Published in final edited form as:

Biotechnol Bioeng. 2014 December ; 111(12): 2407–2411. doi:10.1002/bit.25317.

Unrelated solubility-enhancing fusion partners MBP and NusA utilize a similar mode of action

Sreejith Raran-Kurussi and David S. Waugh*

Protein Engineering Section, Macromolecular Crystallography Laboratory, Center for Cancer Research, National Cancer Institute, Frederick, MD 21702, USA

Abstract

The tendency of recombinant proteins to accumulate in the form of insoluble aggregates in *Escherichia coli* is a major hindrance to their overproduction. One of the more effective approaches to circumvent this problem is to use translation fusion partners (solubility-enhancers, SEs). *E. coli* maltose binding protein (MBP) and N-utilization substance A (NusA) are arguably the most effective solubilizing agents that have been discovered so far. Here, we show that although these two proteins are structurally, functionally, and physiochemically distinct, they influence the solubility and folding of their fusion partners in a very similar manner. These SEs act as “holdases” that prevent the aggregation of their fusion partners. Subsequent folding of the passenger proteins, when it occurs, is either spontaneous or chaperone-mediated.

Keywords

chaperone; fusion partners; inclusion bodies; MBP; NusA; solubility enhancers

Introduction

The ability to produce large quantities of soluble recombinant proteins is crucial for structural and functional studies. Over-expression in *Escherichia coli* is still the most common means of accomplishing this, mainly due to the ease of genetic manipulation, low cost, and scalability. However, the frequent formation of insoluble aggregates (inclusion bodies) in *E. coli* has forced researchers to look for ways to circumvent this bottleneck. One of the most successful strategies to improve the solubility of recombinant proteins is to fuse them to certain highly soluble partners. *E. coli* maltose-binding protein (MBP) and N-utilization substance A (NusA) are among the most commonly used and effective solubility enhancers (SEs) (Hammarstrom et al. 2002; Korf et al. 2005; Shih et al. 2002). The mechanism of solubility enhancement by MBP has been studied in some detail (Douette et al. 2005; Fox et al. 2001; Fox et al. 2003; Nallamsetty and Waugh 2006; Nallamsetty and Waugh 2007; Raran-Kurussi and Waugh 2012). However, similar studies have yet to be conducted with NusA. Consequently, it is unknown whether or not these two unrelated

*Corresponding author: David S. Waugh, Protein Engineering section, MCL, CCR, National Cancer Institute, Frederick, MD 21702, USA, Tel: 301-846-1842, Fax: 301-846-7148, waughd@mail.nih.gov.

The authors declare they have no conflict of interest.

proteins utilize a similar mechanism of action to enhance the solubility of their fusion partners. In the present report, we show that the impact of NusA on the solubility and folding of its fusion partners, both *in vitro* and *in vivo*, closely parallels that of MBP, implying a common mechanism of action.

Materials and Methods

Materials

All chemicals, in the highest available purity, were purchased from Sigma-Aldrich, American Bioanalytical, Merck-Chemicals, or Roche, unless otherwise stated.

Construction of expression vectors

The His₆-NusA fusion protein expression vectors were constructed by Gateway cloning (Life Technologies Inc., Carlsbad, CA), using the pDEST-544 destination vector (Addgene plasmid 11519, Protein Expression Laboratory, Leidos Biomedical Research, Inc., Frederick, MD, USA). The construction of Gateway entry clones for the five passenger proteins used in this study (GFP, G3PDH, DHFR, DUSP14 and TEV protease) was described previously (Raran-Kurussi and Waugh 2012). These entry clones were used in standard recombination reactions (LR) with pDEST-544 to generate the His₆-NusA fusion vectors, according to the manufacturer's instructions (Life Technologies). The resulting protein expression vectors included a canonical TEV protease recognition site (ENLYFQG) between the tags and the passengers, except for the vector encoding the TEV protease fusion, which contained an uncleavable recognition site (ENLYFQP) instead (Kapust et al. 2002). The nucleotide sequences of all entry clones were confirmed experimentally.

Expression of fusion proteins

E. coli BL21-CodonPlus (DE3)-RIL cells (Agilent Technologies Inc., Santa Clara, CA) were used for all expression experiments. Cells containing the protein expression plasmids were grown to mid-log phase (OD₆₀₀ of ~0.5) at 37 °C in Luria broth containing 100 µg ml⁻¹ ampicillin and 30 µg ml⁻¹ chloramphenicol. Overproduction of fusion protein was induced with IPTG at a final concentration of 1 mM for 4 h at 37 °C. Measurements of protein expression and solubility were performed essentially as described previously (Kapust and Waugh 1999).

Purification of proteins, refolding and activity assay

The purification of His₆-NusA fusion proteins under native and denaturing conditions, refolding, and activity assays were performed as described previously without any modifications (Raran-Kurussi and Waugh 2012). The concentration and total yield of the refolded fusion proteins were determined spectrophotometrically on the basis of their absorbance at 280 nm (A_{280nm}) and calculated extinction coefficients using the analysis software that is available through the ExPASy web tool (Gasteiger et al. 2003). Additionally, because some contaminants were present after refolding (most likely truncated forms of the fusion proteins), the yield of the full-length fusion proteins was calibrated by densitometric scanning of those bands on an SDS gel vs. the truncated contaminants, which were assumed to be inactive. All of the passenger proteins except TEV protease were also

assayed for activity after proteolytic removal of the His₆-NusA fusion tag. A 5:1 molar ratio of fusion protein to TEV protease was used to remove the tags with an incubation time of either 1 h at 30 °C or overnight at 4 °C.

Results

Fusion protein design

The His₆-NusA fusion proteins were designed to be as similar as possible to the His₆-MBP fusion proteins that were characterized in a previous study (Raran-Kurussi and Waugh 2012). The same set of five aggregation-prone passenger proteins was fused to His₆-NusA: green fluorescent protein (GFP), glyceraldehyde 3-phosphate dehydrogenase (G3PDH), dihydrofolate reductase (DHFR), dual-specificity phosphatase 14 (DUSP14) and tobacco etch virus (TEV) protease. All five of these proteins are poorly soluble when expressed as His₆- or His₆-GST fusion proteins in *E. coli*, and they all have enzymatic activity (or fluorescence in the case of GFP) that can be used as a quantitative indicator of proper folding (Raran-Kurussi and Waugh 2012).

The N-terminal polyhistidine (His₆) tag enabled the fusion proteins to be purified by immobilized metal affinity chromatography (IMAC) under both native and denaturing conditions. A TEV protease recognition site (ENLYFQG) was located immediately adjacent to the N-termini of the passenger proteins so that the N-terminal tags could be removed if desired. An uncleavable variant of the canonical TEV protease site (ENLYFQP) was used instead in the TEV protease fusion proteins to prevent their autodigestion (Kapust et al. 2002).

In vitro refolding

Refolding of His₆-NusA fusion proteins purified under denaturing conditions typically yielded a high proportion of soluble protein (Fig. 1), ranging between 50–90%, which is comparable to the results obtained previously upon refolding of their His₆-MBP counterparts and in stark contrast to the relatively poor yield of soluble His₆- and His₆-GST fusion proteins (Raran-Kurussi and Waugh 2012). Hence, like MBP, the ability to enhance the solubility of its fusion partners is an intrinsic property of NusA; no extraneous factors are required to achieve this effect.

To ascertain what fraction of the passenger proteins in the population of soluble His₆-NusA fusion proteins was properly folded, we first attempted to assay the activity of the fusion proteins. All of them exhibited less activity than would be expected if 100% of the passenger proteins were folded (data not shown). A substantial amount of protein precipitated when the fusions were cleaved by TEV protease, suggesting the presence of soluble aggregates in the population of refolded fusion proteins, a phenomenon that has been noted previously (Nomine et al. 2001a) and which was also observed when the same passenger proteins were fused to MBP (Raran-Kurussi and Waugh 2012). After removing the precipitated material by high-speed centrifugation, the soluble fractions were assayed for activity, or fluorescence in the case of GFP. These measurements were calibrated/normalized using “standards” of known concentration purchased from commercial sources

(GFP, DHFR, and G3PDH) or crystallization-grade preparations produced in-house (DUSP14 and TEV protease). Assuming that 100% of the protein in the “standard” samples was folded and active, we could estimate the percent moles of active protein recovered per mole of starting material (Fig. 2). In the case of TEV protease, however, the fusion protein could not be cleaved (see Materials and Methods). Therefore, the His₆-NusA-TEV fusion protein rather than the cleaved protease was assayed for activity. As a result, soluble aggregates were not removed before the activity of refolded TEV protease was measured. The final yield of the refolded proteins is listed in Table. 1.

The activity assays confirmed that only a fraction of the passenger proteins were properly refolded, and therefore that the soluble protein obtained after refolding must have contained some improperly folded, soluble aggregates. The percentage of soluble aggregates varied, depending on the passenger protein. For example, by comparing Figs. 1 and 2, one can see that approximately 60–70% of the His₆-NusA-GFP fusion protein was soluble after renaturation, but only 30% of the total GFP (the equivalent of about half the soluble fusion protein) was properly folded. Further, about 50% of the His₆-NusA-DHFR fusion protein was recovered in a soluble form after refolding (Fig. 1), but practically no DHFR activity was detected after the soluble fusion protein was cleaved and the precipitate removed, indicating that virtually all of the His₆-NusA was fused to improperly folded DHFR. Remarkably, in terms of the recovery of properly refolded passenger proteins, the results obtained with His₆-NusA closely parallel those obtained previously with His₆-MBP (Fig. 2).

Because the His₆-tag does not have any solubility-enhancing characteristics (Hammarstrom et al. 2002; Raran-Kurussi and Waugh 2012; Woestenenk et al. 2004), the yield and activity of the His₆ fusion proteins can be taken to approximate the amount of spontaneous folding that occurs. For two of the His₆-tagged passenger proteins (GFP and TEV protease), a significant amount of soluble (and therefore presumably properly folded) material was recovered after refolding. Note that, on a mole per mole basis, approximately as much activity is associated with these two His₆-tagged passengers as was obtained from the corresponding His₆-MBP and His₆-NusA fusions (Fig. 2). In the case of G3PDH and DHFR, little or no active protein was recovered after refolding of the His₆-MBP and His₆-NusA fusions and little to no soluble His₆-tagged protein was obtained after renaturation (Figs. 1, 2). For all four of these passenger proteins, there is rough agreement between the amount of His₆-tagged protein that remained soluble after renaturation and the amount of properly folded passenger protein that was present in the population of soluble His₆-MBP and His₆-NusA fusion proteins after refolding. Hence, MBP and NusA increase the solubility of these four passenger proteins but neither solubility enhancer exerts a positive impact on their folding *in vitro*.

DUSP14, on the other hand, which exhibited little solubility and no enzymatic activity after being refolded as a His₆-fusion, was highly soluble and active after renaturation as either a His₆-MBP or His₆-NusA fusion protein. It seems that both solubility enhancing fusion partners play a more active role in the folding of DUSP14 than the other passenger proteins. The reason for this is unknown, but merits further investigation. It is noteworthy, however, that even in this unusual situation, the behavior of MBP and NusA is very similar.

In vivo folding

The His₆-NusA fusion proteins were also purified under native conditions, cleaved by TEV protease, and centrifuged to remove any precipitate that may have formed. When the soluble material was assayed for activity, all of the samples exhibited a high degree of enzymatic activity (Fig. 3). As before, for technical reasons (see Materials and Methods), the His₆-NusA-TEV fusion protein was not cleaved prior to measuring its activity. The relative activity of GFP, TEV protease and DUSP14 was similar irrespective of which solubilizing fusion partner was utilized and which method was used to generate the samples (native purification or refolding). Interestingly, like their His₆-MBP counterparts, the His₆-NusA-DHFR and His₆-NusA-G3PDH fusion proteins were far more active when they were purified under native conditions than when they were refolded (Fig. 3). This suggests that additional factor(s) in *E. coli* cells contribute to the folding of these passenger proteins *in vivo*.

Discussion

In the present study, we investigated the ability of the solubility-enhancing fusion partner NusA to promote the solubility and folding of aggregation-prone proteins after renaturation *in vitro* as well as after purification under native conditions. We find that the results obtained with five diverse passenger proteins are very similar in all respects to those obtained previously when MBP was used as the solubility-enhancing fusion partner (Raran-Kurussi and Waugh 2012). This is surprising because these two SEs are so structurally, functionally, and physiochemically dissimilar. Renaturation experiments revealed that the ability to enhance solubility is an intrinsic property of NusA and MBP; no other factors are required to mediate this effect. Moreover, like MBP, NusA appears to play a passive role in the folding of its fusion partners, acting mainly as a “holdase” to keep passenger proteins in a soluble form (i.e. to prevent them from forming irreversible aggregates). Except for DUSP14, once they were rendered soluble by fusion to NusA or MBP, passenger proteins either folded spontaneously (GFP, TEV) or not at all (DHFR, G3PDH) *in vitro*. The refolding of DUSP14 appeared to be actively facilitated by NusA and MBP.

The situation is more complicated *in vivo* (*in bacterio*). Some passenger proteins fold spontaneously, as they do *in vitro*, while the folding of others appears to depend upon endogenous factors. Examples of the latter type include DHFR and G3PDH. Virtually no active protein could be recovered after renaturation *in vitro*, but substantial DHFR and G3PDH activity was evident after purification under native conditions. What endogenous factor(s) might promote the proper folding of this class of passenger proteins? Experiments with MBP fusion proteins suggest that it is likely to be the GroESL chaperonin. A mutant form of the chaperonin that promotes the folding of GFP much more effectively than does wild-type GroESL was obtained by directed evolution (Wang et al. 2002). The mutant chaperonin dramatically enhanced the fluorescence of His₆-MBP-GFP and His₆-NusA-GFP fusion proteins, indicating that GroESL is capable of promoting the folding of GFP in the context of rather large fusion proteins (Raran-Kurussi and Waugh 2012). Hence, it is feasible that GroESL participates in the folding of the DHFR and G3PDH fusion proteins as well *in vivo*.

Exactly why only certain highly soluble proteins, like NusA and MBP, are effective at enhancing the solubility (or inhibiting the aggregation) of their fusion partners remains a mystery, yet the results reported here strongly suggest a similar mechanism of action. One property that both of these highly effective yet dissimilar solubility enhancers have in common is their participation in multiple protein-protein interactions. NusA interacts with several proteins involved in transcription and DNA repair (Cohen and Walker 2011), and MBP forms a complex with the multiprotein membrane-spanning maltose transport apparatus (Shilton 2008). The propensity of both solubility enhancers to interact with a variety of proteins may enable them to transiently bind and sequester aggregation-prone folding intermediates of their passenger proteins, thereby preventing their self-association and aggregation, as has been suggested previously (Kapust and Waugh 1999).

The most important findings reported here are, first, that the highly effective solubility enhancing proteins NusA and MBP work in essentially the same manner, despite vast differences between their structures, functions, and physiochemical properties. Moreover, they exert their effect at an early stage in the process, by inhibiting the aggregation of their fusion partners. For the most part, neither solubility enhancer appears to play an active role in the folding of its passenger proteins, when folding occurs at all. In this respect, MBP and NusA are not mimicking the active role played by the prodomains of some proteases in the folding of their zymogens (Bryan 2002). Thus far, the lone exception to this trend is DUSP14. It will be of interest to investigate in greater detail how NusA and MBP facilitate the folding of this protein.

It is well known that many passengers precipitate after removal of the solubility-enhancing tag (Nomine et al. 2001b), presumably because they are incapable of spontaneous folding or chaperone-mediated folding in *E. coli*. Future progress in this area may come from combining the anti-aggregation properties of MBP or NusA with co-expression of additional proteins to facilitate the folding of recalcitrant passenger proteins. Recent examples include the co-expression of oxidoreductases with MBP fusions to obtain folded passenger proteins with disulfide bonds (Austin and Waugh 2012; Nguyen et al. 2011). Co-expression with eukaryotic molecular chaperones such as Rot1 and BiP may also prove effective.

Acknowledgments

This research was supported by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research. We thank the staff of the Biophysics Resource in the Structural Biophysics Laboratory, National Cancer Institute, Frederick, MD, for assistance with spectrofluorometry measurements. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does the mention of trade names, commercial products or organizations imply endorsement by the US Government.

References

- Austin BP, Waugh DS. Isolation of *Metarhizium anisopliae* carboxypeptidase A with native disulfide bonds from the cytosol of *Escherichia coli* BL21(DE3). *Protein Expr Purif*. 2012; 82(1):116–124. [PubMed: 22197595]
- Bryan PN. Prodomains and protein folding catalysis. *Chemical reviews*. 2002; 102(12):4805–4816. [PubMed: 12475207]
- Cohen SE, Walker GC. New discoveries linking transcription to DNA repair and damage tolerance pathways. *Transcription*. 2011; 2(1):37–40. [PubMed: 21326909]

- Douette P, Navet R, Gerkens P, Galleni M, Levy D, Sluse FE. Escherichia coli fusion carrier proteins act as solubilizing agents for recombinant uncoupling protein 1 through interactions with GroEL. *Biochem Biophys Res Commun.* 2005; 333(3):686–693. [PubMed: 15961060]
- Fox JD, Kapust RB, Waugh DS. Single amino acid substitutions on the surface of Escherichia coli maltose-binding protein can have a profound impact on the solubility of fusion proteins. *Protein Sci.* 2001; 10(3):622–630. [PubMed: 11344330]
- Fox JD, Routzahn KM, Bucher MH, Waugh DS. Maltodextrin-binding proteins from diverse bacteria and archaea are potent solubility enhancers. *FEBS Lett.* 2003; 537(1–3):53–57. [PubMed: 12606030]
- Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A. ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* 2003; 31(13):3784–3788. [PubMed: 12824418]
- Hammarstrom M, Hellgren N, van Den Berg S, Berglund H, Hard T. Rapid screening for improved solubility of small human proteins produced as fusion proteins in Escherichia coli. *Protein Sci.* 2002; 11(2):313–321. [PubMed: 11790841]
- Kapust RB, Tozser J, Copeland TD, Waugh DS. The P1' specificity of tobacco etch virus protease. *Biochem Biophys Res Commun.* 2002; 294(5):949–955. [PubMed: 12074568]
- 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(8):1668–1674. [PubMed: 10452611]
- Korf U, Kohl T, van der Zandt H, Zahn R, Schleeger S, Ueberle B, Wandschneider S, Bechtel S, Schnolzer M, Otteleben H, et al. Large-scale protein expression for proteome research. *Proteomics.* 2005; 5(14):3571–3580. [PubMed: 16127724]
- Nallamsetty S, Waugh DS. Solubility-enhancing proteins MBP and NusA play a passive role in the folding of their fusion partners. *Protein Expr Purif.* 2006; 45(1):175–182. [PubMed: 16168669]
- Nallamsetty S, Waugh DS. Mutations that alter the equilibrium between open and closed conformations of Escherichia coli maltose-binding protein impede its ability to enhance the solubility of passenger proteins. *Biochem Biophys Res Commun.* 2007; 364(3):639–644. [PubMed: 17964542]
- Nguyen VD, Hatahet F, Salo KE, Enlund E, Zhang C, Ruddock LW. Pre-expression of a sulfhydryl oxidase significantly increases the yields of eukaryotic disulfide bond containing proteins expressed in the cytoplasm of E.coli. *Microbial cell factories.* 2011; 10:1. [PubMed: 21211066]
- Nomine Y, Ristriani T, Laurent C, Lefevre JF, Weiss E, Trave G. Formation of soluble inclusion bodies by hpv e6 oncoprotein fused to maltose-binding protein. *Protein Expr Purif.* 2001a; 23(1):22–32. [PubMed: 11570842]
- Nomine Y, Ristriani T, Laurent C, Lefevre JF, Weiss E, Trave G. A strategy for optimizing the monodispersity of fusion proteins: application to purification of recombinant HPV E6 oncoprotein. *Protein Eng.* 2001b; 14(4):297–305. [PubMed: 11391022]
- Raran-Kurussi S, Waugh DS. The ability to enhance the solubility of its fusion partners is an intrinsic property of maltose-binding protein but their folding is either spontaneous or chaperone-mediated. *PLoS One.* 2012; 7(11):e49589. [PubMed: 23166722]
- Shih YP, Kung WM, Chen JC, Yeh CH, Wang AH, Wang TF. High-throughput screening of soluble recombinant proteins. *Protein Sci.* 2002; 11(7):1714–1719. [PubMed: 12070324]
- Shilton BH. The dynamics of the MBP-MalFGK(2) interaction: a prototype for binding protein dependent ABC-transporter systems. *Biochimica et biophysica acta.* 2008; 1778(9):1772–1780. [PubMed: 17950243]
- Wang JD, Herman C, Tipton KA, Gross CA, Weissman JS. Directed evolution of substrate-optimized GroEL/S chaperonins. *Cell.* 2002; 111(7):1027–1039. [PubMed: 12507429]
- Woestenenk EA, Hammarstrom M, van den Berg S, Hard T, Berglund H. His tag effect on solubility of human proteins produced in Escherichia coli: a comparison between four expression vectors. *Journal of structural and functional genomics.* 2004; 5(3):217–229. [PubMed: 15503425]

Figure 1. Yield of soluble fusion proteins after refolding

The yield of soluble fusion protein was calculated and expressed as a percentage of the total amount of protein added to the refolding reactions. His₆- and His₆-MBP- data were adapted from ref. 7.

Figure 2. Yield of active protein from refolding experiments

The yield of active protein was calculated and expressed as a percentage of the total amount of protein added to the refolding reactions. His₆- and His₆-MBP- data were adapted from ref. 7.

Figure 3. Relative specific activity of natively purified proteins

The relative activity was estimated by normalizing the values with a standard protein in each case. His₆-MBP- data were adapted from ref. 7.

Table 1

The yield of purified and refolded active protein in mg per liter of *E. coli* culture.

Passenger protein	NusA- tag	MBP- tag	His- tag
G3PDH	0.15	-	-
GFP	1.3	1.67	2.34
DHFR	0.13	0.07	-
DUSP14	0.83	1.47	-
TEV protease	1.48	2.2	1.8