© 2015 Nature America, Inc. All rights reserved.

Rapid screening for protein solubility and expression

The Expresso® Solubility and Expression Screening System uses a simple ligase-free cloning strategy and a single PCR product to construct and test seven distinct protein fusion partners in parallel over the course of a few days. This multiplexed system simplifies the screening of fusion partners and greatly increases the chances of obtaining improved expression and solubility of target proteins.

Many heterologous proteins are insoluble or poorly expressed in *Escherichia coli*. One way to address this problem is to fuse a 'solubility tag' to the target protein. Vectors for cloning and expressing solubility tags are available in numerous formats and include fusion partners such as maltose-binding protein (MBP)[1], glutathione *S*-transferase (GST)[2], small ubiquitin-related modifier (SUMO)[3] and others. For reasons unknown, no single tag is universally successful, and it is currently impossible to predict which fusion partner will perform optimally for a given target protein.

Empirical fusion-tag selection is the best option for important target proteins that are poorly or improperly expressed in *E. coli*; however, there is no unified system for rapid screening to identify the best tag. Variability among designs and disparate cloning strategies not only make it difficult to test different fusion partners against an important target, but also obfuscate meaningful comparisons once the fusions have been achieved.

The pSol suite of fusion vectors

The Expresso Solubility and Expression Screening System features a diverse set of fusion partners in a unified vector design (**Fig. 1**). Seven distinct fusion tags engineered into a suite of cloning-ready pSol vectors enable rapid evaluation in parallel. Several tags were chosen because of their common use and proven success as solubility-enhancing factors (MBP, GST and SUMO). Others were included on the basis of published reports describing them as factors with innate abilities to improve the solubility of proteins fused to them (SlyD[4], Bla[5] and Tsf[6]). Amino acid composition is a major factor contributing to protein solubility, with acidic amino acids having a positive influence on protein folding. In our search for additional solubility-enhancing factors, we found the hypothermostable protein AFV_{1-99} from Acidianus filamentous virus 1 (ref. 7). AFV_{1-99} is a 13.5-kDa protein with a native isoelectric point (pI) of 4.6 that remains folded under high temperature and low pH.

Each vector contains the $rhaP_{BAD}$ promoter for stable cloning and strong tunable expression within a single host strain[8]. The vectors also encode a 6xHis tag at the N terminus of each fusion tag to facilitate affinity-column purification. A TEV protease recognition sequence joins each fusion tag to the N terminus of the target protein. Thus, the fusion tags can be easily cleaved from the protein of interest using SelecTEV™ protease after expression and purification. SelecTEV protease and the fusion tags all contain N-terminal His tags and can be removed from TEV digestion reactions by immobilized-metal affinity chromatography (IMAC) purification.

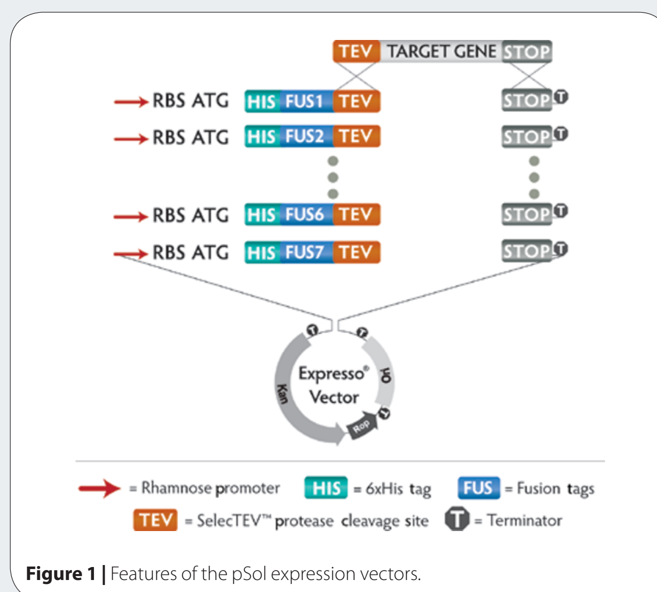

Figure 1 | Features of the pSol expression vectors.

Unique workflow

With current technology, fusion-tag selection is a time-consuming trial-and-error process that lacks uniformity in promoter selection and cloning technologies. The Expresso system uses a simple *in vivo* recombination-based cloning strategy and a single PCR amplicon that is cloned instantly and seamlessly into all of the pSol expression

Mark Maffitt, Michele Auldridge, Saurabh Sen, Sally Floyd, Amanda Krerowicz, Marie Uphoff, Jennifer Thompson, David Mead & Eric Steinmetz

Lucigen Corporation, Middleton, Wisconsin, USA. Correspondence should be addressed to E.S. (esteinmetz@lucigen.com) or M.U. (muphoff@lucigen.com).

vectors[9] (**Fig. 2**). The process does not require DNA purification or treatment with restriction enzymes, ligase or recombinase. Following PCR amplification of the open reading frame, 1 μl of crude PCR product and 2 μl of preprocessed vector are added to tubes containing chemically competent E. cloni™ 10G cells and transformed via standard procedures. The cloning process is directional, seamless and 90% efficient. Transformed colonies are screened by PCR before expression of the fusion proteins is induced with rhamnose. The entire workflow can be completed in 4 d.

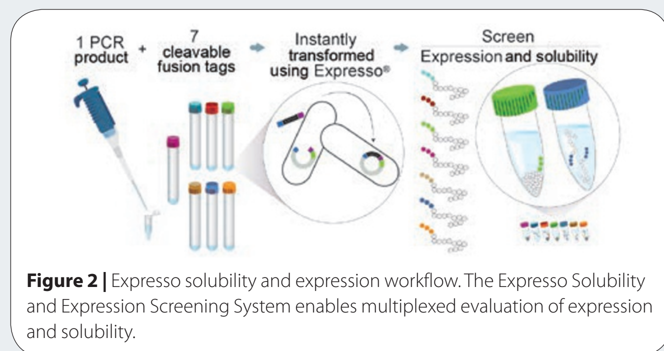

Figure 2 | Expresso solubility and expression workflow. The Expresso Solubility and Expression Screening System enables multiplexed evaluation of expression and solubility.

Enhanced expression and solubility of GH1

The pituitary form of human growth hormone (GH1) is an example of an important protein known for insoluble overexpression in *E. coli*[10]. We subjected GH1 to a fusion-partner optimization screen using the Expresso Solubility and Expression Screening System. The GH1-coding region was first amplified with primers that added 18 base pairs of vector-homologous flanking sequence to each end of the amplicon. Because empty-vector background with the pSol vectors is typically very low (<5%), minimal screening is necessary. Analysis of two candidates for each fusion type by colony PCR verified the presence of the GH1 insert in all cases (data not shown).

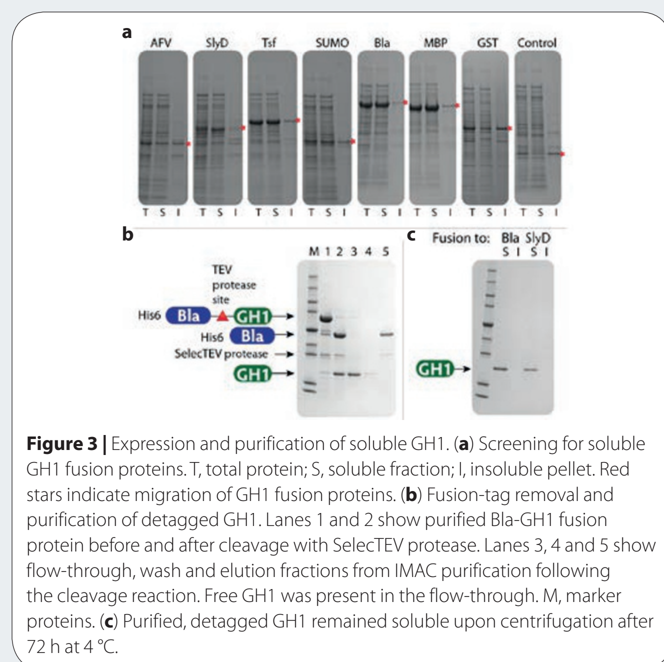

Figure 3 | Expression and purification of soluble GH1. (**a**) Screening for soluble GH1 fusion proteins. T, total protein; S, soluble fraction; I, insoluble pellet. Red stars indicate migration of GH1 fusion proteins. (**b**) Fusion-tag removal and purification of detagged GH1. Lanes 1 and 2 show purified Bla-GH1 fusion protein before and after cleavage with SelecTEV protease. Lanes 3, 4 and 5 show flow-through, wash and elution fractions from IMAC purification following the cleavage reaction. Free GH1 was present in the flow-through. M, marker proteins. (**c**) Purified, detagged GH1 remained soluble upon centrifugation after 72 h at 4 °C.

Protein expression was tested using a convenient autoinduction procedure. Cultivation in LB-Miller medium supplemented with 0.2% rhamnose and 0.05% glucose enabled growth without induction until the glucose had been depleted. As expected, when GH1 was expressed with only an N-terminal His tag, the bulk of the protein partitioned to the insoluble fraction after cell lysis (**Fig. 3**). Marked increases in expression and solubility were observed when GH1 was fused to Tsf, Bla and MBP, but smaller improvements were observed with AFV, SlyD, SUMO and GST. Overexpressed GH1 was purified to homogeneity and remained soluble, even after removal of the fusion tag (**Fig. 3**). Examinations of other target-protein fusions have identified distinct sets of effective fusion partners that generate similar enhanced expression and solubility, but with different partners.

Conclusion

Target proteins respond uniquely and unpredictably to fusion partners, necessitating lengthy and labor-intensive efforts to find tags that promote enhanced expression and solubility. The Expresso Solubility and Expression Screening System represents a significant improvement in the process by which solubility tags are selected. Past methods of selecting fusion partners involved cumbersome cloning methods that were too frequently incompatible with insert sequences, and were thus ill suited to the testing of multiple partners for the expression of important proteins. With the Expresso System, the cloning, expression and evaluation of seven different solubility tags can be completed in a matter of days.

ACKNOWLEDGMENTS

This work was supported by a Small Business Innovation Research grant from the US National Institute of General Medical Sciences.

1. Kapust, R.B. & Waugh, D.S. *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci.* **8**, 1668–1674 (1999).
2. Smith, D.B. & Johnson, K.S. Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene* **67**, 31–40 (1988).
3. Marblestone, J.G. *et al.* Comparison of SUMO fusion technology with traditional gene fusion systems: enhanced expression and solubility with SUMO. *Protein Sci.* **15**, 182–189 (2006).
4. Han, K.Y. *et al.* Solubilization of aggregation-prone heterologous proteins by covalent fusion of stress-responsive *Escherichia coli* protein, SlyD. *Protein Eng. Des. Sel.* **20**, 543–549 (2007).
5. Tokunaga, H. *et al.* Halophilic beta-lactamase as a new solubility- and folding-enhancing tag protein: production of native human interleukin 1alpha and human neutrophil alpha-defensin. *Appl. Microbiol. Biotechnol.* **86**, 649–658 (2010).
6. Han, K.Y. *et al.* Enhanced solubility of heterologous proteins by fusion expression using stress-induced *Escherichia coli* protein, Tsf. *FEMS Microbiol. Lett.* **274**, 132–138 (2007).
7. Goulet, A. *et al.* The thermo- and acido-stable ORF-99 from the archaeal virus AFV1. *Protein Sci.* **18**, 1316–1320 (2009).
8. Giacalone, M.J. *et al.* Toxic protein expression in *Escherichia coli* using a rhamnose-based tightly regulated and tunable promoter system. *Biotechniques* **40**, 355–364 (2006).
9. Bubeck, P., Winkler, M. & Bautsch, W. Rapid cloning by homologous recombination *in vivo*. *Nucleic Acids Res.* **21**, 3601–3602 (1993).
10. Kim, M.J. *et al.* Complete solubilization and purification of recombinant human growth hormone produced in *Escherichia coli*. *PLoS One* **8**, e56168 (2013).

This article was submitted to *Nature Methods* by a commercial organization and has not been peer reviewed. *Nature Methods* takes no responsibility for the accuracy or otherwise of the information provided.

© 2015 Nature America, Inc. All rights reserved.

npg