Proline Residues at the C Terminus of Nascent Chains Induce SsrA Tagging during Translation Termination*§

Christopher S. Hayes‡, Baundauna Bose, and Robert T. Sauer§

From the Department of Biology Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

The SsrA or tmRNA quality control system relieves ribosome stalling and directs the addition of a degradation tag to the C terminus of the nascent chain. In some instances, SsrA tagging of otherwise full-length proteins occurs when the ribosome pauses at stop codons during normal translation termination. Here, the identities of the C-terminal residues of the nascent chain are shown to play an important role in full-length protein tagging. Specifically, a subset of C-terminal Xaa-Pro sequences caused SsrA tagging of the full-length YbeL protein from Escherichia coli. This tagging increased when a less efficient stop codon was used, increased in cells lacking protein release factor-3, and decreased when protein release factor-1 was overexpressed. Incorporation of the analog azetidine-2-carboxylic acid in place of proline suppressed tagging, whereas incorporation of 3,4-dehydroproline increased SsrA tagging of full-length YbeL. These results suggest that the detailed chemical or conformational properties of the C-terminal residues of the nascent polypeptide can affect the rate of translation termination, thereby influencing ribosome pausing and SsrA tagging at stop codons.

All eubacteria contain an RNA molecule known as SsrA or tmRNA that functions in the release of stalled ribosomes from mRNAs and in targeting the nascent polypeptides from such ribosomes for proteolysis (1, 2). SarA has been shown to act as both a transfer RNA and a messenger RNA (1–4). SsrA has a domain that is very similar to a portion of tRNA\[^{4\text{th}}\] and is aminoacylated with alanine at its 3′ end. SsrA also contains a short open reading frame, which in Escherichia coli encodes the sequence ANDENYALAA (3). The tmRNA model of SsrA activity postulates that SsrA enters the empty A-site of stalled ribosomes and adds its charged alanine to the nascent polypeptide by transpeptidation. A conformational change then occurs that allows translation to shift from the original mRNA to the ANDENYALAA reading frame within SsrA. As a result, the stalled ribosome is freed from the problematic mRNA, and the SsrA-encoded peptide tag is added to the C-terminus of the nascent polypeptide. The SarA tag is recognized by multiple proteases in E. coli, resulting in rapid degradation of SarA-tagged proteins (4–6).

The determinants of SsrA tagging are well understood in only a few cases. For example, SsrA tagging has been demonstrated for proteins synthesized from truncated messages lacking in-frame stop codons and at protein positions corresponding to rare codons (4, 7–10). Ribosomes would be expected to stall at the 3′ end of the incomplete mRNA or at the rare codon mRNA position when the cognate charged tRNA was scarce. SsrA tagging of otherwise full-length proteins has also been observed, suggesting that ribosomes stall under some circumstances at stop codons (8, 10). The identity of the stop codon can affect SsrA tagging of full-length proteins, with less efficient stop codons leading to higher levels of tagging (10). Because termination of translation is a relatively slow process compared with translation elongation (11), SsrA and protein release factors probably compete for binding to the A-site, while the ribosome idles at an inefficient stop codon. However, SsrA tagging of full-length proteins encoded by genes with efficient stop signals has also been observed (12). For example, the intact E. coli YbeL protein is tagged (12) even though the ybeL coding region ends with an extended UAAU stop signal which is thought to be the most efficient translation termination signal in E. coli (13, 14).

In this paper, we dissect the YbeL system to identify the determinants that lead to its efficient tagging by SsrA. Mutational analyses of the last two residues of YbeL demonstrated that the proline residue at the C terminus and the identity of the penultimate residue are major determinants of SsrA tagging. Moreover, substituting proline for the C-terminal residue of thioredoxin, a protein that is not normally tagged, resulted in efficient SsrA-tagging of this variant. Tagging of full-length YbeL was modulated by the identity of the stop codon and by the cellular level of protein release factors. The incorporation of proline analogs into YbeL significantly affected SsrA tagging at the stop codon, suggesting that the structure of the nascent polypeptide is a primary determinant of SsrA tagging of YbeL.

EXPERIMENTAL PROCEDURES

Bacterial Strains and Plasmids—Table I lists the E. coli strains and plasmids used in this study. Strain CH522 lacks the P\(^{3}\) episome found in CH113 and was obtained as a Lac- colony following growth of CH113 to stationary phase in 50 μg/ml of acridine orange at 37 °C in the dark. Strain CH522 and its derivatives require L-proline and L-arginine for growth.

The prfC gene for protein release factor-3 (RF-3)\(^{1}\) was deleted from the E. coli chromosome using the phage A Red recombination system according to Datsenko and Wanner (15) with modifications. The FLP recombinase site-flanked kanamycin cassette from plasmid pKD46 (15) was PCR amplified and ligated to SmaI-digested plasmid pBlue-script SK+ to generate plasmid pKAN. A region of DNA upstream of

\[^{1}\] The abbreviations used are: RF, release factor; RRF, ribosome recycling factor; Ni\(^{2+}\)-NTA, Ni\(^{2+}\)-nitrilotriacetic acid; MALDI, matrix-assisted laser desorption/ionization; IPTG, isopropyl-1-thio-β-D-galactopyranoside; Aze, azetidine-2-carboxylic acid.

* This work was supported by Grant AI-16892 from the National Institutes of Health. The costs of publication of this article were defrayed in part by the payment of page charges. This article must be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ To whom correspondence should be addressed: Dept. of Biology, MIT, Rm. 68-571, 77 Massachusetts Ave., Cambridge, MA 02139. Tel.: 617-253-3163; Fax: 617-258-0673; E-mail: bobsauer@mit.edu.

‡ Supported by Fellowship DRG 1686 from the Cancer Research Fund of the Damon Runyon-Walter Winchell Foundation.

Received for publication, May 31, 2002
Published, JBC Papers in Press, July 8, 2002, DOI 10.1074/jbc.M205405200

This paper is available on line at http://www.jbc.org/
SsrA Tagging Induced by Nascent Chains

| Strain or plasmid | Genotype or description* | Ref. |
|------------------|--------------------------|------|
| Strains |
| X90 | F' lacI lacI proI ara D(lac-pro) nalA argEam(ram) rif' thi-1 | 8 |
| CH113 | F' lacI lacI proI ara D(lac-pro) nalA argEam(ram) rif' thi-1 ssrA::cat (DE3) | 10 |
| CH522 | ara D(lac-pro) nalA argEam(ram) rif' thi-1 ssrA::cat (DE3) | This study |
| AT46 | Any | This study |
| Plasmids |
| pKD46 | Red recombinase expression plasmid, Amp' | 15 |
| pCP20 | FLP recombinase expression plasmid; Amp', Cm' | 15 |
| pKAN | pBlueScript SK+ with FLP-flanked kanamycin resistance cassette; Amp', Km' | This study |
| pYbel | pET11d::ybel; wild-type Ybel expression plasmid, Amp' | This study |
| pYbel(His6) | pET24d::ybel(His6); Ybel expression plasmid with His6 tag, Km' | This study |
| pKW23 | Contains ssaA gene that encodes ANDENYALDD peptide tag, Tet' | 8 |
| pCH201 | Contains ssaA gene that encodes ANDH 6D peptide tag, Tet' | 10 |
| pCH405 | Contains multiple cloning site, and an ssaA gene that encodes ANDH 6D peptide tag, Tet' | This study |
| pRF-3 | Contains prfC (encoding RF-3), and an ssaA gene that encodes ANDH 6D peptide tag, Tet' | This study |
| pRF-1 | Contains prfC (encoding RF-1), and an ssaA gene that encodes ANDH 6D peptide tag, Tet' | This study |
| pRRF | Contains frf (encoding RRF), and an ssaA gene that encodes ANDH 6D peptide tag, Tet' | This study |

* Abbreviations as are follows: Amp', ampicillin resistance; Km', kanamycin resistance; Tet', tetracycline resistance.

Identification of YbeL-AANDH6D Junction Peptides from Protein Expression Libraries—Strain CH113/pCH201 was transformed with Ybel expression plasmid libraries followed by overnight culture in LB broth supplemented with 100 μg/ml ampicillin and 20 μg/ml tetracycline. The overnight cultures were diluted into 50 ml of fresh media to an optical density at 600 nm (OD600) of 0.1 and grown at 37 °C until the culture reached an OD600 of ~1. Expression of YbeL was induced by the addition of IPTG to final concentration of 1.5 mM, followed by growth at 37 °C for 2.5 h. Cultures were harvested, frozen at −80 °C and lysed in 12 ml of 0.05 M Tris HCl (pH 8.0), 8 M urea, 100 mM NaCl with stirring for 20 min at room temperature. SsrA(ANDH6D)-tagged proteins were purified by Ni2+−nitrilotriacetic acid (Ni2+-NTA) affinity chromatography (Qiagen) using urea lysis buffer as the column buffer. SsrA(ANDH6D)-tagged protein was eluted with 1.6 M of elution buffer (2 M urea, 100 mM acetic acid), the pH of the eluate was adjusted to pH ~8 with ~2 M Tris HCl (pH 9.5), and trypsin (0.04 mg/ml; Worthington Biochemicals) was added for 3 h at 37 °C. YbeL-AANDH6D junction peptides were purified by a second round of Ni2+-NTA affinity chromatography, eluted with 0.06% trifluoroacetic acid, and chromatographed on a C4 reverse phase column (Vydac). Peptides were detected by UV absorption at 214 nm and identified by matrix-assisted laser desorption/ionization (MALDI) mass spectrometry at the MIT Biopolymers laboratory.

SsrA Tagging of YbeL Mutants in Vivo—Strains of E. coli CH113 and CH468 containing plasmid pCH201 and Ybel expression plasmids were grown overnight at 37 °C in LB broth supplemented with 100 μg/ml ampicillin and 20 μg/ml tetracycline. New cultures were grown the following day by resuspending samples to an OD600 of 0.2 in fresh medium. After growth for 2 h at 37 °C, protein synthesis was inhibited by the addition of IPTG to a final concentration of 1.5 mM, followed by growth for 2 h. 1.0-ml samples were removed and harvested by centrifugation. The supernatant was aspirated, and the cell pellet was frozen at −80 °C.

Samples for Western blot analysis were extracted from cell pellets by vigorous vortexing in 60 μl of urea lysis buffer, centrifuged in a microcentrifuge at top speed for 10 min, and the urea-soluble total protein was collected as the supernatant. Protein concentration was estimated by Bradford assay and equal amounts of total urea-soluble protein were resolved by 10% Tris-tricine SDS-PAGE. Proteins were transferred to nitrocellulose membranes by electroblotting, and probed with the following antibodies: SsrA(ANDH6D)-tagged proteins were detected by using His-probe H15 antibodies (Santa Cruz Biotechnology), and SsrA(ANDH6D)-tagged proteins were detected with rabbit polyclonal antisera as described (8, 10).

Incorporation of Proline Analogs in YbeL—Overnight cultures of strain CH113/pCH468 containing Ybel expression plasmids and plasmids pCH201 or pKW23 were resuspended at an OD600 of 0.05 in glucose minimal medium supplemented with 1 μM/ml thiamine, 50 μM proline, 450 μM arginine, and appropriate antibiotics. Minimal medium cultures (100 ml) were grown at 37 °C to an OD600 of 0.6 with 0.6 and 0.9 IPTG was added to a final concentration of 1.5 mM, and growth was continued for 10 min. The culture was then split into four aliquots (25 ml each), and proline analogs were added to a final concentration of 1 mM. Proline
RESULTS

Mutation of the Last Two Residues of YbeL Affects SsrA Tagging—We initially used a library approach to determine whether the identity of the C-terminal residues of E. coli YbeL or their coding sequences influenced tagging. The codons for the C-terminal residue or the penultimate residue were randomized in separate experiments (Fig. 1A), and the mixture of resulting YbeL proteins was expressed in a strain lacking wild-type YbeL but containing a SsrA variant encoding an ANDH6D tag. Proteins tagged by SsrA(ANDH6D) are not degraded rapidly, can be purified by Ni²⁺-NTA affinity chromatography, and can be detected by anti-His6 antibodies. For these experiments, tagged proteins were purified by Ni²⁺-NTA chromatography and digested with trypsin. Following additional purification, junction peptides between YbeL and the AANDH6D tag were identified by mass spectrometry. YbeL expressed from the wild-type gene yielded a single YbeL-AANDH6D junction peptide with a mass corresponding to the C-terminal tryptic peptide of YbeL joined to the AANDH6D tag (Table II, Fig. 1A). Following randomization of the C-terminal or (−1) YbeL position, this same junction peptide was recovered, but no other peptides corresponding to full-length protein tagging were identified (Table II). This result suggested that substitution of the C-terminal proline of YbeL eliminated or reduced tagging. Numerous full-length junction peptides were identified from the (−2) position library (Table II), including those containing Glu, Asp, Ala, Ser, Gly, Pro, Val, Tyr, Phe, and His residues at this position. Full-length junction peptides containing either Ile/Leu or Lys/Gln at the (−2) position were also identified but could not be differentiated because of mass equivalence (Table II). The recovery of a large number of different tagged peptides in this experiment indicated that efficient tagging occurs with many different amino acids or codons at the penultimate position when the C-terminal residue is proline.

Individual clones representing four amino acid substitutions at the (−1) position and all possible substitutions at the (−2) position were isolated from the libraries and Western blot analysis was used to examine YbeL tagging (Fig. 1). Mutation of the wild-type proline at the (−1) position to Ala, Leu, Lys, or Ser resulted in large reductions in the level of SsrA tagging (Fig. 1B). At the (−2) position, high levels of tagging were seen for proteins containing Glu (wild type), Asp, Ile, Val, and Pro; moderate levels of tagging were observed for Ala, Gly, Ser, Tyr, Glu, Phe, and Lys; and low levels of tagging resulted from substitution by Met, Thr, His, Cys, Trp, Leu, Asn, and Arg (Fig. 1C).

Transfer of Full-length Tagging to Another Protein—Is a C-terminal proline sufficient to cause SsrA tagging at stop codons in other genetic contexts? As shown in Fig. 2A, wild-type thioredoxin (TrxA) from E. coli is not tagged significantly by SsrA(ANDH6D). We constructed variant trxA genes encoding C-terminal Ala → Pro or Leu-Ala → Glu-Pro substitutions (Glu-Pro is the C-terminal sequence of YbeL). Both variants were efficiently tagged by SsrA(ANDH6D) (Fig. 2A). Mass spectrometry of the purified TrxA-AANDH6D junction peptides confirmed that tagging occurred at the C terminus of the full-length thioredoxin variants (data not shown). Thus, a single C-terminal proline was sufficient to induce tagging at a stop codon where it did not occur normally.

Tagging Does Not Depend on the Identity of the C-terminal Proline Codon—Studies of translation termination have suggested that interactions between the protein release factors that decode stop codons and certain P-site tRNAs may lead to inefficient translation termination (16–19). Proline is encoded by four codons (CCA, CCC, CCG, and CCU), which are decoded by three tRNAPro species in E. coli. To determine whether a specific proline codon or tRNAPro was required for tagging, we constructed ybeL variants with all possible proline codons at the (−1) position. As shown in Fig. 2B, YbeL was tagged efficiently in each case. These results are consistent with a model in which the proline residue in the nascent chain causes tagging. Alternatively, codon-specific or tRNA-specific sequence determinants may be shared by all proline codons and/or tRNAPro combinations.

Stop Codon Effects—If SsrA tagging at the C terminus of YbeL is caused, in part, by prolonged pausing at the ybeL stop codon, then mutations that increase pausing should acerbate tagging and vice versa. The ybeL gene uses an ochre (UAU) stop codon. To determine the importance of the termination codon, new (−2) position YbeL libraries were made with amber (UAG) or opal (UGA) termination codons. The YbeL-AANDH6D junction peptides from the amber library were identical to those from the original ochre library, although peptide yields were lower from the amber library (data not shown). By contrast, there was an increase in the relative yields of several peptides from the opal library, including those containing His, Phe, Tyr, Ile/Leu, and Glu/Lys residues at the (−2) position (data not shown). Fig. 3 shows a comparison of YbeL tagging for constructs with ochre or opal stop codons and Glu, Ala, Phe, His, Lys, Leu, Asn, Trp, and Tyr at the (−2) position. For all pairs except one where tagging could not be detected (Fig. 3), the level of tagging was higher for the construct with the less efficient (UGA) than more efficient (UAU) stop codon. This increase was particularly striking with His, Lys, Leu, or Tyr at the (−2) position (Fig. 3).

| Table II | Mutants (A) | Pro | Ala | Leu | Lys | Ser | Ile/Leu | Glu | Asp | Ile | Val | Tyr | Gly | Pro | stop stop |
|---------|-------------|-----|-----|-----|-----|-----|--------|-----|-----|-----|-----|-----|-----|-----|--------|
| ybeL (wt) |    Pro     |   Ala |    Leu |   Lys |    Ser | Ile/Leu |  Glu |  Asp | Ile | Val | Tyr | Gly | Pro | stop stop |
| ybeL (wt) |    Pro     |   Ala |    Leu |   Lys |    Ser | Ile/Leu |  Glu |  Asp | Ile | Val | Tyr | Gly | Pro | stop stop |
| ybeL (wt) |    Pro     |   Ala |    Leu |   Lys |    Ser | Ile/Leu |  Glu |  Asp | Ile | Val | Tyr | Gly | Pro | stop stop |
| ybeL (wt) |    Pro     |   Ala |    Leu |   Lys |    Ser | Ile/Leu |  Glu |  Asp | Ile | Val | Tyr | Gly | Pro | stop stop |
SsrA Tagging Induced by Nascent Chains

Table II

| Random codon library | Predicted YbeL-AANDH₆₋D junction peptide | Measured mass of junction peptide | Calculated average mass of predicted junction peptide |
|----------------------|------------------------------------------|----------------------------------|-----------------------------------------------------|
| Wild-type (-1)       | RPFEP-tag                                | 1954.84                          | 1955.0                                              |
| Wild-type (-2)       | RPFDP-tag                                | 1996.83                          | 1997.0                                              |
|                       | RPFEP-tag                                | 1940.91                          | 1941.0                                              |
|                       | RPFDP-tag                                | 1954.83                          | 1955.0                                              |
|                       | RPFPP-tag                                | 1972.96                          | 1973.1                                              |
|                       | RPFGF-tag                                | 1882.82                          | 1883.0                                              |
|                       | RPFHP-tag                                | 1962.93                          | 1963.1                                              |
|                       | RPFI(L)P-tag                             | 1938.96                          | 1939.1 (1939.1)                                     |
|                       | RPFKQ(P)-tag                             | 1954.04                          | 1954.1 (1954.1)                                     |
|                       | RPFEP-tag                                | 1922.94                          | 1923.0                                              |
|                       | RFFBP-tag                                | 1912.82                          | 1913.0                                              |
|                       | RFFVP-tag                                | 1924.81                          | 1925.1                                              |
|                       | RFPYV-tag                                | 1988.91                          | 1989.1                                              |

*YbeL portion of the junction peptide is given in one-letter code.

Fig. 2. C-terminal prolines induce full-length protein tagging. A, Western blot of thioredoxin tagging. Strains containing SsrA(ANDH₆₋D) as well as wild-type thioredoxin or mutant proteins containing C-terminal proline residues were grown with or without IPTG, and cell proteins were analyzed by SDS-PAGE and Western blot. The positions of SsrA(ANDH₆₋D)-tagged thioredoxin and overexpressed thioredoxin are indicated by arrows on the Western blot and Coomassie Blue-stained gel, respectively. B, analysis of SsrA(ANDH₆₋D)-tagged YbeL expressed from constructs containing different proline codons by SDS-PAGE and Western blot. The position of SsrA(ANDH₆₋D)-tagged YbeL is indicated by the arrow on the Coomassie Blue-stained gel.

Altered Levels of Translation Termination Factors Affect SsrA Tagging—The efficiency of translation termination depends on the levels of the decoding protein release factors as well as the identity of the stop codon. Ochre stop codons are recognized by release factors RF-1 and RF-2. Tagging of wild-type YbeL, which terminates with an ochre codon, was substantially decreased in cells carrying a multi-copy plasmid encoding the prfC gene for RF-1 (Fig. 4A). Overexpression of RF-1 also decreased tagging of a YbeL variant ending in Pro-Pro, which is normally tagged at very high levels (Fig. 4A). We did not examine the effects of overexpression of RF-2, which has been reported to be lethal (20). RF-3 increases the rate of RF-1 and RF-2 dissociation from the ribosome after peptidyl-tRNA hydrolysis (21, 22) and increases termination efficiency by maintaining the pool of free RF-1 and RF-2 available to decode stop codons (23). YbeL tagging increased significantly in ΔprfC cells lacking RF-3 (Fig. 4B) but occurred at normal levels in these cells when the prfC gene was provided on a plasmid (Fig. 4B). RRF facilitates the dissociation of ribosomal subunits after peptidyl-tRNA hydrolysis (24, 25). Overexpression of RRF caused a small decrease in the levels of SsrA tagging of full-length YbeL (Fig. 4A). Thus, changes in the levels of release factors that would be expected to make translation termination less efficient result in increased levels of YbeL tagging by SsrA tagging and vice versa.

Proline Analogs Affect Full-length Tagging—The proline analogs azetidine-2-carboxylic acid, γ-thiapropyl, and 3,4-dehydroproline are incorporated into proteins in E. coli (26, 27). Using a proline auxotrophic strain, we examined the effects of these proline analogs on SsrA tagging of a variant of YbeL containing a His₆ sequence before the last four residues of the wild-type protein (Fig. 5). The His₆ epitope allowed the C-terminal tryptic peptide of YbeL (Fig. 5A) to be purified by Ni²⁺-NTA affinity chromatography and then analyzed by reverse-phase HPLC and MALDI mass spectrometry to assay incorporation of the analogs (Fig. 5B; Table III). Incorporation of 3,4-dehydroproline for both proline residues in the C-terminal tryptic peptide of YbeL was complete, and incorporation of azetidine-2-carboxylic acid occurred at >95% efficiency (Fig. 5; Table III). γ-Thiapropyl incorporation was more variable, but the major C-terminal peptide species contained two γ-thiapropyl residues (Fig. 5B; Table III).

Because the YbeL protein used for these analog studies contained a His₆ sequence, we used an SsrA variant that encodes an ANDENYALDD tag (SsrA(DD)) and assayed YbeL tagging by Western blots using antisera against this tag (Fig. 6). Strik-
ingly, azetidine-2-carboxylic acid incorporation almost completely suppressed SsrA tagging of the full-length YbeL(His6) variant (Fig. 6). Because SsrA tagging of other cellular proteins was still observed, this effect of azetidine-2-carboxylic acid cannot be explained as a general nonspecific suppression of tagging (Fig. 6). Incorporation of /H9253-thiaproline into YbeL(His6) reduced the level of tagging compared with the proline control, whereas 3,4-dehydroproline incorporation increased the level of tagging (Fig. 6). Increased full-length tagging with 3,4-dehydroproline was confirmed by peptide analysis which showed a higher proportion of tagged C-terminal peptide in the 3,4-dehydroproline sample compared with the proline control (cf. peak 1 (untagged) and peak 2 (tagged) in Fig. 5B; Table III).

**DISCUSSION**

The results presented here show that several factors contribute to tagging of full-length YbeL by SsrA, including stop codon identity, release factor activity, and the amino acid sequence of the nascent peptide. These observations are consistent with a model in which inefficient translation termination at the ybeL stop codon permits SsrA to compete with RF-1 and RF-2 for entry into the ribosomal A-site, leading to higher levels of full-length protein tagging. The RF-1 and RF-2 proteins are present in ~10-fold and 50-fold excess, respectively, over SsrA RNA in *E. coli* during log-phase growth (28). The affinities of these protein release factors for the A-site depends on the identities of the stop codon and following nucleotide (29). Because SsrA does not appear to have an anticodon stem-loop (30, 31), we assume that its affinity for the A-site is independent of the stop signal. As a result, the competition model predicts the observed dependence of full-length YbeL tagging on stop-codon identity and release factor levels. Moreover, if peptidyl-tRNA hydrolysis is slow for YbeL, then several rounds of release factor association and dissociation may be required before normal termination of translation. This would allow SsrA multiple opportunities to compete for A-site binding.

In principle, the release factor/SsrA competition model should operate for termination at any stop codon and, by itself, does not explain why full-length YbeL is tagged at such a high level. Several lines of evidence indicate that the last two residues of the nascent YbeL polypeptide also play a critical role in determining the level of tagging. High levels of tagging occurred only when proline occupied the C-terminal position, irrespective of the proline codon and decoding tRNAPro species used. Moreover, the incorporation of certain proline analogs, such as azetidine-2-carboxylic acid, led to dramatic changes in the levels of YbeL tagging. Because the same mRNA codons and tRNAPro molecules direct the incorporation of proline and its analogs, this result provides the strongest evidence that it is the C-terminal proline, rather than mRNA or tRNA determinants, that causes full-length protein tagging. The identity of the penultimate residue of the nascent YbeL chain was also
important for full-length tagging, with Asp, Glu, Pro, Ile, and Val leading to the highest tagging levels.

How might the identity of the C-terminal residues of the nascent polypeptide affect translation termination and/or SsrA tagging? Interactions mediated by these amino acids could decrease the affinity of release factors for the A-site, increase the affinity of SsrA for the A-site, or slow hydrolysis of the peptidyl-tRNA bond. The ester bond of prolyl-tRNA hydrolyzes more rapidly than other aminoacyl-tRNAs (32), suggesting that the YbeL-Pro-tRNAPro linkage should not be more resistant to hydrolysis from a chemical perspective. Although this observation does not rule out the slow hydrolysis model, it favors versions that focus on conformational effects of the C-terminal residues in positioning the ester bond for catalyzed hydrolysis. Effects mediated via release factors are supported by studies that conclude that the nascent peptide chain can regulate ribosome activity, including translation termination (reviewed in Ref. 33). Indeed, Isaksson and co-workers (34–36) have used termination read-through assays to show that the last two residues of proteins influence translation termination, with the Asp-Pro sequence resulting in the most severe inhibition of E. coli translation termination. Although no read-through or frameshift products were observed in the studies reported here, the C-terminal Asp-Pro dipeptide resulted in one of the highest levels of SsrA tagging of full-length YbeL.

Proline is unique among naturally occurring protein residues in being an imino acid and having its dihedral angle constrained to roughly 60° by the pyrrolidine ring. Neither of these properties by themselves explain why proline causes tagging because the azetidine analog (Aze), which does not cause tagging, also lacks an amide proton and has a constrained angle similar to proline (37, 38). Is tagging related to the fact that the peptide bond of Xaa-Pro dipeptides can isomerize from the trans to the cis conformation at faster rates than other bonds?

**Table III**

Proline analog incorporation into C-terminal YbeL(His6) tryptic peptides

| HPLC peak | Predicted peptide | Measured mass | Calculated average mass |
|-----------|-------------------|---------------|------------------------|
| Proline   |                  | 1312.53       | 1312.4                 |
| Azetidine-2-carboxylate | H₃PFE(P) or H₆(Az)FEP | 1328.55       | 1328.4                 |
| γ Thiaproline | H₃(P)FEP or H₆(tP)FEP | 1330.64       | 1330.5                 |
| 3,4-Dehydroproline | H₆(dP)FEP | 1308.52       | 1308.4                 |

**Fig. 6.** Proline analogs affect tagging of YbeL(His6). YbeL(His6) was expressed in pro cells grown in the presence of proline or proline analogs. SsrA(DD)-tagged YbeL(His6) in cell lysates or following Ni²⁺-NTA purification was analyzed by Western blot and SDS-PAGE. The position of SsrA(DD)-tagged YbeL is indicated on the Western blot, and the position of YbeL(His6) is indicated on the Coomassie Blue-stained gel. The major protein below YbeL(His6) in the lysate prepared from cells grown in azetidine-2-carboxylic acid had an N-terminal sequence identical to the E. coli IbpA/IbpB proteins. Position of molecular standards (in kDa) are indicated along the left margin. Pro, proline; Az, azetidine-2-carboxylic acid; Thia, γ-thiaproline; and Dehyd, 3,4-dehydroproline.

SsrA Tagging Induced by Nascent Chains
dipeptides lacking proline? *Cis-trans* isomerization in solution occurs on a time scale of minutes, whereas translation termination is complete in a few seconds (11). Moreover, the identity of the residue before proline influences isomerization (39) in a fashion that is not correlated with the effects of the (−2) residue on YbeL tagging. For example, Tyr-Pro, Phe-Pro, and Trp-Pro peptide bonds are most likely to adopt the *cis* conformation in solution (39) but do not result in the highest levels of YbeL tagging. Finally, both the isomerization rate and equilibrium population of the cis isomer is higher for Xaa-Aze than Xaa-Pro peptide bonds (40), and yet Xaa-Aze dipeptides at the C terminus of the nascent chain do not cause tagging. These observations argue against isomerization being an important determinant of full-length tagging, but do not rule out this model, because *cis-trans* isomerization could be very different in solution than on the ribosome. Energy calculations show that Xaa-Aze and Xaa-Pro have roughly similar conformational preferences but the smaller four-member ring of azetidine is predicted to lead to slightly more flexibility (37, 38). This additional flexibility may permit the Xaa-Aze nascent chains to avoid sterice clashes with release factors or help position the ester bond for hydrolysis. The role of the (−2) residue in Pro-dependent tagging may also result from conformational effects on the C-terminal residue and/or from direct interactions between the (−2) side chain and the termination machinery.

In our studies, Asp-Pro and Pro-Pro resulted in the highest relative levels of tagged to untagged YbeL protein. For the Pro-Pro sequence, for example, roughly 40% of the total YbeL protein was tagged (Fig. 4A). Asp-Pro and Pro-Pro are highly underrepresented at the C terminus of proteins in most bacteria (Table IV; Supplemental Material). In *E. coli*, only one of 4355 proteins ends with Asp-Pro (10 expected) and none end with Pro-Pro (7 expected). Among 23,558 proteins in the *Bacillus/Clostridium* family, just one protein ends with Asp-Pro (43 expected) and four end with Pro-Pro (20 expected). Some bacteria, however, show only a slight bias against these C-terminal dipeptides. In the gamma proteobacterium *Xylella fastidiosa*, for example, six proteins end with Asp-Pro (8 expected) and six terminate with Pro-Pro (8 expected). In fact, Xaa-Pro dipeptides are found more frequently at the C terminus (5.7%) than at internal positions (5.0%) in *X. fastidiosa*. By contrast, significant biases against C-terminal Xaa-Pro sequences are observed for many other gamma proteobacteria, including *E. coli* (2.9% terminal/4.5% internal), *Buchnera* sp. APS (0.7% terminal/3.0% internal), *Haemophilus influenzae* (1.7% terminal/3.7% internal), and *Pasteurella multocida* (1.9% terminal/3.9% internal). The translation-termination machinery or SsrA system of *X. fastidiosa* may have evolved to avoid interactions with C-terminal prolines in nascent polypeptides. Interestingly, archaeabacteria, which have no SsrA system and have protein release factors that are unrelated to the bacterial RF-1 and RF-2 proteins, show no systematic biases against C-terminal prolines.

In previous work (10), we showed that tagging of *E. coli* ribokinase (RbsK) at or near the C terminus depended on the presence of a rare arginine codon at the mRNA position encoding the C-terminal residue, as well as the levels of the cognate charged tRNA*Arg*, and the efficiency of the terminal signal. In unpublished work, we have found that RbsK tagging increases in a RF-3 deletion strain. Thus, tagging of both YbeL and RbsK has the same dependence on the stop codon and the levels of release factors, as would be expected if SsrA competes with release factors. However, RbsK tagging depends in large part on the scarcity of the rare tRNA whereas full-length YbeL tagging is largely dependent on the chemical nature of the C-terminal residues of the nascent chain. Full-length tagging of the λ CI repressor (12) and UDP-galactose-4-epimerase have also been observed but neither of these systems involves a C-terminal proline or a rare C-terminal codon. Perhaps other C-terminal amino acid sequences can also influence tagging via interactions with release factors or SsrA.

Full-length protein tagging seems inherently wasteful. Why should some fraction of protein synthesis result in SsrA-tagged proteins that are rapidly degraded shortly after release from the ribosome? In a few cases, this may be an inevitable consequence of a functional requirement for a specific C-terminal sequence like Asp-Pro or Pro-Pro. In other instances, such as regulatory leader peptides, the full-length protein/peptide may have no function other than influencing downstream gene expression and thus recycling of its amino acids would be beneficial. Three of the ten leader peptides in *E. coli* end with Xaa-Pro sequences, and another leader coding sequence ends with two consecutive rare arginine codons. Still other proteins may use inefficient translation termination as a final opportunity to regulate gene expression at the level of protein synthesis and degradation.

Acknowledgments—We thank Drew Endy, Sean Moore, and Peter Chivers for helpful discussions.

REFERENCES

1. Muto, A., Ushida, C., and Himeno, H. (1998) *Trends Biochem. Sci.* 23, 25–29
2. Karzai, A. W., Roche, E. D., and Sauer, R. T. (2000) *Nat. Struct. Biol.* 7, 449–455
3. Tu, G. F., Reid, G. E., Zhang, J. G., Moritz, R. L., and Simpson, R. J. (1995) *J. Biol. Chem.* 270, 9322–9326
4. Keiler, K. C., Waller, P. R., and Sauer, R. T. (1996) *Science* 271, 990–993
5. Gottesman, S., Roche, E., Zhou, Y., and Sauer, R. T. (1998) *Genes Dev.* 12, 1338–1347
6. Herman, C., Thevenet, D., Boulouc, P., Walker, G. C., and D’Ari, R. (1998) *Genes Dev.* 12, 1348–1355
7. Williams, K. A., Martin, D. A., and Bartel, D. P. (1999) *EMBO J.* 18, 5423–5430
8. Roche, E. D. and Sauer, R. T. (1999) *EMBO J.* 18, 4579–4589
9. Abe, T., Inada, T., Ogawa, K., and Aiba, H. (2000) *EMBO J.* 19, 3762–3769
10. Hayes, C. S., Bose, B., and Sauer, R. T. (2002) *Proc. Natl. Acad. Sci. U. S. A.* 99, 3440–3445
11. Freistroffer, D. V., Kwiatkowski, M., Buckingham, R. H., and Ehrenberg, M. (2000) *Proc. Natl. Acad. Sci. U. S. A.* 97, 2046–2051
12. Roche, E. D., and Sauer, R. T. (2001) *J. Biol. Chem.* 276, 28509–28515
13. Poole, E. S., Brown, C. M., and Tate, W. P. (1995) *EMBO J.* 14, 151–158
14. Tate, W. P., and Manning, S. A. (1996) *Mol. Microbiol.* 21, 213–219
15. Datsenko, K. A., and Wanner, B. L. (2000) *Proc. Natl. Acad. Sci. U. S. A.* 97, 6640–6645
16. Arkov, A. L., Korolev, S. V., and Kisselev, L. L. (1995) *Nucleic Acids Res.* 23, 5991–5997
17. Zhang, S., Ryder-Aulin, M., and Isaksson, L. A. (1996) *J. Biol. Chem.* 271, 98–107
18. Zhang, S., Ryder-Aulin, M., and Isaksson, L. A. (1998) *J. Mol. Biol.* 284, 1243–1246
19. Zhang, S., Ryder-Aulin, M., and Isaksson, L. A. (1999) *FEBS Lett.* 455, 355–358
20. Dincbas-Renqvist, V., Engstrom, A., Mora, L., Heurige-Hamard, V., and cucumber, R. (2000) *EMBO J.* 19, 6905–6907
21. Freistroffer, D. V., Pavlov, M. Y., MacDougall, J., Buckingham, R. H., and Ehrenberg, M. (1997) *EMBO J.* 16, 4126–4133
22. Zavialov, A. V., Buckingham, R. H., and Ehrenberg, M. (2001) *Cell* 107, 115–124
23. Crawford, D.J.-G., Ito, K., Nakamura, Y., and Tate, W. P. (1999) *EMBO J.* 18, 727–732
24. Pavlov, M. Y., Freistroffer, D. V., MacDougall, J., Buckingham, R. H., and Ehrenberg, M. (1999) *EMBO J.* 18, 4334–4341
25. Janosi, L., Mottagui-Tabar, S., Isaksson, L. A., Sekine, Y., Ohtsubo, E., Zhang, S., Goun, S., Nelken, S., Shuda, M., and Kaji, A. (1998) *EMBO J.* 17, 1141–1151
26. Bouda, N., Minks, C., Medrano, F. J., Lutz, J., Huber, R., and Moroder, L. (1998) *Proc. Natl. Acad. Sci. U. S. A.* 95, 455–459
27. Grant, M. M., Brown, A. S., Corwin, L. M., Troxler, R. F., and Franzblau, S. (1975) *Biochim. Biophys. Acta* 302, 180–187
28. Adamski, F. M., McCaughan, K. K., Jorgensen, F., Kurland, C. G., and Tate, W. P. (1984) *J. Mol. Biol.* 183, 352–358
29. Pavlov, M. Y., Freistroffer, D. V., Dinchas, V., MacDougall, J., Buckingham, R. T. Sauer, unpublished results.
SsrA Tagging Induced by Nascent Chains

30. Felden, B., Himeno, H., Muto, A., McCutcheon, J. P., Atkins, J. F., and Gesteland, R. F. (1997) RNA (N. Y.) 3, 89–103
31. Williams, K. P., and Bartel, D. P. (1996) RNA (N. Y.) 2, 1306–1310
32. Hentzen, D., Mandel, P., and Garel, J.-P. (1972) Biochim. Biophys. Acta 281, 228–232
33. Lovett, P. S., and Rogers, E. J. (1996) Microbiol. Rev. 60, 366–385
34. Mottagui-Tabar, S., Bjornsson, A., and Isaksson, L. A. (1994) EMBO J. 13, 249–257
35. Bjornsson, A., Mottagui-Tabar, S., and Isaksson, L. A. (1996) EMBO J. 15, 1696–1704
36. Mottagui-Tabar, S., and Isaksson, L. A. (1997) FEBS Lett. 414, 165–170
37. Zagari, A., Nemethy, G., and Scheraga, H. A. (1990) Biopolymers 30, 951–959
38. Zagari, A., Nemethy, G., and Scheraga, H. A. (1990) Biopolymers 30, 961–966
39. Reimer, U., Scherer, G., Drewello, M., Kruber, S., Schutkowski, M., and Fischer, G. (1996) J. Mol. Biol. 279, 449–460
40. Kern, D., Schutkowski, M., and Drakenberg, T. (1997) J. Am. Chem. Soc. 119, 8403–8408
