Short Technical Reports

SHORT TECHNICAL REPORTS

Manuscripts published in the Short Technical Reports section are shorter and less comprehensive in scope than full Research Reports.

Construct for High-Level Expression and Low Misincorporation of Lysine for Arginine During Expression of pET-Encoded Eukaryotic Proteins in *Escherichia coli*

BioTechniques 27:950-954 (November 1999)

ABSTRACT

The arginine codon AGA is rarely used in *E. coli* but is common in eukaryotic genes. Prior studies have shown that the low level of tRNA^{Arg}_{UCA} can lead to low expression and misincorporation of lysine for arginine, during expression of genes containing AGA codons in *E. coli*. The chloramphenicol-selectable plasmid pJY2 is designed to facilitate the expression of such genes cloned into pET vectors: it encodes T7 lysozyme (to depress constitutive expression of the cloned gene) and tRNA^{Arg}_{UCA} (to suppress lysine misincorporation at AGA codons). Using pJY2, we observed robust and translationally faithful expression of mutant ubiquitin genes in which 14% (11 out of 76) of the total codons were AGA. Competent BL21(DE3)pJY2 cells can be used to suppress lysine misincorporation

and achieve high-level expression of pET-encoded target genes without modification of AGA codons in the target gene sequence.

INTRODUCTION

The arginine codon AGA occurs commonly in eukaryotic genes but is extremely rare in *Escherichia coli* (7). When eukaryotic genes harboring AGA codons are expressed in *E. coli*, this discordance in codon usage can lead to (i) low expression level and (ii) misincorporation of lysine for arginine (3–5,7,9, 15,17). That these problems arise from a low level of tRNA^{Arg}_{UCA} is indicated by the restoration of translational fidelity and robust expression on the substitution of AGA with a different arginine codon (5,7,9). In some cases, it has been possible to achieve the same effect by using a helper plasmid carrying the argU gene (also known as dnaY), which encodes tRNA^{Arg}_{UCA} (3–5,7,9,16,17).

We are interested in ubiquitin conjugation reactions. Conjugation involves the formation of an isopeptide bond between the carboxyl terminus of ubiquitin and a lysine residue of a target protein (12,14). A favored in vitro strategy for identifying ubiquitin conjugation sites involves the mutation of candidate lysines to arginines. This strategy will be defeated if there is misincorporation of lysine for arginine during target-protein expression.

Among the targets of ubiquitin conjugation is ubiquitin itself (14). The unmutated ubiquitin gene has four AGA-encoded arginine residues (R42, R54, R72 and R74) and seven lysine residues. Here, we show that these endogenous AGA codons, and other AGA codons introduced by mutagenesis, are subject to extensive lysine misincorporation during ubiquitin expression in *E. coli* strain BL21(DE3)pLysS (18). Attempts to overcome this problem by replacing pLysS with an argU-encoding plasmid caused a severe reduction in the expression of ubiquitin. We have developed a new plasmid, pJY2, which carries both the argU and T7 lysozyme (LysS) genes. The use of pJY2 (in place of pLysS) allowed for robust expression of the above-described ubiquitins while reducing lysine misincorporation to a nearly undetectable level.

MATERIALS AND METHODS

Plasmids and Strains

pLysS was obtained from F. Studier (Brookhaven National Laboratory, Upton, NY, USA) (18). pDC952 (5,17) was obtained from J. Walker (University of Texas, Austin, TX, USA) and is identical to pACYC184 (6) except that pDC952 carries a 650-bp *Bam*HI/*Hind*III fragment containing the *E. coli* argU gene (15). A set of pET3a vectors encoding “single-lysine” ubiquitins were provided by M. Ellison (University of Alberta, Edmonton, Alberta, Canada) (1,13). K0 is a ubiquitin gene in which all seven lysine codons are mutated to AGA, while K6-, K11-, K27-, K29-, K33-, K48- and K63-Ub are genes in which every lysine codon, except the one indicated, has been changed to AGA. DNA sequence analysis showed that relative to the yeast ubiquitin sequence, pET3a-K63-Ub specified P19 (vs. S19) and E24 (vs. D24). The calculated mass of K63-ubiquitin is 8732 Da. K48R-ubiquitin, which has a calculated mass of 8593 Da, was expressed from pET3d-K48R-Ub (2).

A plasmid carrying both the argU and LysS genes was constructed by cloning the *Bam*HI fragment of pLysS (carrying the T7 lysozyme gene) into the *Bam*HI site of pDC952. We obtained clones with the lysozyme gene in both orientations relative to the *tet* promoter upstream of the argU gene. Clone pJY2, carrying the lysozyme gene in the antisense orientation, was selected for this work (Figure 1). As in pLysS (18), a weak T7 polymerase promoter (Ø3.8) immediately following the lysozyme gene allows for weak constitutive expression of lysozyme by transcription around the plasmid.

Protein Expression and Purification

The pET vectors encoding mutant ubiquitins were transformed into mineral-competent BL21(DE3) cells harboring one of the following helper plasmids: pLysS, pDC952 or pJY2 (each helper plasmid carried a chloramphenicol marker). Ubiquitin expression was induced with 0.4 mM isopropyl-1-thio-β-D-galactoside (IPTG) (2). Ubiquitins were purified at neutral pH by sequen-

Polyubiquitin Chain Assembly

The assembly of ubiquitin into K48-linked chains, which was catalyzed by the ubiquitin-conjugating enzyme E2_{25K}, was assayed in the presence of 10 μ M E2_{25K} and 0.3 mg/mL purified ubiquitin (pH 7.3, 37°C) (11). Incubations were sampled at time zero and 2 h. Chains were resolved by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), transferred to ImmobilonTM-P (Millipore, Bedford, MA, USA) for 3 h at 68 V (5°C) and probed with affinity-purified anti-ubiquitin antibodies (10).

Electrospray Ionization Mass Spectrometry

Mass spectra were obtained at the University of Iowa Mass Spectrometry Facility using an Autospec Mass Spectrometer (Micromass, Manchester, England, UK) equipped with an electrospray interface. Samples in water/acetonitrile/formic acid (50:50:1) were introduced by direct infusion at a flow rate of 10 $\mu\text{L}/\text{min}$. A mass range (m/z) of 700–2200 Da was scanned, and for each sample, approximately 30 scans were acquired, averaged and transformed using OPUSTM software (Micromass). The instrument was calibrated with tetraethyl ammonium iodide to within 0.2 Da.

Extensive Lysine Misincorporation During Expression of Ubiquitin Genes with Multiple AGA Codons in Strain BL21(DE3)pLysS

Eight pET3a plasmids encoding mutant ubiquitins were expressed in BL21(DE3)pLysS cells. Each ubiquitin gene specified four natural AGA-encoded arginine residues (R42, R54, R72 and R74). In seven of the plasmids, six (of seven) natural lysine residues were mutated to arginine using the AGA codon (K6-, K11-, K27-, K29-, K33-, K48- and K63-ubiquitin); the eighth plasmid encoded a lysine-free protein, K0-ubiquitin (1,13). Each mutant protein was expressed at a high level, as shown for K63-ubiquitin in Figure 2. Coomassie® Blue staining revealed a prominent band of approximately 8 kDa (Figure 2A, lane 2); the identity of this protein as ubiquitin was confirmed by western blotting (Figure 2B, lane 2). Thus, the presence of a high density of AGA codons (14%–16% of total codons) did not inhibit ubiquitin expression.

The presence of AGA codons can also cause misincorporation of lysine for arginine (5,9). As an initial test of whether such misincorporation oc-

Figure 1. Map of pJY2. Top: diagram of vector; except for the LysS and argU genes, the sequence of the vector is identical to the sequence of pACYC184. Selected unique restriction sites are shown with their map positions. Bottom: sequences flanking argU gene; the *Bam*HI (5') and *Hind*III sites (3') used in cloning are underlined; the start and stop codons of the argU gene are bracketed and shown in bold.

Figure 2. K63-ubiquitin expression in different strains. (A) Coomassie Blue-stained gel. Lane 1, molecular mass standards (in kDa); lanes 2-4, μ L of soluble extract from 4-h induced BL21 (DE3) *E. coli* cells transformed with pET3a-K63Ub and the following helper plasmid: pLysS (lane 2); pDC952 (lane 3); pJY2 (lane 4). (B) Western blot developed with anti-ubiquitin antibodies. Lane 1, standard (20 ng); lanes 2-4, 1 μ L of 1000-fold dilutions of samples in lanes 2-4 of panel A; lane 5, 2 μ L of undiluted extract from lane 3 of panel A. All lanes in panel B derived from one blot; intervening lanes were excised between lanes 4 and 5.

Short Technical Reports

curred with the mutant ubiquitins, we evaluated the purified proteins as substrates for ubiquitin-ubiquitin conjugation (i.e., polyubiquitin chain assembly) catalyzed by ubiquitin-conjugating enzyme E2_{25K}. This enzyme exclusively used K48 in chain assembly, as determined by direct tryptic peptide mapping (8) and its inability to polymerize K48R-ubiquitin (2,11). If there was no lysine misincorporation, then only K48-ubiquitin would be a substrate for E2_{25K}. Instead, all of the mutant ubiquitins, including the "lysine-free" protein K0-ubiquitin, were active as substrates (Figure 3A and data not shown). Thus, in spite of robust expression, extensive misincorporation of lysine at the AGA₄₈ codon occurred during the expression of each mutant ubiquitin gene.

The extent of misincorporation was quantitatively analyzed by electrospray ionization mass spectrometry (ESIMS) of purified K63-ubiquitin. The spectrum displayed seven peaks: one of approximately the expected mass (8732 Da) and six peaks of lower mass (Figure 4A). Each peak differed in mass from its neighbor(s) by an increment of 28 Da, corresponding to the mass difference between lysine and arginine. Thus, successively lower-mass peaks represented proteins that contain an increasing number of misincorporated lysine residues. Only a tiny fraction (ca. 3%) of the molecules corresponded to the translation product actually specified by the K63-ubiquitin gene. Calderone et al. reported 36%–42% misincorporation at AGA codons during recombinant protein expression in *E. coli* (5). The distribution of translation products seen in Figure 4A fits well to the binomial distribution predicted by a 35%–40% probability of lysine misincorporation at each of the 11 AGA codons in the K63-ubiquitin gene.

Poor Expression in BL21(DE3)pDC952 Cells

Co-expression of tRNA_{UCU}^{Arg}, encoded by the argU gene in pDC952, has been reported to suppress lysine misincorporation at AGA codons (5). However, when BL21(DE3)pDC952 cells were used for mutant ubiquitin expression, the level of expressed ubiquitin was reduced to less than 1% (Figure 2, A and

B, lane 3). Although it is likely that lysine misincorporation was significantly suppressed (5,9), the low expression of K63-ubiquitin prevented us from carrying out ESIMS analysis to address this issue. We suspected that the low expression was because of the absence of T7 lysozyme, which was expressed from pLysS in the studies described above. T7 lysozyme suppresses constitutive expression of pET-encoded genes while allowing a high level of expression upon induction of T7 RNA polymerase; the reduction in constitutive expression can reduce toxicity correlated with recombinant protein expression (18).

Restoration of Translational Fidelity and Robust Expression in BL21(DE3) pJY2 Cells

To reduce constitutive ubiquitin expression and simultaneously reduce lysine misincorporation, we sought to co-express lysozyme and tRNA_{UCU}^{Arg}. Because pLysS and pDC952 are both derived from pACYC184 (17,18) and thus cannot be stably maintained in the same cell (6), we generated a new plasmid, pJY2, that carried the lysozyme and argU genes (see Materials and Methods and Figure 1). pJY2 was transformed into BL21(DE3) cells. Competent cells were then transformed with the pET-ubiquitin plasmids.

The presence of pJY2 restored high

expression of the mutant proteins (Figure 2, A and B, lane 4), suggesting that the low expression in BL21(DE3)-pDC952 cells was indeed due to toxicity associated with premature ubiquitin expression. Moreover, the presence of the argU gene in pJY2 strongly enhanced translational fidelity, as shown by the finding that only K48-ubiquitin was well-utilized by E2_{25K} (Figure 3B, the band labeled E2-Ub corresponds to an auto-ubiquitinated form of E2_{25K}, i.e., it is not a polyubiquitin chain). Traces of diubiquitin seen in several reactions in Figure 3B may be due to a low level of lysine misincorporation at AGA₄₈ or to a weak ability of E2_{25K} to utilize other lysine residues.

A strong reduction in lysine misincorporation was rigorously confirmed by ESIMS analysis of purified K63-ubiquitin. Figure 4B shows a single major translation product, whose mass (8732 Da) agreed with that calculated from the deduced amino acid sequence. The several minor peaks do not show the differences in mass (± 28 Da) that are expected if these peaks arise from lysine misincorporation (see Figure 4B legend). ESIMS analysis of K0-ubiquitin purified from BL21(DE3)pJY2 cells indicated that this protein also exhibited negligible lysine misincorporation (not shown).

The presence of AGG arginine codons can also be problematic for gene expression in *E. coli*. The pres-

Figure 3. E2_{25K}-catalyzed assembly of K48-linked polyubiquitin chains from wild-type and mutant ubiquitins. Chain assembly reactions were carried out as described in Materials and Methods, with wild-type ubiquitin (Sigma, St. Louis, MO, USA) or purified mutant ubiquitins (recombinant, as indicated at the bottom). Odd-numbered lanes: time zero; even-numbered lanes, after 2 h of incubation. E2-Ub, auto-mono-ubiquitinated E2_{25K} (text). (A) Ubiquitins purified from BL21(DE3)pLysS cells. (B) Ubiquitins purified from BL21(DE3)pJY2 cells. Intervening lanes were excised between lanes 6 and 7 each panel.

ence of tandem AGG codons can lead to frameshifting and thereby decrease the expression of recombinant proteins, while the presence of multiple dispersed AGG codons can reduce expression level (3,4,7,16,17). Prior studies of tRNA^{Arg}_{UCU}-dependent suppression of frameshifting at tandem AGG and tandem AGA codons showed that this tRNA is AGA-specific when it is expressed from the endogenous argU gene, but is able to decode AGG when overexpressed (17). This divergence in properties has been suggested to reflect incomplete modification of bases in the overexpressed tRNA (17). Because pJY2 is a high copy plasmid (10), most likely pJY2 will ameliorate problems associated with the presence of multiple AGG codons. However, we did not address this question experimentally.

In contrast to frameshifting or pausing, lysine misincorporation at AGG codons has not been reported. In the case of ubiquitin, we did not observe misincorporation at a single AGG codon in the gene encoding a different mutant protein, K48R-ubiquitin. This

gene has four natural AGA codons, plus an AGG-encoded arginine at residue 48 (2). As in the case of K63-ubiquitin, K48R-ubiquitin was highly expressed in BL21(DE3)pLysS cells and was subject to extensive lysine misincorporation (Figure 4C), indicating that tRNA^{Arg}_{UCU} was limiting. However, this misincorporation occurred exclusively at the AGA codons, because K48R-ubiquitin expressed in BL21(DE3)pLysS cells cannot be polymerized by E2_{25K} (2,11). This finding is consistent with the conclusion that the product of the endogenous argU gene specifically decodes AGA vs. AGG (17). Lysine misincorporation at the AGA codons of K48R-ubiquitin was suppressed when this protein was expressed in BL21(DE3)pJY2 cells (Figure 4D).

Our results demonstrate that co-expression of tRNA^{Arg}_{UCU} and T7 lysozyme provides great advantages: tRNA^{Arg}_{UCU} expression from the plasmid-borne argU gene virtually eliminated lysine misincorporation at AGA codons, even when these codons were

Figure 4. ESIMS analysis of purified recombinant ubiquitins. (A) K63-ubiquitin expressed in BL21(DE3)pLysS cells. (B) K63-ubiquitin expressed in BL21(DE3)pJY2 cells. (C) K48R-ubiquitin expressed in BL21(DE3)pLysS cells. (D) K48R-ubiquitin expressed in BL21(DE3)pJY2 cells. The arrow in each panel indicates the peak for the molecular ion (MH^+) of the ubiquitin species with the expected composition; in panels A and C, the major peaks with lower mass than the expected MH^+ species are smaller by increments of 28 Da and correspond to protein with arginine(s) replaced by lysine(s). The minor peaks X, Y and Z correspond to oxidized (as MetSO) ubiquitin, a potassium adduct and ubiquitin containing both sodium and potassium.

Short Technical Reports

present at a high density, while lysozyme expression reduced constitutive expression of the target gene and allowed for robust expression upon induction. Thus, eukaryotic genes cloned into pET vectors (or other T7 polymerase-based vectors) can be efficiently and accurately expressed in BL21-(DE3)pJY2 cells even when these genes harbor multiple AGA codons.

These benefits are documented here for the case of ubiquitin; we have also observed that pJY2 affords enhanced expression of non-ubiquitin proteins whose genes harbor AGA codons (data not shown). The use of pJY2 is more convenient than expressing tRNA-UCU^{Arg} and T7 lysozyme from separate plasmids (e.g., 4). Moreover, pJY2 is more versatile than the previously described argU-encoding plasmids pUBS520 (3) and pSBET (16), which carry kanamycin markers and therefore cannot be used with kanamycin-selectable pET vectors; the chloramphenicol

marker in pJY2 is compatible with all currently available pET vectors. pJY2 can also be used as a source of tRNA-UCU^{Arg} with expression systems, such as the pGEX series of vectors, that do not rely on T7 polymerase.

REFERENCES

1. Arnason, T. and M.J. Ellison. 1994. Stress resistance in *Saccharomyces cerevisiae* is strongly correlated with assembly of a novel type of multiubiquitin chain. *Mol. Cell. Biol.* 14:7876-7883.
2. Beal, R., Q. Deveraux, G. Xia, M. Rechsteiner and C.M. Pickart. 1996. Surface hydrophobic residues of multiubiquitin chains essential for proteolytic targeting. *Proc. Natl. Acad. Sci. USA* 93:861-866.
3. Brinkmann, U., R.E. Mattes and P. Buckel. 1989. High-level expression of recombinant genes in *Escherichia coli* is dependent on the availability of the *dnaY* gene product. *Gene* 85:109-114.
4. Buhler, C., D. Gabelle, P. Forterre, J.C. Wang and A. Bergerat. 1998. Reconstitution of DNA topoisomerase VI of the thermophilic archaeon *Sulfolobus shibatae* from subunits separately overexpressed in *Escherichia coli*. *Nucleic Acids Res.* 26:5157-5162.
5. Calderone, T.L., R.D. Stevens and T.G. Oas. 1996. High-level misincorporation of lysine for arginine at AGA codons in a fusion protein expressed in *Escherichia coli*. *J. Mol. Biol.* 262:407-412.
6. Chang, A.C. and S.N. Cohen. 1978. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. *J. Bacteriol.* 134:1141-1156.
7. Chen, G.T. and M. Inouye. 1994. Role of the AGA/AGG codons, the rarest codons in global gene expression in *Escherichia coli*. *Genes Dev.* 8:2641-2652.
8. Chen, Z. and C.M. Pickart. 1990. A 25-kilodalton ubiquitin carrier protein (E2) catalyzes multi-ubiquitin chain synthesis via lysine 48 of ubiquitin. *J. Biol. Chem.* 265:21835-21842.
9. Forman, M.D., R.F. Stack, P.S. Masters, C.R. Hauer and S.M. Baxter. 1998. High level, context dependent misincorporation of lysine for arginine in *Saccharomyces cerevisiae* a1 homeodomain expressed in *Escherichia coli*. *Protein Sci.* 7:500-503.
10. Haas, A.L. and P.M. Bright. 1985. The immunochemical detection and quantitation of intracellular ubiquitin-protein conjugates. *J. Biol. Chem.* 260:12464-12473.
11. Haldeman, M.T., G. Xia, E.M. Kasperek and C.M. Pickart. 1997. Structure and function of ubiquitin conjugating enzyme E2-25K: the tail is a core-dependent activity element. *Biochemistry* 36:10526-10537.
12. Hochstrasser, M. 1996. Ubiquitin-dependent protein degradation. *Annu. Rev. Genet.* 30:405-439.
13. Hodgins, R., C. Gwozd, T. Arnason, M. Cummings and M.J. Ellison. 1996. The tail of a ubiquitin-conjugating enzyme redirects multi-ubiquitin chain synthesis from the lysine 48-linked configuration to a novel nonlysine-linked form. *J. Biol. Chem.* 271:28766-28771.
14. Pickart, C.M. 1997. Targeting of substrates to the 26S proteasome. *FASEB J.* 11:1055-1066.
15. Saxena, P. and J.R. Walker. 1992. Expression of argU, the *Escherichia coli* gene coding for a rare arginine tRNA. *J. Bacteriol.* 174:1956-1964.
16. Schenk, P.M., S. Baumann, R. Mattes and H.-H. Steinbiss. 1995. Improved high-level expression system for eukaryotic genes in *Escherichia coli* using T7 RNA polymerase and rare Arg tRNAs. *BioTechniques* 19:196-200.
17. Spanjaard, R.A., K. Chen, J.R. Walker and J. van Duin. 1990. Frameshift suppression at tandem AGA and AGG codons by cloned tRNA genes: assigning a codon to argU tRNA and T4 tRNA(Arg). *Nucleic Acids Res.* 18:5031-5036.
18. Studier, F.W. 1991. Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system. *J. Mol. Biol.* 219:37-44.

We thank Dr. Lynn Teesch for her help with the mass spectrometry and members of our laboratories for helpful comments on the manuscript. Funded by National Institutes of Health Grant Nos. DK46984 (to C.M.P.) and GM37666 (to R.E.C.). Address correspondence to Cecile M. Pickart, Department of Biochemistry, School of Public Health, Johns Hopkins University, 615 North Wolfe Street, Baltimore, MD 21205, USA. Internet: cpickart@welchlink.welch.jhu.edu

Received 5 April 1999; accepted 23 June 1999.

Jianxin You, Robert E. Cohen¹ and Cecile M. Pickart
*Johns Hopkins University
Baltimore, MD*
¹*University of Iowa
Iowa City, IA, USA*