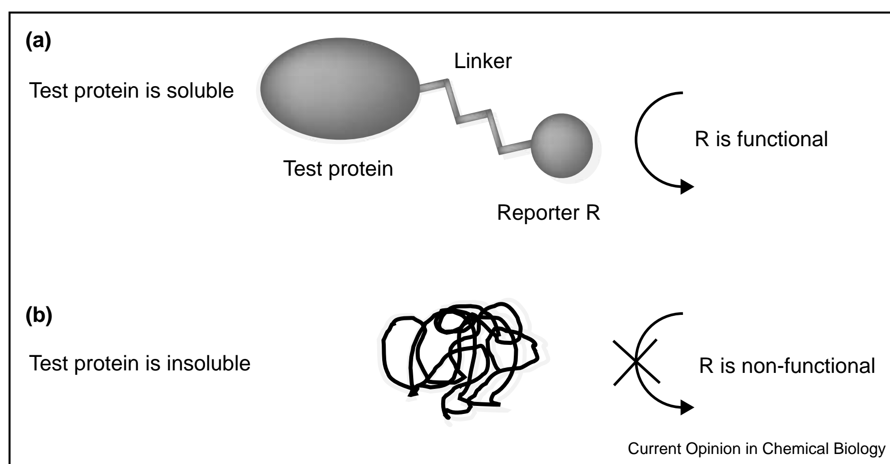
Conventional approaches to obtain soluble, correctly folded protein include low-temperature expression, promoters with different strengths, a variety of solubility-enhancing fusion tags [1], and modified growth media (reviewed in [2]). Insoluble proteins can sometimes be chemically unfolded then refolded [3]. These approaches can occasionally be successful, but are not ideal for high-throughput applications. First, many different kinds of operations must be carried out on each target. Second, only a relatively small number of conditions can be tried before the target is committed to the production pipeline. Finally, these methods fail to modify the intrinsic folding stability and solubility of the protein. Even when a fraction of the protein is obtained in a soluble form during initial expression or refolding experiments, the protein may still aggregate irreversibly during subsequent workup and concentration. This problem is especially costly for high-throughput structural genomics during the steps of crystallization or NMR structure determination.

An alternative approach is to somehow engineer the target protein to improve its intrinsic folding yield, stability and solubility. When structural information is available, site-directed rational mutation can sometimes result in proteins with improved solubility [4]. *In silico* methods can be used to predict stabilizing mutations, but the potentially deleterious effects of these mutations on folding pathways cannot be anticipated [5]. Function-based *in vivo* screening and selection methods can sometimes be devised [6]. Cell lysates may be screened for increased activity using high-throughput liquid handling techniques [7], but are limited by the throughput of current liquid-handling systems. All these methods lack generality because they require structural or functional information about the target. Not all proteins have easily measured activities, and it may be impractical to devise high-throughput functional assays for large numbers of different targets. This review describes recent advances in methods for selecting and screening soluble proteins, focusing on the time-frame 2001 through 2002. Particular emphasis is given to methods suitable for high-throughput structural genomics.

Structure and function independent screens for improved protein solubility

A more general approach is to find better-performing, more soluble versions of a recalcitrant protein by directed evolution. Libraries of protein variants (random point mutations, deletions, fragments) are generated using well-established protocols (reviewed in [8••]), and the expressed gene products are screened or selected *in vivo* or *in vitro* for improved solubility. Recently developed protein solubility screening and selection methods fall into three classes. None rely on the function of the protein of interest. In fusion reporter methods (Figure 1), a test protein and a reporter protein with an easily detected function or biological activity are expressed as a genetic fusion. Information about the folding and solubility of the test protein is transduced to a screenable or selectable activity by the fused reporter domain. These approaches are simple to implement but can perturb the solubility of the test protein. Stress-reporter methods use the innate host cell response to misfolded proteins. In principle, no tagging is required. Instead, stress-responsive promoters are cloned in front of a gene product with an easily measured function. Expression of the reporter is activated during expression of a misfolding target protein. Some independent means of confirming clones that evoke a reduced stress response is necessary, since these are

Figure 1



Principle of fusion solubility reporter. A reporter domain with an easily measured function is fused to a test domain via a flexible linker. **(a)** When the test protein is soluble, the reporter domain is functional and its activity can be detected. **(b)** On the other hand, when the test protein is insoluble or misfolded, the function of the reporter domain is compromised. The reporter domain can be any moiety with an easily measured function. The mode of the assay depends only on the function of the reporter, which can be a selectable marker protein such as a metabolic enzyme or antibiotic resistance protein or a screenable marker such as an enzyme or fluorescent protein. Intact proteins or fragments of proteins can be used, such as split protein interactors.

positive reporters of stress, not of expression and solubility. Direct methods involve the physical separation of soluble protein from insoluble protein, and the subsequent quantification of these fractions by immunological means. Antibodies to tags or to intrinsic domains of the test protein can be used.

Fusion reporter methods

In the green fluorescent protein (GFP) folding reporter method [9], a test domain is expressed as an N-terminal fusion with GFP. The fluorescence yield of the GFP transduces information about the folding success or failure of the upstream fusion partner. Cells expressing GFP fused to well-folded soluble proteins are brighter than those expressing GFP fused to poorly folded insoluble proteins. The GFP is used only to monitor the folding yield of the test protein, which is subsequently expressed without the GFP tag. *In vitro* recombination at a low to moderate rate of mutagenesis is used to efficiently generate and recombine folding mutants [10,11]. *Escherichia coli* expressed libraries of 1–5 $\times$ 10^4 variants are screened during 3–5 rounds of forward evolution for improved brightness on plates or by flow cytometry. Several examples of the application of this method to generate soluble variants of intractable proteins have been published. Bowie and co-workers [12•] engineered a soluble variant of the SAM/TEL domain of a leukemogenesis protein. From the structure of the evolved variant, they were able to rationally design second-site mutations that restored the ability of the protein to multimerize. Suh and co-workers [13] evolved a soluble variant of a metabolic enzyme from the human pathogen *Mycobacterium tuberculosis*. The detailed structural information of the active site provided clues as to the unique substrate specificity of the enzyme. Waldo and co-workers [14••] engineered a soluble variant of a hexameric nucleoside diphosphate kinase from the hyperthermophile *Pyrobaculum aerophilum*. Refolding trials of the wild-type protein had previously yielded insufficient correctly folded material for crystallization trials. The structural information explained in part the enzyme's preference for the less-bulky substrate TTP. Interestingly, the evolved protein selected a pair of compensatory charge interactions at the dimer interface, also seen in a naturally occurring human homologue. None of the mutations were near the active site, and the evolved protein maintained high structural similarity with its closest human homologue. Hecht and co-workers [15•] used the GFP folding reporter to discover more soluble variants of the Alzheimer α/β precursor protein. They found mutations that had previously been described, as well as several new solubilizing mutations. Kawasaki and Inagaki [16•] used a random PCR approach to clone fragments of a large, insoluble protein in front of GFP. The approach yielded several subdomains with improved solubility. This 'divide and conquer' approach should have considerable utility obtaining soluble subdomains of large, intractable proteins, especially for NMR structural applications.

Davidson and co-workers recently demonstrated that the survival of *E. coli* cells on media containing the antibiotic chloramphenicol, was positively correlated with the solubility of test proteins expressed in the cells as N-terminal fusions with chloramphenicol acetyltransferase (CAT)

protein. Growing the cells on media containing progressively higher levels of chloramphenicol provides the selective pressure. The method was used to discriminate a minority population of known soluble variants from a pool of cells containing an excess of less soluble wild-type clones [17]. Recently, Arnold and co-workers [18•] used the method to enrich a library of mutants for more soluble variants before the application of a lower-throughput functional assay. The authors screened a library of interspecies hybrids of a membrane-associated human cytochrome P450 (1A2) and the heme domain of a soluble bacterial P450 (BM3). Genetic fusions to CAT facilitated selection for soluble protein variants. Screening for 1A2 activity (deethylation of 7-ethoxyresorufin) identified two functional P450 hybrids that were more soluble than the wild-type 1A2 enzyme. It remains to be seen whether the CAT fusion method is applicable to the wide range of targets encountered in structural genomics, because the CAT protein is an obligate trimer, and may lead to the formation of higher-order aggregates of otherwise soluble multimeric proteins. A similar problem has been encountered using tetrameric red fluorescent protein as a genetic marker of multimeric proteins [19].

In the lacZα (β galactosidase α peptide) complementation solubility reporter assay, the small ca. 100 amino acid lacZα peptide tag is fused to a test protein [20••]. If the fusion protein remains soluble and the tag is not hidden, the lacZα peptide can bind inactive lacZΩ supplied by the host *in trans*, or *in vitro*, restoring β galactosidase activity. Split protein genetic complementation could provide some advantages. First, a small tag might be expected to have less of an effect on the folding and solubility of the test protein compared with a larger protein such as CAT. Second, because the complementation depends only on the accessibility of the tag in soluble form, peptide tag solubility reporters might be especially useful for monitoring slow protein aggregation processes. Third, the lacZα system can function as a selectable marker, conferring survival in the appropriate host using minimal media in which lactose is the carbon source. This would enable large libraries to be selected for improved solubility ($1 \times 10^{7-9}$). It remains to be seen whether the lacZα solubility reporter is suitable for high-throughput engineering of proteins for structural genomics. To date, the published examples of the application of the lacZα system amount to tantalizing proof-of-principle experiments in which a limited number of highly soluble or very insoluble proteins were discriminated *in vivo* and *in vitro*. No naïve libraries have been screened for improved solubility using the lacZα system, but a known soluble protein variant gave a higher level of complementation than the less soluble wild type [20••]. Though small, the lacZα peptide is not entirely innocent. Many partially soluble proteins are insoluble when expressed as lacZα fusions (Waldo GS *et al.*, unpublished data). This effect is not unexpected, as many split proteins have poor solubility and folding characteristics compared with their full-length counterparts [21].

Barberis and co-workers have developed a fusion reporter system they term 'Quality Control' using a selectable genetic marker for the identification of soluble single chain antibodies (scFv). The system operates in yeast, using fusion to components analogous to the classical 'two hybrid' system [22•]. The scFv is fused to a selectable marker protein comprising a transcription activation domain (AD) and a peptide fragment derived from Gal11P. If the scFV is stable and soluble, the fused AD-Gal11P domain can associate with the DNA-bound Gal4 (1–100) fragment, thereby activating transcription of HIS3 and lacZ reporter genes. In the appropriate strain background and on selective media, HIS3 expression allows survival of the host cell (selection), whereas lacZ allows blue/white screening. Thus, cells expressing soluble scFV will survive, while those expressing unstable or insoluble scFV will die because of the lack of expression of the selectable marker HIS3. The authors showed a good correlation between solubility and cell survival for four previously characterized scFvs, three soluble and one insoluble. It remains to be seen whether the system will function well for the wide variety of protein scaffolds encountered in structural genomics.

In the S-protein split protein assay, an inactive fragment of RNase-A (S-protein) missing its N-terminal 15 amino acids is rescued by binding to the S-peptide derived from the N-terminal 15 amino acids of RNase-A. Recently, Raines and co-workers [23] developed a highly sensitive fluorogenic donor-quencher ribonucleotide substrate for RNase-A. Cleavage of the substrate by RNase-A separates the donor and quencher, leading to a dramatic (>180-fold) increase in the fluorescence. This opened the door for sensitive, rapid high-throughput detection of femtomolar quantities of S-peptide-tagged proteins using the S-protein complementation reaction. The fluorogenic assay is marketed in kit form by Novagen, Inc. (FRETWorks®). To date, the fluorogenic substrate for the S-peptide assay has been limited to *in vitro* applications, and is best suited to screening protein expression and solubility in microtiter plate format. Higher throughputs will require an *in vivo* version of the assay or miniaturization of *in vitro* screening methodologies.

Protein solubility is often highly dependent on stability and folding robustness. Recent enhancements of phage display techniques enable the selection of peptides with increased stability, independent of protein function (i.e. the PROSIDE technique that links increased protease resistance with display of the infective phage tip) [24,25]. Schmid and co-workers [26] applied the PROSIDE technique to select a thermostable version of a small cold shock protein. They demonstrated two selection pressures, denaturant or elevated temperatures, providing two

different strategies for protein stabilization. However, the utility of this method for larger multimeric proteins remains to be demonstrated. Baker and co-workers [27••] devised a clever selection stratagem for trapping peptides with reduced entropy (increased foldedness). In their approach, peptides are displayed at a permissive site of an SH2 domain, which is in turn displayed on phage. Peptides with low entropy (i.e. stability and close N and C termini) preserve the function of the SH2 domain. The phage pools were panned on phosphotyrosine beads to recover folded sequences. It remains to be seen whether a similar strategy can be applied to larger, multimeric proteins.

Solubility and folding reporters using host cell stress response

Bentley and co-workers [28] used a non-invasive stress reporter based on heat shock promoter driven GFP expression to monitor cellular stress to various stimuli such as heat shock, osmotic stress, ethanol and induction. The authors used differential mRNA expression and a DNA gene chip to examine *E. coli* cellular responses to various stressors, and identified upregulation of heat shock responsive promoters ftsH, clpP, lon, ompT, degP, groEL, aceA and ibpA during recombinant protein overexpression [29]. Cortazzo *et al.* [30] showed they could identify silent mutations that strongly affected the folding yield of a protein *in vivo*, using the stress-responsive promoter to drive a reporter gene. Recently, Lesley *et al.* [31••] did a more rigorous search for *E. coli* promoters that were most strongly and specifically activated during overexpression of misfolded proteins. A gene expression chip containing an array of *E. coli* open reading frames was used to assess differential gene expression of *E. coli* overexpressing insoluble proteins versus soluble proteins. Certain genes were specifically activated when misfolded proteins were expressed. On the basis of these experiments, the promoter for the small heat shock protein ibpA was selected and used to control the expression of a lacZ reporter. The authors showed they could discriminate soluble, partially soluble, and insoluble polyhistidine-tagged recombinant proteins when the *in vivo* stress reporter was used in conjunction with a secondary dot-blot screen for soluble expression using an anti-polyhistidine antibody conjugate. The authors used the method to screen a random fragment library of a large, insoluble protein. They were able to find a small, soluble N-terminal domain. Because stress reporters are *positive* reporters of protein misfolding, but not of productive protein folding and solubility, it remains to be seen whether they can be used reliably alone to select for full-length, high expresser variants.

Direct detection methods

Knaust and Nordlund [32] have developed a convenient screen for polyhistidine-tagged soluble protein expression using filter plates and dotblots. Total cell lysate is applied to a multiwell filter plate, and vacuum is applied. The soluble fraction passes through the filter membrane and is collected in a multiwell plate. The insoluble fraction, retained on the filter plate, is solubilized using a denaturant and collected in a second multiwell plate. The fractions are applied to membranes, washed, probed with antipolyhistidine antibodies, and developed. The group demonstrated the ability to easily discriminate two proteins (one soluble, the other insoluble) randomly arrayed in a 96-well plate in a double-blind experiment. Currently, the size of the libraries that can be screened is limited by the available plastic ware. Furthermore, performing the antibody blots requires lengthy cycles of binding, washing and probing.

Peabody and co-workers [33••] developed an interesting screen for proteins with altered multimerization and solubility. The authors grew colonies of *E. coli* expressing plasmid-encoded variants of an MS2 viral capsid coat protein on nutrient agar plates. The colonies were overlayed with a layer of agar containing a lysis reagent and an antibiotic to inhibit satellite culture growth. Soluble proteins diffused from the colonies and into the agar. Capture membranes were applied to bind the released proteins, probed using an anti-MS2 coat protein antibody, and developed using conventional ELISA techniques. The authors noted that spot diameter and intensity were positively correlated with the amount of soluble protein, as well as the rate of diffusion of the protein. Soluble variants that were dimeric diffused faster and gave larger diameter spots compared with clones expressing the higher-order soluble multimers characteristic of the wild-type protein. Soluble proteins gave much more intense spots compared with those of an insoluble mutant. Further work is needed to accurately discriminate increased solubility from decreased multimerization. Obviously, the method could also be used with an antibody to a generic tag (i.e. a polyhistidine antibody). Currently, the method can screen a few hundred clones per 8 cm Petri plate, as adequate spacing between colonies is required to avoid overlap of signals from individual colonies.

Conclusion

An ideal method for screening proteins for improved solubility would operate on a large number of variants (1×10^4–1×10^9), be able to detect improvements resulting from as few as one or two point mutations, would be inexpensive, operate on any protein, and would have a high reliability of predicting which variants will be well-behaved during subsequent large-scale expression, purification and concentration steps. The method would be useable *in vivo* as well as *in vitro*, and would not perturb the behaviour of the protein of interest. No single reporter so far devised embodies all these characteristics. However, given the expanding armamentarium of screening and selection-based techniques, researchers can often devise a selection method suitable for a particular protein.

In the immediate future, hybrid methods could be envisioned in which a series of different screens is applied in tandem. For example, an inexpensive (but less precise) high-throughput screen or selection for improved folding and solubility might be followed by more expensive (but more precise) screens for solubility, aggregation or function. Finally, for structural genomics applications, proteins and their variants passing this gauntlet would be screened using the most expensive and time-consuming method of all — crystallizability.

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. Kapust RB, Waugh DS: ***Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused**. *Protein Sci* 1999, **8**:1668-1674.

2. Makrides SC: **Stratagies for achieving high-level expression of genes in *Escherichia coli***. *Microbiol Rev* 1996, **60**:512-538.

3. Armstrong N, de Lencastre A, Gouaux E: **A new protein folding screen: application to the ligand binding domains of a glutamate and kainate receptor and to lysozyme and carbonic anhydrase**. *Protein Sci* 1999, **8**:1475-1483.

4. Dale GE, Broger C, Langen H, D'Arcy A, Stuber 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.

5. Malakauskas SM, Mayo SL: **Design, structure and stability of a hyperthermophilic protein variant**. *Nat Struct Biol* 1998, **5**:470-475.

6. Suzuki T, Yasugi M, Arisaka F, Yamagishi A, Oshima T: **Adaptation of a thermophilic enzyme 3-isopropylmalate dehydrogenase to low temperatures**. *Protein Eng* 2001, **14**:85-91.

7. Sun L, Petrounia IP, Yagasaki M, Bandara G, Arnold F: **Expression and stabilization of galactose oxidase in *Escherichia coli* by directed evolution**. *Protein Eng* 2001, **14**:699-704.

8. •• Farinas ET, Butler T, Arnold F: **Directed enzyme evolution**. *Curr Opin Biotechnol* 2001, **12**:545-551.
A comprehensive review of recent advances in methodologies for engineering enzymes with enhanced functionality and stability, with emphasis on methods for creating and screening functional diversity.

9. Waldo GS, Standish BM, Berendzen J, Terwilliger TC: **Rapid protein-folding assay using green fluorescent protein**. *Nat Biotechnol* 1999, **17**:691-695.

10. Stemmer WPC: **Rapid evolution of a protein *in vitro* by DNA shuffling**. *Nature* 1994, **370**:389-391.

11. Zhao H, Arnold FH: **Optimization of DNA shuffling for high fidelity recombination**. *Nucleic Acids Res* 1997, **25**:1307-1308.

12. • Kim CA, Phillips ML, Kim W, Gingery M, Tran HH, Robinson MA, Faham S, Bowie JU: **Polymerization of the SAM domain of TEL in leukemogenesis and transcriptional repression**. *EMBO J* 2001, **20**:4173-4182.
The first example of a published crystal structure of a recalcitrant protein engineered by a single round of mutagenesis and screening using the GFP fusion folding reporter assay.

13. Yang JK, Park M, Waldo GS, Suh S: **Crystallization and preliminary X-ray crystallographic analysis of the Rv2002 gene product from *Mycobacterium tuberculosis*, a β-ketoacyl carrier protein reductase homologue**. *Acta Crystallogr D* 2002, **58**:303-305.

14. •• Pedelacq J-D, Piltch E, Liong EC, Berendzen J, Kim C-Y, Rho B-S, Park MS, Terwilliger TC, Waldo GS: **Engineering soluble proteins for structural genomics**. *Nat Biotechnol* 2002, **20**:927-932.
The first example of a published crystal structure of a recalcitrant protein engineered using multiple rounds of directed evolution driven by the GFP fusion folding reporter assay.

15. • Wurth C, Guimard NK, Hecht MH: **Mutations that reduce aggregation of the Alzheimer's A beta 42 peptide: an unbiased search for the sequence determinants of A beta amyloidogenesis**. *J Mol Biol* 2002, **319**:1279-1290.
Demonstration that the GFP folding reporter can be used to select more soluble variants of a protein normally associated with slow aggregation.

16. • Kawasaki M, Inagaki F: **Random PCR-based screening for soluble domains using green fluorescent protein**. *Biochem Biophys Res Commun* 2001, **280**:842-844.
The authors demonstrate that the GFP folding reporter can be used to find soluble subdomains of a large insoluble protein by screening a random fragment library. The approach is especially promising for NMR and crystallographic studies of otherwise intractable proteins.

17. Maxwell KL, Mittermaier AK, Forman-Kay JD, Davidson AR: **A simple *in vivo* assay for increased protein solubility**. *Protein Sci* 1999, **8**:1908-1911.

18. • Sieber V, Martinez CA, Arnold FH: **Libraries of hybrid proteins from distantly related sequences**. *Nat Biotechnol* 2001, **19**:456-460.
The authors selected an interspecies hybrid library for protein variants with improved solubility using fusions with chloramphenicol resistance protein, before the application of a lower-throughput functional assay.

19. Gavin P, Devenish RJ, Prescott M: **An approach for reducing unwanted oligomerisation of DsRed fusion proteins**. *Biochem Biophys Res Commun* 2002, **298**:707-713.

20. •• Wigley WC, Stidham RD, Smith NM, Hunt JF, Thomas PJ: **Protein solubility and folding monitored *in vivo* by structural complementation of a genetic marker protein**. *Nat Biotechnol* 2001, **19**:131-136.
The authors fused lacZα to several test proteins and complementation with the lacZΩ fragment as a solubility reporter assay. Restored enzymatic activity of the lacZ was positively correlated with test protein solubility.

21. Michnick SW: **Exploring protein interactions by interaction-induced folding of proteins from complementary peptide fragments**. *Curr Opin Struct Biol* 2001, **11**:472-477.

22. • Auf der Maur A, Escher D, Barberis A: **Antigen-independent selection of stable intracellular single-chain antibodies**. *FEBS Lett* 2001, **508**:407-412.
Fusion of single-chain antibodies to a genetically selectable marker protein allows the identification of soluble single chains expressed in yeast cytoplasm.

23. Keleman BR, Klink TA, Behhlke MA, Eubanks SR, Leland PA, Raines RT: **Hypersensitive substrate for ribonucleases**. *Nucleic Acids Res* 1999, **27**:3696-3701.

24. Sieber V, Pluckthun A, Schmid FX: **Selecting proteins with improved stability by a phage-based method**. *Nat Biotechnol* 1998, **16**:955-960.

25. Jung S, Honegger A, Pluckthun A: **Selection for improved protein stability by phage display**. *J Mol Biol* 1999, **294**:163-180.

26. Martin A, Sieber V, Schmid FX: ***In-vitro* selection of highly stabilized protein variants with optimized surface**. *J Mol Biol* 2001, **309**:717-726.

27. •• Minard P, Scalley-Kim M, Watters A, Baker D: **A "loop entropy reduction" phage-display selection for folded amino acid sequences**. *Protein Sci* 2001, **10**:129-134.
A permissive site for the presentation of peptides on phage-displayed SH2 domain was used as a stratagem to select folded from unfolded sequences, using the function of the SH2 domain for selection. This provides an example of fusion selection in which the test polypeptide is displayed at an internal permissive site, rather than on the ends of the reporter domain.

28. Cha HJ, Srivastava R, Vakharia VN, Rao G, Bentley WE: **Green fluorescent protein as a noninvasive stress probe in resting *Escherichia coli* cells**. *Appl Environ Microbiol* 1999, **65**:409-414.

29. Gill RT, Valdes JJ, Bentley WE: **A comparative study of global stress gene regulation in response to overexpression of**

recombinant proteins in *Escherichia coli*. *Metab Eng* 2000, **2**:178-189.

30. Cortazzo P, Cerveñansky C, Marín M, Reiss C, Ehrlich R, Deana A: **Silent mutations affect *in vivo* protein folding in *Escherichia coli***. *Biochem Biophys Res Comm* 2002, **293**:547-551.

31. Lesley SA, Graziano J, Cho CY, Knuth MW, Klock HE: **Gene expression response to misfolded protein as a screen for soluble recombinant protein**. *Protein Eng* 2002, **15**:153-160.
•• The heat shock promoter IbpA is used to drive the expression of the lacZ reporter gene, whose expression is up-regulated during expression of misfolded recombinant polypeptides. The authors use the stress reporter and an antipolyhistidine antibody blot assay to distinguish soluble, partially soluble, and insoluble protein expression, and to identify a small soluble domain from a library of fragments of a larger insoluble protein.

32. Knaust RKC, Nordlund P: **Screening for soluble expression of recombinant proteins in a 96-well format**. *Anal Biochem* 2001, **297**:79-85.

33. Peabody DS, Al-Bitar L: **Isolation of viral coat protein mutants with altered assembly and aggregation properties**. *Nucleic Acids Res* 2001, **29**:E113.
•• Colonies expressing recombinant proteins are lysed in an agarose overlay and the liberated proteins diffuse at rates dependent on their molecular sizes. The proteins are transferred to a membrane and visualized by antibody probe.