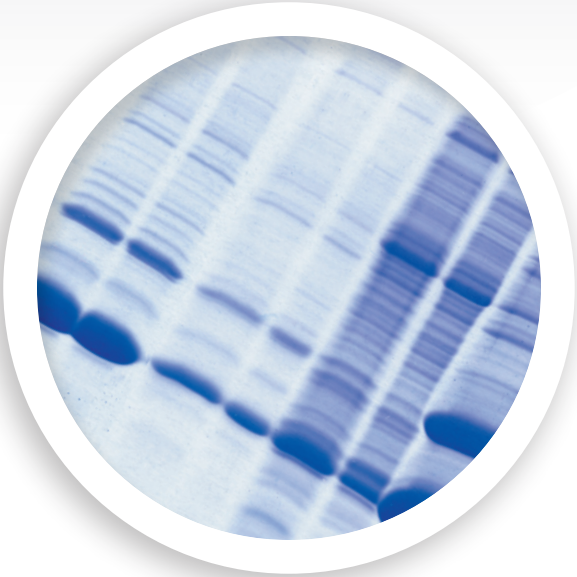
Cambridge Isotope Laboratories, Inc.
**www.isotope.com**

Eurisotop
A CAMBRIDGE ISOTOPE LABORATORIES COMPANY

APPLICATION NOTE 15

# Top Ten Tips for Producing $^{13}C$, $^{15}N$ Protein in Abundance

Deborah A. Berthold, Victoria J. Jeisy, Terry L. Sasser, John J. Shea, Heather L. Frericks, Gautam Shah and Chad M. Rienstra
Departments of Chemistry and Biochemistry
University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA



What could be easier than overexpressing an *E. coli* protein in *E. coli*? You don't have to be an old hand at protein expression to know that this can often be more difficult than it sounds. We tested our skills recently with DsbA, a 20 kDa protein that catalyzes disulfide bond formation in the *E. coli* periplasm. The wildtype DsbA expressed well in LB medium and also in a BioExpress®-supplemented $^{13}C$, $^{15}N$-labeling medium. Likewise, the DsbA C33S mutant expressed well in LB. But when we first tried to label C33S, our luck ran out – we saw no expression at all. Today we are producing $^{13}C$, $^{15}N$ DsbA C33S at a yield of 100 mg per liter. Here are our top ten tips for expressing recalcitrant proteins.

**10. Stop the leaks.** Leaky expression (i.e. expression in the absence of inducer) of a "toxic" protein or even a "less than healthful" mutant protein can slow cell growth, resulting in a suboptimal level of expression. In addition, because the half-life of ampicillin in a dense culture is less than 30 minutes, at later stages of growth there is a loss of selection for cells with ampicillin-resistant expression plasmids. So, any selection pressure stemming from leaky expression would give advantage to cells that have lost their expression plasmids. At best, the portion of plasmid-free cells in the culture at the time of induction are taking up $^{13}C$-labeled glucose without contributing to expression. At worst, leaky expression slows cell growth to the point that there is no growth at all in the labeling medium.

Promoters show variation in their leakiness, so use of a highly regulated promoter, such as the T7 promoter of the pET vector system, may be the solution. In our case, DsbA is on a high-copy plasmid (pUC119) and is expressed from a *lac* promoter (Kisigami, et al.,1995). In addition to this promoter being well-known for its leakiness, in optimizing the expression of the wildtype DsbA, we had changed *E. coli* strains. The original strain, *E. coli* M15/pREP4, contained additional copies of the *lac* repressor on pREP4; our preferred strain, *E. coli* C43(DE3), had no additional repressor other than the copy in the host genome. It was likely that the repressor binding sites on our expression plasmid titrated out all the copies of the repressor in the cell, leaving some promoters unrepressed. Our response: we put the compatible pREP4 plasmid into our *E. coli* C43(DE3) so that the cell would produce enough repressor for all

*(continued)*

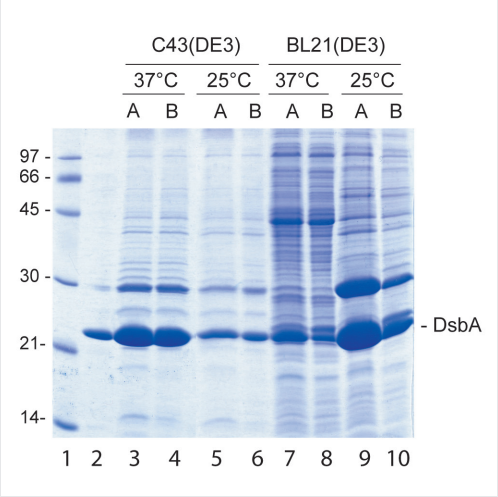



**Figure 1. Expression of DsbA C33S in supplemented Studier Medium P containing BioExpress® 1000.** DsbA C33S expression plasmid was transformed into chemically competent BL21(DE3)/pREP4 and C43(DE3)/pREP4. Several colonies were used to inoculate an overnight culture in LB containing 60 µg/ml ampicillin and 25 µg/ml kanamycin, which in turn was used to inoculate a pregrowth culture of LB. When the pregrowth culture reached an $OD_{600}$ of 0.7, the cells were harvested in a sterile manner and resuspended in ¼ volume (relative to the pregrowth culture) of supplemented Medium P (Table 1). After one hour of growth, expression was induced with 0.8 mM IPTG, and growth continued at either 25 or 37°C, as indicated. The 37°C cultures were harvested at four hours, and the 25°C at eight hours. The periplasmic proteins were obtained by resuspending the pellet from 100 mL of culture in 8 mL of 20% sucrose, 0.3 M Tris, pH 8, 0.5 mM EDTA, mixing at 4°C for five minutes, and repelleting the cells. Eight microliters of the supernatant from this centrifugation were loaded in the lanes labeled "A." The pellets were then resuspended in 4 mL water, mixed for 15 minutes, and centrifuged again. Four microliters of this supernatant were loaded in the lanes labeled "B." Lanes A and B, combined for each *E. coli* strain (C43(DE3), BL21(DE3)) and induction temperature (25°C or 37°C), are representative of the total yield for each growth condition. Lane 1, molecular mass standards (kDa); Lane 2, purified DsbA. Proteins were separated on a Tris-glycine (Laemmli) gel containing 15% acrylamide and 2.5 M urea and stained with Coomassie Blue.

the plasmid binding sites. Another solution might be to clone *lacI*$^q$ (a constitutive *lac* repressor) directly onto the expression plasmid; this is found, for example, in Qiagen's QE80 series of His-tag expression vectors.

**9. Slow down the train.** For high levels of protein expression, the rate of transcription needs to be coupled to that of translation, which in turn needs to be coupled to any essential co- or post-translational events, such as folding, cofactor binding or membrane insertion. When transcription outstrips translation, loss of cell viability can occur (along with the destruction of ribosomal RNA and induction of proteases; see Dong, et al., 1995). There are several methods to tweak the rates of cellular metabolism to try to bring transcription, translation and post-translational processing in line. One can change promoters (Makrides, 1996; Baneyx, 1999). Strong promoters, such as T7, can be replaced with weaker promoters (arabinose, T5, tac). One can change cell lines. Two *E. coli* strains particularly suited for expression optimization, C41(DE3) and C43(DE3), were originally obtained from a selection for mutations that overcame lethality associated with overexpression from a T7 promoter (Miroux and Walker, 1996). These two strains also show an increase in plasmid stability relative to their parent BL21(DE3) strain (Dumon-Seignovert, et al., 2004). And, perhaps most easily, one can change the temperature during expression. For DsbA, the interplay between growth temperature and *E. coli* strain can be seen in Figure 1. Using BL21(DE3), slowing cell processes by growing at 25°C gives a large increase in expression over growth at 37°C (lanes 9-10 vs. lanes 7-8). In contrast, the strain C43(DE3), which is thought to have a mutation slowing the rate of transcription, gives a greater yield of expressed protein at 37°C (compare lanes 3-4 vs. 5-6). For DsbA C33S, we can choose between expressing at 37°C in C43 or expressing at 25°C in BL21(DE3). For some other proteins we have seen that both C43(DE3) and a lower temperature are required for the highest expression level.

**8. Eat yer spinach.** If Popeye was a microbiologist, he might well claim that his cultures were "strong to the finish," because he feeds them iron. Studier recommends that trace metals be added to defined media, and, remarkably, that if a trace metal mixture is not available, 100 µM $FeCl_3$ alone will give the nearly the same level of expression (Studier, 2005). It is recommended that trace metals be added to rich media, as a precaution against batch-to-batch variation of trace metal content. We have found that addition

**Table 1.** Supplemented Studier Medium P, with BioExpress® 1000 adapted from Studier (2005)

| Medium P salts – pH ~7 | nitrogen and carbon sources | cofactor and antibiotic supplements | "0.2x" Studier trace metals |
|---|---|---|---|
| 50 mM $Na_2HPO_4$ | 3 g/L (56 mM) $NH_4Cl$[1] | 10 µg/mL thiamin | 10 µM $FeCl_3$ |
| 50 mM $KH_2PO_4$ | 4 g/L glucose[2] | 10 µg/mL biotin | 4 µM $CaCl_2$ |
| 5 mM $Na_2SO_4$ | 10 mL/L BioExpress® 1000[3] | 60 µg/mL ampicillin | 2 µM $MnCl_2$ |
| 2 mM $MgSO_4$ | | | 2 µM $ZnSO_4$ |
| | | | 0.4 µM $CoCl_2$ |
| | | | 0.4 µM $CuCl_2$ |
| | | | 0.4 µM $NiCl_2$ |
| | | | 0.4 µM $Na_2MoO_4$ |
| | | | 0.4 µM $Na_2SeO_3$ |
| | | | 0.4 µM $H_3BO_3$ |

**1:** For $^{15}N$ labeling: CIL No. NLM-467, ammonium chloride ($^{15}N$, 99%).

**2:** When following the method of Marley (2001) 4 g/L glucose is used, and otherwise 2 g/L glucose. For $^{13}C$ labeling: CIL No. CLM-1396, D-glucose, (U-$^{13}C_6$, 99%).

**3:** For expression tests (unlabeled): CIL No CGM-1000-U-S, BioExpress® 1000 (unlabeled) 10X concentrate. For uniform $^{13}C$, $^{15}N$ labeling: CIL No. CGM-1000-CN-S, BioExpress® 1000 (U-$^{13}C$, 98%; U-$^{15}N$, 98%), 10X concentrate.

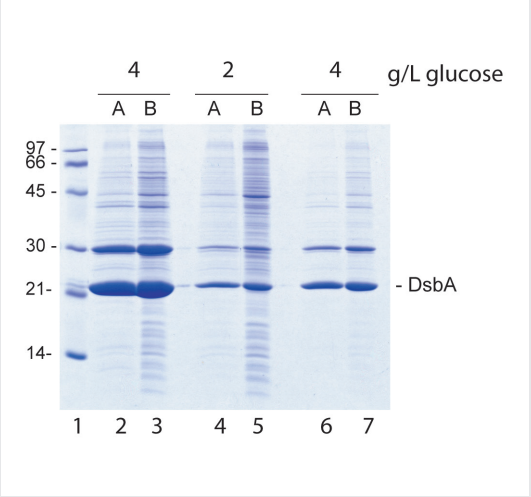



**Figure 2. Comparison of expression of DsbA C33S in supplemented Studier Medium P containing BioExpress® 1000 using two concentrations of glucose.** DsbA C33S was expressed as described in the legend to Figure 1, using *E. coli* strain BL21(DE3), at 25°C, with a harvest time of 14.5 hours post-induction. Lanes A and B are two fractions of DsbA obtained from each culture condition, and combined represent the total yield. Lanes 2, 3, 6 and 7 are from medium containing 4 g/L glucose; lanes 4 and 5 from 2 g/L glucose. Lanes 6 and 7 are a 1:5 dilution of the samples of lanes 2 and 3, respectively, demonstrating that the 4 g/L glucose culture contains at least 5x the DsbA of the culture with 2 g/L glucose. Lane 1 contains molecular mass standards (kDa). Proteins were separated on a Tris-glycine (Laemmli) gel containing 15% acrylamide and 2.5 M urea, and stained with Coomassie Blue.

of trace metals improves the yield of expressed protein even in BioExpress®-supplemented (10 mL/L) media.

**7. Be media savvy.** Media acidification is detrimental to *E. coli* growth and protein yield (Swartz, 2001), but can be prevented to some extent by using a well-buffered medium with good aeration. We are currently using Studier Medium P (a phosphate-buffered medium) for expression of DsbA and several other proteins (Table 1). In addition, it has been found that calcium, although often added in relatively high amounts in traditional *E. coli* media as M9, is actually not essential at such high levels (and, we have noted, can sometimes cause precipitant to form in the medium). On the other hand, Studier (2005) reports that using 2 mM $MgSO_4$, rather than 1 mM, can increase the cell density anywhere from 50% to five-fold, which would correspond to comparable increases in protein yield.

**6. Exercise "aerobically."** The least expensive (and perhaps also least appreciated) supplement for cultures is oxygen, in the form of increased aeration. Glucose utilized by the aerobic respiratory chain provides more than 10 times the energy of glucose fermented. For growth on glycerol, the situation is even more dramatic: *E. coli* cannot utilize glycerol as an anaerobic energy source. And, as noted above, culture acidification, as occurs under anaerobic and microaerophilic conditions, is often detrimental to protein expression. To enhance aeration, we routinely use a culture volume of only 250 mL in a baffled 2 L flask. Note, however, that all expressions cannot tolerate all volume/baffling geometries, as certain expressions may cause cells to become fragile and sensitive to shear forces. Therefore, optimal culture volume and rate of shaking with a given baffle geometry should be tested, not presumed.

**5. Timing is everything. (Part I: Induction)** A common protocol for protein expression requires that IPTG be added when a culture has reached a cell density ($A_{600}$) of 0.8. But, following induction the growth of the culture will often slow or stop entirely. Thus it could happen that while the labeling medium supports uninduced growth

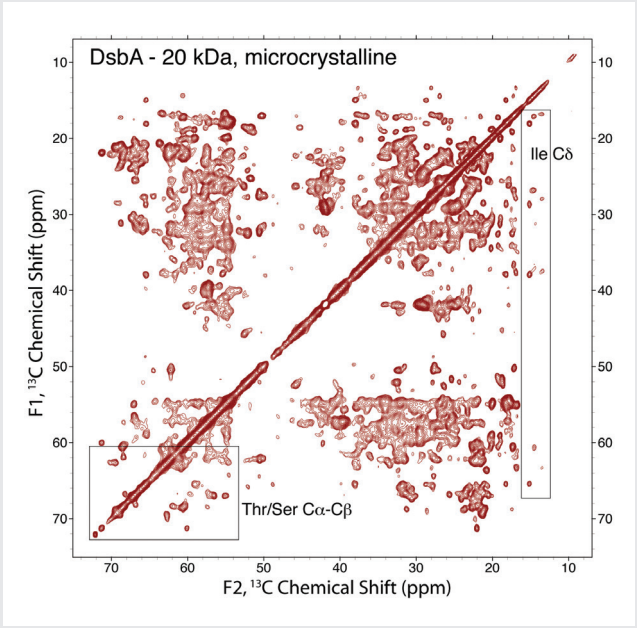



**Figure 3. Solid-state NMR C-C 2D spectrum of microcrystalline DsbA acquired at 750 MHz with 100 ms DARR mixing.** The aliphatic region of the spectrum is shown, with the particularly well-resolved isoleucine delta carbon correlations and serine and threonine Cα-Cβ correlations outlined in boxes. The sample is U-$^{13}C$, $^{15}N$-labeled DsbA (5 mg) in a 3.2 mm standard Varian rotor. Data was acquired for 18 hours at -10°C, utilizing 12.5 kHz magic-angle spinning and high-power proton decoupling.

to cell density of $A_{600}$ = 4.0, the density of an induced culture ends up at harvest at only $A_{600}$ = 1.0, giving perhaps only one quarter the potential yield. A better "rule of thumb" for time of induction would be to induce expression at 50% the density of an uninduced culture at harvest, and to also test the effect of induction at 80% uninduced density (i.e. at $A_{600}$ = 3.2 for a culture that could grow to 4.0).

*(continued)*

**4. Timing is everything. (Part II: Harvest)** A typical protocol for a 37°C expression calls for harvest of *E. coli* cells two-three hours after protein expression is induced. However, it can be worthwhile to determine where expression is maximal, especially if one is using temperatures lower than 37°C. We now harvest our 25°C DsbA expression at 20 hours post-induction, and this more than doubles the protein yield of a six-hour harvest.

**3. Quadruple up on the cells.** Marley, et al. (2001) have reported increases in expression of labeled protein by growing an *E. coli* culture in LB medium to density of $A_{600}$ = 0.8, followed by harvest and resuspension at four-fold the density in labeling medium. Because cells containing the DsbA C33S expression plasmid grew well in rich medium, but poorly in labeling medium, this was an ideal solution for us. The four expression conditions of Figure 1 all used this method.

**2. Just a spoonful of sugar (or four).** Marley, et al. (2001) also reported using 4 g/L glucose in their dense cell cultures. We tested 2 g/L glucose and 4 g/L glucose and found an increase of more than five-fold in yield of DsbA C33S with a two-fold increase in glucose (Figure 2). Additional glucose (6 and 8 g/L) did further increase the yield, but not cost-effectively. At 4 g/L glucose, increasing the $NH_4Cl$ from 2 g/L to 3 g/L gave a corresponding ~50% increase in expressed protein, but an increase of $NH_4Cl$ above 3 g/L had no further effect (data not shown).

**1. Save both the baby *and* the bathwater.** A common protocol for extracting periplasmic proteins such as DsbA directs that the cells be suspended in a hyperosmotic medium (we use 20% sucrose), followed by resuspension in water to burst the outer membrane, releasing the periplasmic protein. We found our DsbA overexpressing cells to be so fragile that breakage occurred even in the sucrose solution (Figure 1, lanes A), and had we not saved this supernatant, we would have lost a substantial part of our yield (compare Figure 1, lanes B). Another occasion where it can be useful to save and assay all fractions is when assessing growth conditions and expression strains. A low-speed centrifugation of a well-sonicated *E. coli* sample will pellet inclusion bodies. If one is attempting to optimize for a high yield of protein in a native conformation, a large amount of protein in an inclusion body fraction can indicate that expression needs to be slowed somehow (see Tip #9 above) to allow time for proper folding or post-translational events.

By applying these tips to our DsbA C33S expression, we were able to produce sufficient $^{13}C$, $^{15}N$-labeled protein for solid-state NMR studies. Figure 3 shows a highly resolved 100 ms DARR spectrum of 5 mg microcrystalline DsbA taken on a 750 MHz spectrometer (Franks, et al. (2005)). We have found these tips to be generally useful: we have not only improved the expression of the *E. coli* periplasmic protein DsbA, but have been able to increase the expression yield of membrane proteins (e.g. *E. coli* cytochrome $bo_3$ oxidase (Frericks, et al. (2006) and DsbB (Li, et al. (2007)), as well as heterologous (non-*E. coli*) proteins expressed in *E. coli*.

## References

Baneyx, F. **1999**. Recombinant protein expression in *Escherichia coli*. *Curr Opin Biotech, 10*, 411-421.

Dong, H.; Nilsson, L.; Kurland, C.G. **1995**. Gratuitous overexpression of genes in *Escherichia coli* leads to growth inhibition and ribosome destruction. *J Bacteriol, 177*, 1497-1504.

Dumon-Seignovert, L.; Cariot, G.; Vuillard, L. **2004**. The toxicity of recombinant proteins in *Escherichia coli*: a comparison of overexpression in BL21(DE3), C41(DE3), and C43(DE3). *Protein Expr Purif, 37*, 203-206.

Frericks, H. L.; Zhou, D.H.; Yap, L.L.; Gennis, R.B.; Rienstra, C.M. **2006**. Magic-angle spinning solid-state NMR of a 144 kDa membrane protein complex: *E. coli* cytochrome bo3 oxidase. *J Biomol NMR, 36*, 55-71.

Franks, W.T.; Zhou, D.H.; Wylie, B.J.; Money, B.G.; Graesser, D.T.; Frericks, H.L.; Sahota, G.; Rienstra, C.M. **2005**. Magic-angle spinning solid-state NMR spectroscopy of the beta-1 immunoglobulin binding domain of protein G (GB1): $^{15}N$ and $^{13}C$ chemical shift assignments and conformational analysis. *J Am Chem Soc, 127*, 12291-12305.

Kishigami, S.; Kanaya, E.; Kikuchi, M.; Ito, K. **1995**. DsbA-DsbB interaction through their active site cysteines. Evidence from an odd cysteine mutant of DsbA. *J Biol Chem, 270*, 17072-17074.

Li, Y.; Berthold, D.A.; Frericks, H.L.; Gennis, R.B.; Rienstra, C. **2007**. Partial $^{13}C$ and $^{15}N$ chemical-shift assignments of the disulfide-bond-forming enzyme DsbB by 3D magic-angle spinning NMR spectroscopy. *Chembiochem, 8*, 434-442.

Makrides, S.C. **1996**. Strategies for achieving high-level expression of genes in *Escherichia coli*. *Microbiol Rev, 60*, 512-538.

Marley, J.; Lu, M.; Bracken, C. **2001**. A method for efficient isotopic labeling of recombinant proteins. *J Biomol NMR, 20*, 71-75.

Miroux, B.; Walker, J.E. **1996**. Over-production of proteins in *Escherichia coli*: Mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J Mol Biol, 260*, 289-298.

Studier, F.W. **2005**. Protein production by auto-induction in high-density shaking cultures. *Protein Expr Purif, 41*, 207-234.

Swartz, J.R. **2001**. Advances in *Escherichia coli* production of therapeutic proteins. *Curr Opin Biotech, 12*, 195-201.