

NusA Protein as a Solubility-Promoting Tag: A Review

Roger G. Harrison

School of Chemical, Biological and Materials Engineering – University of Oklahoma – Norman, OK, USA

Introduction

The NusA fusion protein system was discovered 9 years ago (Davis 1999) and was first made commercially available by Novagen in 2000. This system is based on the solubility-promoting properties of NusA when it is fused at the amino-terminus of a target protein that is otherwise insoluble when it is expressed by itself in *Escherichia coli*. Since its introduction, the NusA system has been evaluated in several high-throughput expression screens. In addition, there have been a number of reports in the literature of success in obtaining active, functional proteins after removal of the NusA tag by protease cleavage, and there have been reports of the target protein being active while still fused to NusA. The NusA system is now one of the three most widely used fusion tag systems for *E. coli* expression (Cabrita 2006). The mechanism for the solubility promotion by NusA has also been studied. The purpose of this article is to review these studies with respect to the performance of the NusA system.

Evaluations of the solubility of NusA fusion proteins with a large number of target proteins

Three studies of the solubility in *E. coli* of NusA fusion proteins that have been reported in the literature stand out because of the number and wide size range of heterologous target proteins that were studied. A summary of the results of these studies is shown in Table 1. The study by Shih et al. (2002)

is notable because target proteins from four different types of organisms were studied. In the study by Korf et al. (2005), it was found that NusA fusion proteins were more soluble when induced at 20°C than at 30°C. In addition, the expression of soluble forms of seven out of eight large proteins (with molecular weight exceeding 70 kDa) was induced using the NusA system, compared to four of eight using three other fusion tags (GST, MBP, and hexahistidine).

Another finding by Korf et al. was that the NusA fusion tag was the best option compared to the other fusion tags for eukaryotic target proteins normally residing in intracellular organelles, the plasma membrane, or on the cytoskeleton. Kohl et al. (2008) also found that the NusA tag is useful for increasing the solubility of difficult-to-express proteins such as membrane proteins, as long as the induction of protein expression is performed at 20–25°C. Consistent with the findings of Korf et al., Kohl et al. found that more NusA fusion proteins could be purified when protein expression was induced at 25°C compared to induction at 30°C or 37°C.

Cleavage of NusA fusion proteins to obtain active or correctly folded proteins

Studies on 16 target proteins that have been expressed as NusA fusions and then cleaved to obtain active or correctly folded proteins are summarized in Table 2. Most of these studies were performed with target proteins that are 20 kDa or less in size. The yields of purified target protein ranged from 1.5–100 mg per liter of culture. High yields in the range of 30–100 mg/liter were obtained for the chemokines and cytokines. Additional information of interest about the expression and purification of these proteins includes:

- **Plant phosphoenolpyruvate-carboxylate kinase (Ermolova 2003)** – The target protein was further purified after cleavage using a BDA (blue dextran-agarose) dye affinity column. The purified target protein had 50 times greater catalytic efficiency than the fusion protein.
- **Xklp3a, Tep3Ag, and E8R (De Marco 2004)** – After cleavage, His-tagged TEV and NusA were selectively removed by Ni²⁺ immobilized metal af-

Table 1. Studies of the solubility of NusA fusion proteins using a large number of target proteins.

Literature paper ^a	Number of target proteins	Target insert source (species)	Size range of target proteins	% of NusA fusion proteins soluble
Shih et al. 2002	40	yeast, mammalian, plant, insect	9–100	60
Korf et al. 2005	75	human	6–127	60 ^b
Kohl et al. 2008	96	human	1–118	44 ^c

^a A hexahistidine tag was included in the Korf et al. and Kohl et al. studies.

^b Fusion protein counted as soluble if solubility equaled or exceeded 10%.

^c Fusion protein counted as soluble if it could be purified to yield a clear band of the expected molecular weight by Coomassie staining following SDS-PAGE.

finity chromatography. The affinity tags remained tightly bound to the resin, while pure target proteins were recovered in the unbound fractions. Each of the three target proteins were monodispersed and correctly folded after purification. Following purification of the membrane-associated E8R vaccinia virus protein, the protein precipitated after the removal of NusA in Tris buffer. However, adding 0.02% dodecylmaltoside and 150 mM sodium chloride resulted in the E8R protein remaining soluble after removal of NusA.

■ Cyclomaltodextrinase (Turner 2005) –

This protein is from the α -amylase protein family. Members of this family have frequently been difficult to express in active form in *E. coli*. Incubation with enterokinase over 24 hours resulted in a steadily increasing activity until activity levels were two times higher than for the untreated fusion protein, indicating that the presence of the fusion decreased the activity of the enzyme. To remove the affinity tags, the cleaved protein was passed through a column with immobilized Cu^{2+} .

■ Eight human chemokines (Magistrelli 2005) –

All of the target proteins were expressed in the *E. coli* Origami™ B strain to enhance disulfide formation in the cytoplasm. The sequence for the AviTag™ (avidity) biotinylation site was introduced at the C-terminus of the chemokine coding sequence. Cleaved chemokines were further purified from the NusA tag and protease mixture by affinity chromatography using monomeric avidin resin. While all of the purified proteins were active in a cell chemotaxis assay, none of the fusion proteins were active.

■ Haemerythrin (Karlsen 2005) –

After cleavage, haemerythrin was further purified by gel filtration chromatog-



ROGER HARRISON is a Professor of Chemical, Biological, and Materials Engineering at the University of Oklahoma. As a chemical engineer, Harrison states that his research interests include "the application of biotechnology to solve medical problems." Towards this end, his group has focused its efforts in large part on developing methods for recombinant protein engineering, expression, and purification. The groundwork that led to development of the NusA system began in 1991 when he and a graduate

student published a new mathematical model for predicting the tendency of a recombinant protein to accumulate in inclusion bodies upon expression, based solely on amino acid sequence (Wilkinson 1991). Harrison explains, "I applied that model to *E. coli* to look for proteins that would be solubility partners for target proteins that would otherwise be insoluble." Continuation of this work was funded by NSF in 1995. "I had a really excellent Ph.D. student, Greg Davis, who devoted himself to the project," states Harrison; this work was later published in *Biotechnology and Bioengineering*. NusA was one of the top proteins identified in the screen. "We noticed it was large in comparison to other fusion tags, but we decided to test it anyway," remarks Harrison. Davis and colleagues empirically tested for its solubility-enhancing properties, including assessing the activity of a target protein (hIL-3) when fused to NusA. The system was licensed for commercial use in 2000. Harrison states that other current research projects include "the design, expression, and purification of proteins to be used to focus on diseases including cancer, obesity, and hemophilia." He has also published an award-winning textbook, *Bioseparations Science and Engineering*, which has been adopted for use at over 53 colleges and universities worldwide.

– ROBYN M. PERRIN

raphy. The α -helical structure content determined by circular dichroism for purified haemerythrin agreed closely with prediction by modeling. The purified protein was stable in a monomeric form.

■ Human IL-29 (Li 2006) –

Purer fusion protein was obtained with S-protein affinity chromatography than with immobilized Ni^{2+} affinity chroma-

tography. After cleavage of the amino-terminal NusA/His•Tag®/S•Tag™ fusion, the biotinylated thrombin was removed with streptavidin agarose. The antiviral activity of purified IL-29 was demonstrated by its protective effect on the human immortalized amnion epithelial cell line (WISH cells) against vesicular stomatitis virus (VSV).

Table 2. Cleavage of NusA fusion proteins to obtain active or correctly folded target proteins.

Reference	Target protein(s)	MW of target protein (kDa)	Protease for cleavage	Ligand immobilized for affinity purification of fusion protein	Yield of purified target protein (mg/liter of culture)
Ermolova et al. (2003)	Plant phosphoenolpyruvate-carboxylate kinase	32	Thrombin	Ni ²⁺	1.5
De Marco et al. (2004)	Xklp3A from <i>Xenopus laevis</i>	15	TEV protease	Ni ²⁺	5.0
	Tep3Ag from <i>Anopheles gambiae</i>	NR ^a	TEV protease		2.5
	E8R Vaccinia virus protein	32 ^b	TEV protease		4.0
Turner et al. (2005)	Cyclomaltodextrinase from <i>A. flavithermus</i>	69	Enterokinase	Cu ²⁺	1.6
Magistrelli et al. (2005)	Eight human chemokines	8-21	Factor Xa protease	Ni ²⁺	30-100
Karlsen et al. (2005)	Haemerythrin from <i>M. capsulatus</i>	15	TEV protease	Ni ²⁺	NR ^a
Li and He (2006)	Human IL-29 (class II cytokine)	20	Thrombin (biotinylated)	S-protein	60
Li and Huang (2007)	Human IFN-λ2 (a cytokine)	20	Enterokinase	Ni ²⁺	65

^a Not reported.

^b Molecular weight of predicted full-length protein, as reported in NCBI.

■ **Human IFN-λ2 (Li 2007)** – After cleavage, the recombinant enterokinase was removed using Novagen® EKapture™ Agarose. Purified IFN-λ2 was able to protect WISH cells from a virus-induced cytopathic effect when applied 24 hours before adding VSV.

NusA fusion proteins for which the target protein is active

In addition to the NusA fusion proteins that have been cleaved to remove NusA in order to obtain active purified target protein, there have been a number of reports of NusA fusions in which the target protein retained activity as a NusA:target protein fusion. Such examples include ScFv catalytic antibody 14D9 (Zheng 2003), green fluorescent protein from *Aequorea victoria* (Nallamsetty 2006), human dihydrofolate reductase (Nallamsetty 2006), arginase kinase from the razor clam *Ensis directus* (Compaan 2003), modified δ-endotoxin from *B. thuringiensis* (Kumar 2005), human BCMA transmembrane receptor (Guan 2006), plant α-dioxygenase 1 (Liu 2006), and β-ketoacyl-acyl carrier protein synthase from *Plasmodium falciparum* (Lack 2006). Thus, target proteins from widely different sources have been shown to be active when fused to NusA.

Possible mechanism for solubility promotion by NusA

The NusA protein has been demonstrated to be an obligatory substrate of the chaperonin GroEL *in vivo* (Houry 1999). GroEL, along with its cofactor GroES, is the only chaperone system in *E. coli* that is essential under all growth conditions. Douette et al. (2005) studied the expression yield of the soluble NusA-UCP1 (uncoupling protein 1) fusion protein, where UCP1 is a mitochondrial membrane protein. These authors found that the solubility increased in response to co-overexpression of GroEL at 16°C. This result suggested that NusA interacts with the chaperone pathway, preventing the partner protein from aggregation.

Summary

The NusA system has been proven to facilitate solubilization of many diverse target proteins in *E. coli* that are insoluble when expressed by themselves. There are numerous examples of this system being used to produce active or correctly folded target proteins after cleaving off the NusA tag, with several different proteases used for the cleavage. There are also a number of examples of target proteins being active while still linked to NusA. The success of NusA for solubilization appears likely to be, at least in part, a result of its interaction with the *E. coli* chaperone pathway. ■

References

- Cabrita, L.D., et al. 2006. *BMC Biotechnol.* 6, 12.
- Compaan, D.M. and Ellington, W.R. 2003. *J Exp. Biol.* 206, 1545.
- Davis, G.D., et al. 1999. *Biotechnol. Bioeng.* 65, 382.
- De Marco, V., et al. 2004. *Biochem. Biophys. Res. Commun.* 322, 766.
- Douette, P., et al. 2005. *Biochem. Biophys. Res. Commun.* 333, 686.
- Ermolova, N.V., et al. 2003. *Protein Expr. Purif.* 29, 123.
- Guan, Z.B., et al. 2006. *Sheng Wu Gong Cheng Xue Bao* 22, 46.
- Houry, W.A., et al. 1999. *Nature* 402, 147.
- Karlsen, O.A., et al. 2005. *FEBS J* 272, 2428.
- Kohli, T., et al. 2008. *Proteome Sci.* 6, 4.
- Korf, U., et al. 2005. *Proteomics* 5, 3571.
- Kumar, S., et al. 2005. *J Inverteb. Pathol.* 88, 83.
- Lack, G., et al. 2006. *J Biol. Chem.* 281, 9538.
- Li, M. and He, S. 2006. *J Biotechnol.* 122, 334.
- Li, M. and Huang, D. 2007. *Biotechnol. Lett.* 29, 1025.
- Liu, W., et al. 2006. *Plant Physiol. Biochem.* 44, 284.
- Magistrelli, G., et al. 2005. *Biochem. Biophys. Res. Commun.* 334, 370.
- Nallamsetty, S. and Waugh, D.S. 2006. *Prot. Expr. Purif.* 45, 175.
- Shih, Y.P., et al. 2002. *Protein Sci.* 11, 1714.
- Turner, P., et al. 2005. *Prot. Expr. Purif.* 39, 54.
- Wilkinson, D.L. and Harrison, R.G. 1991. *Biotechnol. (N.Y.)* 9, 443.
- Zheng, L., et al. 2003. *J Biochem.* 133, 577.

Product	Size	Cat. No.	Price
pET NusA Fusion System 43.1	1 system	70942-3	\$388
pET NusA Fusion System 44	1 system	71125-3	\$388
pET-50b(+) DNA	10 µg	71464-3	\$168
Gateway® Nova pET-57-DEST™ DNA	10 µg	71848-3	\$200
Gateway® Nova pET-58-DEST™ DNA	10 µg	71849-3	\$200
Recombinant Enterokinase	50 U	69066-3	\$106
Biotinylated Thrombin	50 U	69672-3	\$179
His•Bind® Resin	10 ml	69670-3	\$88
	25 ml	69670-4	\$358
	100 ml	69670-5	\$657