The C-terminal amino acid sequence of nascent peptide is a major determinant of SsrA tagging at all three stop codons

T. Sunohara, T. Abo, T. Inada and H. Aiba

RNA 2002 8: 1416-1427

RNA (2002), *8*:1416–1427. Cambridge University Press. Printed in the USA.

DOI: 10.1017/S1355838202020198

The C-terminal amino acid sequence of nascent peptide is a major determinant of SsrA tagging at all three stop codons

TAKAFUMI SUNOHARA, TATSUHIKO ABO, TOSHIFUMI INADA, and HIROJI AIBA
Division of Biological Science, Graduate School of Science, Nagoya University, Chikusa, Nagoya 464-8602, Japan

ABSTRACT

Recent studies on endogenous SsrA-tagged proteins have revealed that the tagging could occur at a position corresponding to the normal termination codon. During the study of SsrA-mediated LacI tagging (Abo et al., *EMBO J*, 2000 *19*:3762–3769), we found that a variant LacI (LacIΔC1) lacking the last C-terminal amino acid residue is efficiently tagged in a stop codon-dependent manner. SsrA tagging of LacIΔC1 occurred efficiently without LacI binding to the *lac* operators at any one of three stop codons. The C-terminal (R)LESG peptide of LacIΔC1 was shown to trigger the SsrA tagging of an unrelated protein (CRP) when fused to its C terminus. Mass spectrometry analysis of the purified fusion proteins revealed that SsrA tagging occurs at a position corresponding to the termination codon. The alteration of the amino acid sequence but not the nucleotide sequence of the C-terminal portion eliminated the tagging. We also showed that the tagging-provoking sequences cause an efficient translational readthrough at UGA but not UAA codons. In addition, we found that C-terminal dipeptides known to induce an efficient translation readthrough could cause an efficient tagging at stop codons. We conclude that the amino acid sequence of nascent polypeptide prior to stop codons is a major determinant for the SsrA tagging at all three stop codons.

Keywords: nascent peptide; stop codon; tmRNA; *trans*-translation; translation readthrough

INTRODUCTION

Bacterial SsrA RNA, also known as tmRNA or 10Sa RNA, is a central player in the cotranslational tagging process called *trans*-translation (Keiler et al., 1996; Karzai et al., 2000). When a ribosome stalls on an mRNA during translation, SsrA RNA charged with alanine is recruited to the ribosome together with an associated protein, SmpB, and donates an alanine to the growing polypeptide chain by acting as an alanyl-tRNA. The SsrA RNA then serves as an mRNA to direct the addition of a short peptide tag to the C terminus of the nascent polypeptide. The tagged polypeptide is degraded by several ATP-dependent proteases (Keiler et al., 1996; Gottesman et al., 1998; Herman et al., 1998). The major biological roles of the SsrA system are believed to rescue stalled ribosomes and to eliminate incomplete polypeptides (Keiler et al., 1996; Karzai et al., 2000). More specific aspects regarding biological functions of the SsrA system have been addressed in recent studies (Abo et al., 2000, 2002; Muto et al., 2000; De La Cruz & Vioque, 2001; Ranquet et al., 2001; Ueda et al., 2002; Withey & Friedman, 2002).

The SsrA-mediated protein tagging occurs extensively in *Escherichia coli* cells (Roche & Sauer, 1999, 2001; Abo et al., 2000). However, it remains to be elucidated how a large number of endogenous proteins are tagged by the SsrA system and what mRNA features are responsible for the SsrA-mediated protein tagging in living cells. There are several situations where mRNAs become targets for the SsrA system and protein tagging occurs (Gillet & Felden, 2001). An obvious target for the SsrA system is the 3′ end of a truncated mRNA lacking an in-frame stop codon where the ribosome stalling may occur due to the lack of the normal translation termination signal (Keiler et al., 1996). The truncated mRNA could be generated either by nuclease cleavages of an mRNA or by incomplete transcription. The SsrA-mediated tagging of LacI typically represents this situation because LacI binding to the *lac* operators could block the completion of *lac* transcription resulting in truncated *lac* mRNAs (Abo et al., 2000). A ribosome would also reach the 3′ end of an intact mRNA when a normal stop codon is erroneously translated either in the presence of nonsense suppressor tRNAs (Ueda et al., 2002) or in the presence of

Reprint requests to: Hiroji Aiba, Division of Biological Science, Graduate School of Science, Nagoya University, Chikusa, Nagoya 464-8602, Japan; e-mail: i45346a@nucc.cc.nagoya-u.ac.jp.

miscoding drugs (Abo et al., 2002). The SsrA system appears to also act at a run of rare codons on an mRNA where ribosome stalling may occur due to the deficiency of cognate aminoacyl-tRNAs, although it remains obscure whether the ribosome stalling itself leads to tagging at internal sites of the mRNA or if stalling leads to tagging by generating somehow truncated mRNAs lacking a stop codon (Roche & Sauer, 1999; Karzai et al., 2000).

Recent studies on endogenous SsrA-tagged proteins in *E. coli* showed that the SsrA tag can be added immediately after the normal C-terminal residue of the protein in at least several cases (Roche & Sauer, 2001; Collier et al., 2002; Hayes et al., 2002a). Thus, the ribosome stalling and SsrA tagging could occur at a position corresponding to the normal termination codon in certain conditions. It has been proposed that the combination of a rare arginine codon at the C terminus and the adjacent inefficient UGA termination codon is responsible for SsrA tagging at stop codons (Hayes et al., 2002a). In this article, we report a new finding regarding an important determinant of SsrA tagging at all three stop codons. During the study of SsrA-mediated LacI tagging (Abo et al., 2000), we found that a variant LacI (LacIΔC1) lacking the last C-terminal amino acid residue was efficiently tagged at the stop codon by the SsrA system. Neither an inefficient UGA codon nor an adjacent rare codon is required for the LacIΔC1 tagging. Instead, we demonstrate that the amino acid sequence of the nascent polypeptide prior to a stop codon is critical for an efficient SsrA tagging at stop codons.

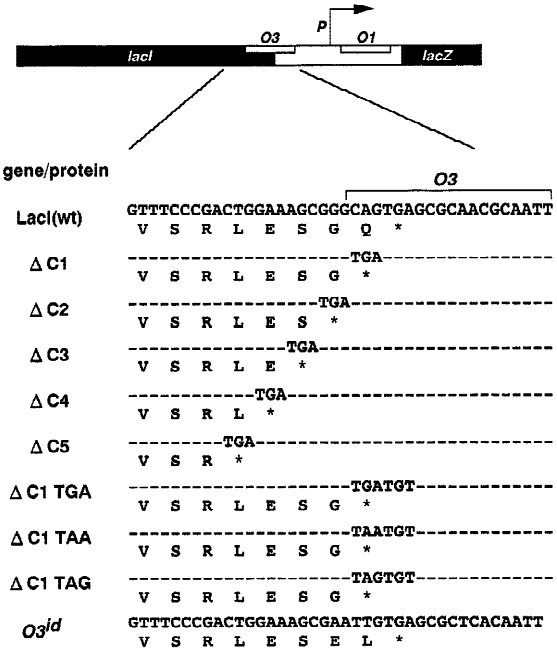

FIGURE 1. Nucleotide and amino acid sequences of the C-terminal region of the *lacI* variants. Plasmid pIT613 carrying the *lacIZYA* with the $lacI^q$ promoter on pBR322 is the parent plasmid for a series of plasmids carrying *lacI* gene variants encoding truncated LacI. The mutated codons and/or sequences are shown together with the amino acid sequences of LacI variant proteins. pTA100 is the derivative of pIT613, which has an ideal operator sequence at the position of *O3* (Abo et al., 2000).

RESULTS

A variant LacI (LacIΔC1) is strongly tagged by the SsrA system

The *E. coli* Lac repressor (LacI) is the first endogenous protein for which SsrA tagging was demonstrated (Abo et al., 2000). The observation that LacI tagging was dependent on the binding of LacI to the *lac* operators led us to propose that LacI–operator interaction blocks the completion of *lacI* transcription, resulting in truncated *lacI* mRNAs that are, in turn, recognized by the SsrA system. Although the precise tagging sites in LacI have not been determined yet, it was shown by Roche and Sauer (2001) that the tagging must occur within the C-terminal nine residues of LacI. To gain further insight into the mechanism of LacI tagging, we constructed a series of derivatives of pIT613 carrying *lac* variants in which the UGA stop codon replaced one of the last five C-terminal codons, as shown in Figure 1. Plasmid pIT613 carries the entire *lac* region under the control of $lacI^q$ promoter (Abo et al., 2000) and the pIT613 derivatives carry the *lac* gene variants encoding truncated LacI proteins lacking 1 to 5 amino acid residues at its C terminus. Each of the LacI variants was coexpressed with $SsrA^{DD}$ RNA and tested for its tagging property either in the presence or absence of IPTG by western blotting (Fig. 2).

As reported previously (Abo et al., 2000), the IPTG-sensitive SsrA tagging of LacI was typically observed when pIT613 was introduced into *ssrA* cells expressing $SsrA^{DD}$ RNA (Fig. 2A, B, lanes 1 and 2). When *O3* was replaced by the ideal operator, the tagging was markedly increased and it was still IPTG sensitive (Fig. 2A, B, lanes 3 and 4). In the absence of IPTG, all of the truncated LacI proteins were tagged to various extents whereas the tagging of the truncated LacI proteins was markedly reduced in the presence of IPTG except for LacIΔC1 (Fig. 2A, B, lanes 5–14). The fact that these truncated proteins including a ΔC5 variant are SsrA-tagged like wild-type LacI implies that transcription termination occurs upstream of these codons. In fact, it was shown by Roche and Sauer (2001) that the tagging must occur within the last nine residues of LacI, although the precise tagging sites in LacI remain to be determined yet. It is clear that both the ΔC4 and ΔC5 variants are subject to IPTG-suppressible SsrA tagging. A striking observation was that a variant LacI (LacIΔC1) lacking the last C-terminal amino acid residue was strongly tagged and the tagging showed only a slight reduction in the presence of IPTG (Figs. 2A, B,

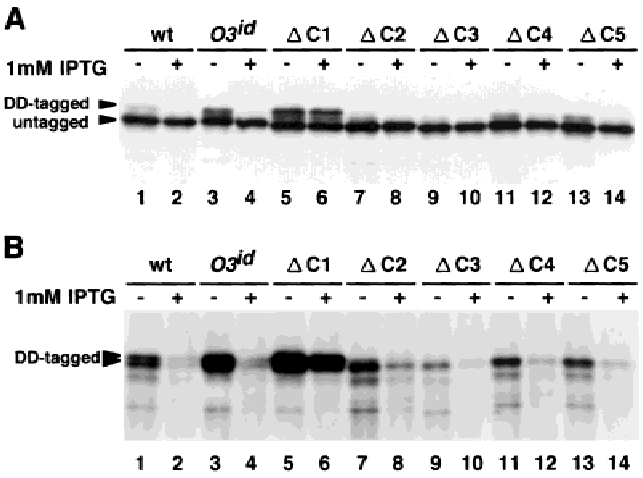

FIGURE 2. SsrA tagging of variant LacI proteins. TA411 (Δ*lacIZYA*, *ssrA*^{DD}) cells harboring pIT613 (lanes 1 and 2), pTA100 (lanes 3 and 4), pIT604 (lanes 5 and 6), pIT605 (lanes 7 and 8), pIT607 (lanes 9 and 10), pIT606 (lanes 11 and 12), and pIT608 (lanes 13 and 14) were grown in the presence (lanes 2, 4, 6, 8, 10, 12, and 14) or absence (lanes 1, 3, 5, 7, 9, 11, and 13) of 1 mM of IPTG. Total extracts equivalent to OD_{600} = 0.01 were analyzed by western blotting using anti-LacI (**A**) or anti-DD-tag (**B**) antibodies.

lanes 5 and 6). In fact, the DD-tagged LacIΔC1 was produced at more than 30% of the level of untagged LacIΔC1 as judged by the intensities of bands in western blots. These results suggest that the mechanism responsible for LacIΔC1 tagging is different from that of the wild-type LacI tagging.

The LacIΔC1 tagging is independent on LacI binding to operators

The IPTG-resistant tagging of LacIΔC1 could be explained if LacIΔC1 protein somehow retains the ability to bind to operators even in the presence IPTG. To test this possibility, we determined the β-galactosidase activities in cells carrying pIT613 or pIT604 in the presence and absence of IPTG (Fig. 3). The β-galactosidase activity of cells expressing LacIΔC1 was significantly higher than that of cells expressing LacI in the absence of IPTG, suggesting that the LacIΔC1 is less active compared to the intact LacI. An important observation, however, is that the *lac* operon was efficiently induced in the presence of IPTG in cells expressing LacIΔC1 protein as in cells expressing the normal LacI. This implies that LacIΔC1 binding to operators was eliminated by IPTG, suggesting that the SsrA tagging of LacIΔC1 is not due to the LacI binding to operators. In fact, the SsrA tagging of LacIΔC1 protein occurred efficiently even when the *lac* operators were deleted (data not shown).

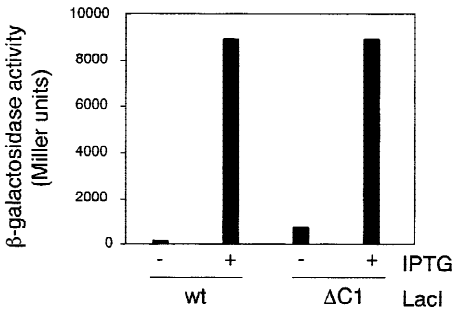

FIGURE 3. Effect of LacI variants on the *lac* operon expression. MC4100Δ*ssrA* cells harboring pIT613 and pIT604 were grown in LB medium. The cultures were divided into two parts at OD_{600} = 0.2 and 0.4 mM IPTG (final concentration) were added to one aliquot. The cells were grown for another 2 h and β-galactosidase activities were measured.

Any one of three stop codons at the –1 position of LacI is sufficient for SsrA tagging of LacIΔC1

How does the SsrA tagging of LacIΔC1 occur without the LacI–operator interaction? The results mentioned above indicate that the UGA codon introduced at position −1 of LacI may be responsible for the SsrA tagging of LacIΔC1. It should be noted that the variant *lacIΔC1* gene encoding LacIΔC1 possesses two consecutive UGA codons (Fig. 1). It is possible that the existence of tandem UGA codons somehow causes the SsrA tagging of LacIΔC1 protein. To test this possibility, the original UGA stop codon of LacI was replaced with a sense codon to construct pST201 (Fig. 1). The SsrA tagging of LacIΔC1 was examined in cells carrying pST201 in the presence of IPTG by western blotting. The substitution of the original UGA stop codon with a sense codon did not affect the tagging of LacIΔC1 (Fig. 4, lane 4), indicating that only one UGA

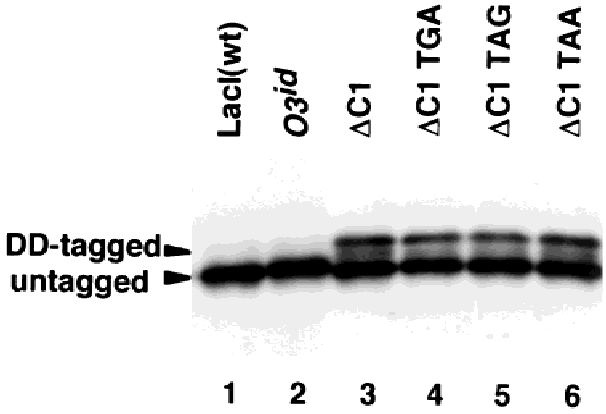

FIGURE 4. Effect of the stop codon nature on the SsrA tagging of LacIΔC1. TA411 (Δ*lacIZYA ssrA*^{DD}) cells harboring pIT613 (lane 1), pTA100 (lane 2), pIT604 (lane 3), pST201 (lane 4), pST202 (lane 5), and pST203 (lane 6) were grown in LB medium containing 1 mM IPTG. Total extracts equivalent to OD_{600} = 0.01 were analyzed by western blotting using anti-LacI antibodies.

codon at position −1 of LacI is sufficient to cause the efficient tagging of LacIΔC1. To examine how the identity of the stop codon influences the SsrA tagging of LacIΔC1, the UGA stop codon of LacIΔC1 on pST201 was changed to either UAA or UAG to construct plasmid pST202 or pST203, respectively (Fig. 1). These plasmids were also introduced in cells expressing $SsrA^{DD}$ RNA, and the SsrA tagging of LacIΔC1 in the presence of IPTG was examined by western blotting. The alteration of the UGA codon to UAA or UAG caused little difference in the SsrA tagging of LacIΔC1 in the presence of IPTG (Fig. 4, lanes 5 and 6). Thus, the stop codon-dependent SsrA tagging of LacIΔC1 occurs efficiently at any one of the three stop codons.

The C-terminal portion of LacIΔC1 is sufficient to cause the stop codon-dependent tagging

To examine how amino acid and/or nucleotide sequences before the stop codon affect the stop codon-dependent SsrA tagging of LacIΔC1, we constructed a series of genes encoding variants of cAMP receptor protein (CRP) in which the last one to seven codons of LacIΔC1 were fused to the CRP ORF (Fig. 5A). These fusion proteins were coexpressed with $SsrA^{DD}$ RNA, and the stop codon-dependent SsrA tagging was examined by western blotting using anti-CRP and anti-DD-tag antibodies. No significant SsrA tagging of the

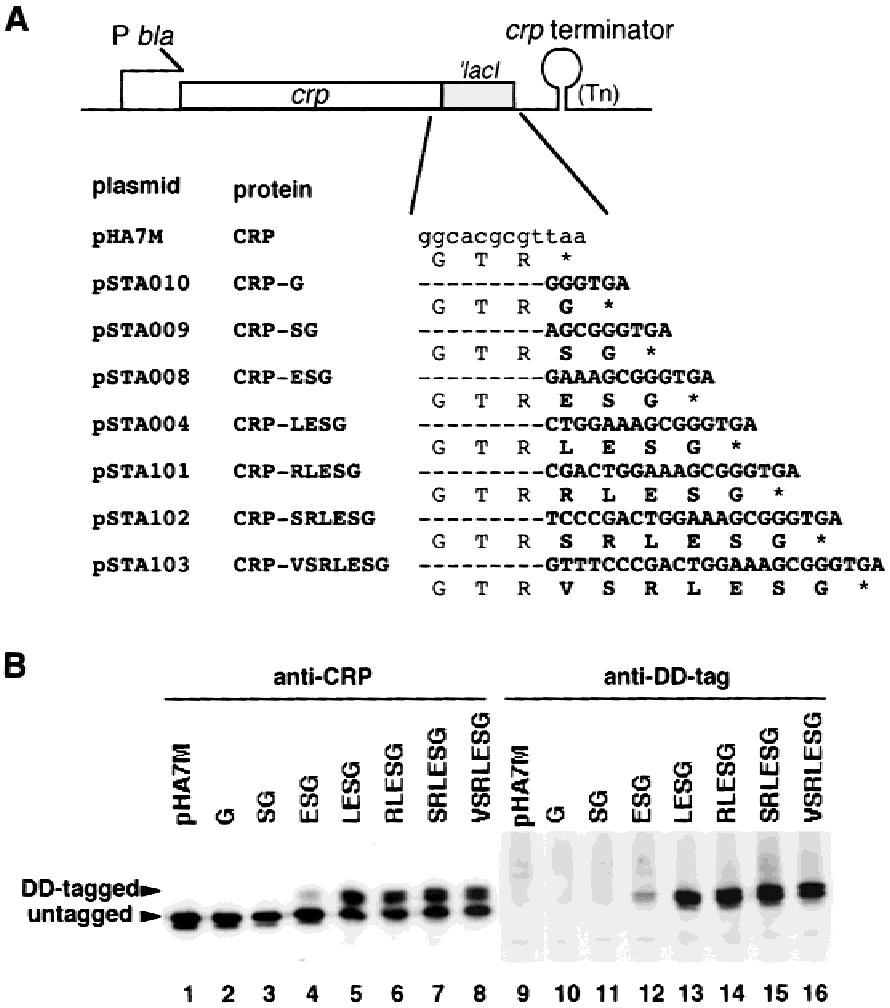

FIGURE 5. SsrA tagging of variant CRP proteins. **A:** Schematic drawing and sequences of the *crp* gene variants. The open box is the CRP ORF and the shaded box represents the sequence derived from the C terminus of LacIΔC1. The nucleotide sequences around the junction are given together with the amino acid sequences. The lowercase letters correspond to the nucleotide sequence of the original *crp* gene and bold uppercase letters correspond to those of the *'lacI* gene. The C-terminal amino acid sequence of wild-type CRP is GTR. The amino acid sequences corresponding to the C terminus of LacIΔC1 are shown in bold letters. All the fusion genes are followed by the *crp* terminator sequence with the same spacing. **B:** Western blot analysis of variant CRP proteins. TA481 (Δ*crp ssrA*^{DD}) cells harboring pHA7M (lanes 1 and 9), pSTA010 (lanes 2 and 10), pSTA009 (lanes 3 and 11), pSTA008 (lanes 4 and 12), pSTA004 (lanes 5 and 13), pSTA101 (lanes 6 and 14), pSTA102 (lanes 7 and 15), and pSTA103 (lanes 8 and 16) were grown in LB medium. Total extracts equivalent to OD_{600} = 0.01 were analyzed by western blotting using anti-CRP (lanes 1–8) and anti-DD-tag (lanes 9–16) antibodies.

fusion protein was observed when the last one or two codons of LacIΔC1 were fused to CRP ORF (Fig. 5B, lanes 2, 3, 10, and 11). However, a weak but significant SsrA tagging of the fusion protein was observed when the last three codons of LacIΔC1 were fused to CRP ORF (Fig. 5B, lanes 4 and 12). The tagging of the fusion protein markedly increased when the last four codons of LacIΔC1 were fused to CRP (Fig. 5B, lanes 5 and 13). The addition of more than five codons did not cause further increase in the tagging of the fusion protein (Fig. 5B, lanes 6–8 and 14–16). The efficiency of SsrA tagging of CRP variants possessing the last four or more codons of LacIΔC1 was comparable to that of LacIΔC1. We conclude that the C-terminal portion of LacIΔC1 could confer the property to be tagged by the SsrA system to an unrelated protein when fused to its C terminus. The data indicate that the tetra- or pentapeptide of the C-terminal region of LacIΔC1, LESG or RLESG, is responsible and sufficient for an efficient stop codon-dependent SsrA tagging of nascent peptide. The DD-tagged proteins apparently consist of two bands, suggesting that the tagging may occur at two sites around the stop codon.

SsrA tagging occurs at a position corresponding to the termination codon

To estimate the SsrA tagging sites in the LacIΔC1, we constructed a plasmid, pTA402, carrying a variant *crp* gene encoding His_6-CRP-LESG that contains a His_6 tag at the N terminus of CRP-LESG. The His_6-tagged proteins were expressed in cells carrying the *ssrA*^{DD} gene and purified by Ni^{2+}-NTA affinity chromatography. The purified proteins were subjected to SDS-polyacrylamide gel electrophoresis followed by Coomassie Brilliant Blue staining (Fig. 6A). The DD-tagged proteins consisting of two bands (bands II and III) were clearly detected along with untagged protein (band I). The purified proteins were analyzed by MALDI-TOF mass spectrometry. The signal at 24,635 Da corresponds to the "full-length" His_6-CRP-LESG (band I) whereas the broad signal of higher molecular weight represents the DD-tagged proteins (bands II + III; Fig. 6B). This broad signal has a peak at the position of 25,803 Da that corresponds to His_6-CRP-LESG sequence plus the 11-residues DD-tag. To determine more precisely the SsrA tagging sites, the three bands were excised from the gel and digested separately in-gel with lysyl endopeptidase that specifically cleaves the peptide bond after lysine residues. The eluted peptides were analyzed by MALDI-TOF mass spectrometry (Fig. 6C). The peptidase digestion of untagged band I gave a signal with a mass of 1,294.68 Da that corresponds to that expected for the C-terminal fragment size of His_6-CRP-LESG along with several other signals (Fig. 6C, top). When the lysyl endopeptidase digest of band III was analyzed, the 1,294.68-Da signal was no longer observed and a new signal with a mass of 2,489.69 Da that corresponds to a junction peptide containing the C-terminal fragment of His_6-CRP-LESG plus the tag appeared (Fig. 6C, bottom). All other signals observed in the digest of the untagged band I were also detected in the digest of band III. These results clearly indicate that the stop codon-dependent SsrA tagging of CRP-LESG occurs just after the last C-terminal residue. The additional DD-tagged band II that is slightly smaller than the DD-tagged band III suggests that the tagging may occur also near the stop codon. Although the lysyl endopeptidase digest of band II gave a new signal with a mass of 2,292.14 Da (data not shown), we failed to identify the precise junction peptide within a permissible error.

Amino acid sequence of the C-terminal end is the determinant for the stop codon-dependent SsrA tagging

It has been shown that both an inefficient UGA stop codon and an arginine rare codon prior to the UGA are responsible for the stop codon-dependent SsrA tagging of RbsK (Collier et al., 2002; Hayes et al., 2002a). Our results clearly indicate that these features cannot be responsible for SsrA tagging of LacIΔC1 because the tagging occurs even at a strong UAA codon and there is no rare codon prior to the stop codon. One possible determinant of the stop codon-dependent SsrA tagging of LacIΔC1 would be the amino acid sequence of the C-terminal end. It is also possible that the nature of codons and/or the nucleotide sequence corresponding to the C terminus of LacIΔC1 influences the stop-codon dependent SsrA tagging. To address these questions, all of the last four codons of a variant CRP expressed from pSTA004 were replaced by other synonymous codons to construct pST026 (Fig. 7A). This extensive alteration of the nucleotide sequence had little effect on the tagging of the full-length CRP-LESG, although the tagging of the lower band was moderately decreased (Fig. 7B, lanes 2 and 3). The result suggests that neither the nature of codons nor the nucleotide sequence before the stop codon are primarily responsible for the SsrA tagging at the position of the stop codon. We next examined how the changes in the amino acid sequence affect the SsrA tagging of variant CRP proteins. We first replaced the Gly at the −1 position of LacIΔC1 to an acidic Glu (Fig. 7A). This amino acid substitution dramatically reduced the SsrA tagging of the variant CRP (Fig. 7B, lane 5). The substitution of Gly at the −1 position to a basic (Arg) or hydrophobic (Leu) amino acid also completely eliminated the SsrA tagging of the variant CRP (data not shown). These results indicate that the Gly at −1 is an important element for the efficient SsrA tagging at the stop codon. We also found that some amino acid substitutions at −2 (e.g., Ser to Leu) essentially eliminated the SsrA

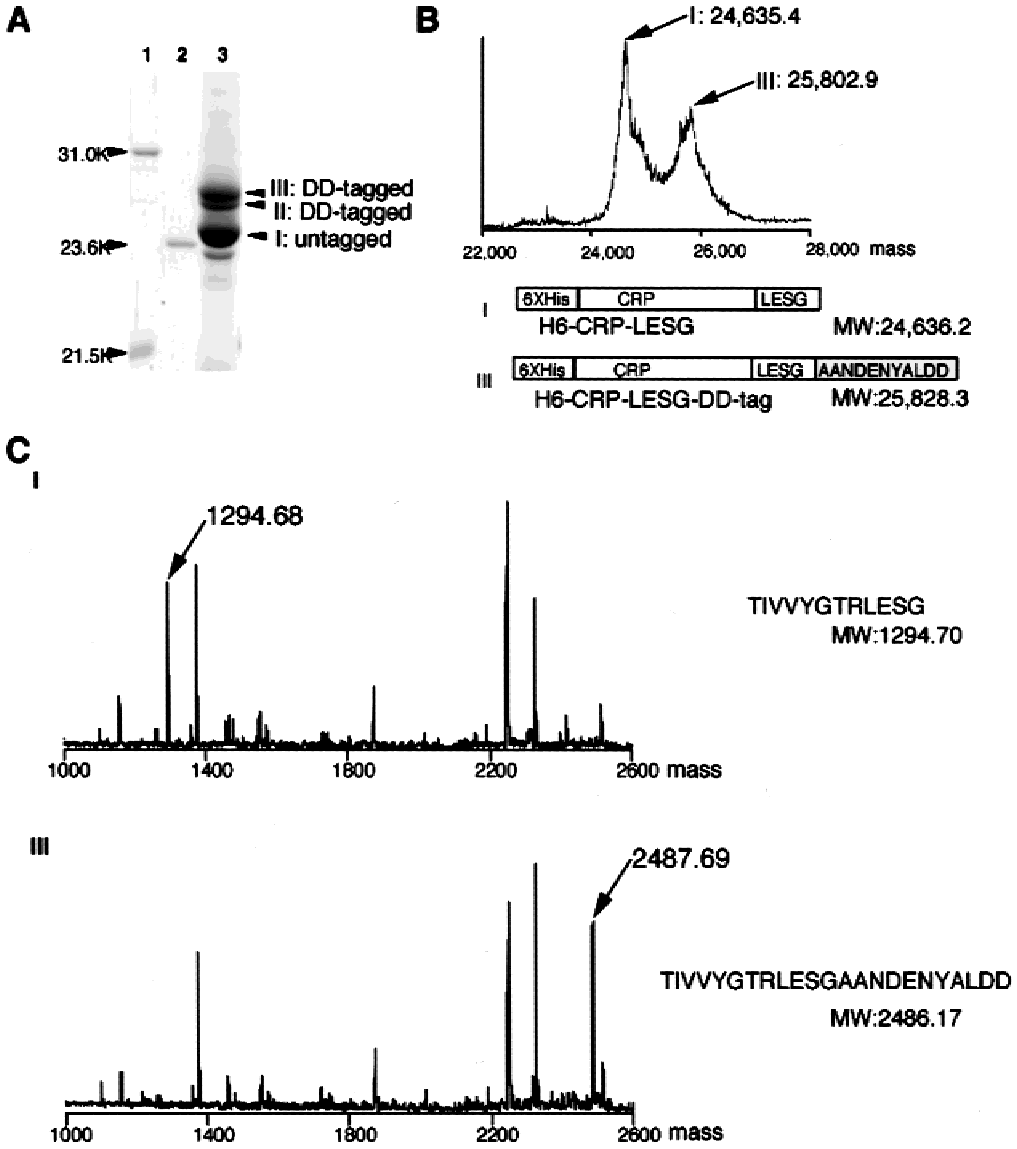

FIGURE 6. Identification of $SsrA^{DD}$ tagging sites in His_6-CRP-LESG. **A:** SDS-polyacrylamide gel electrophoresis analysis of purified His_6-CRP chimeric proteins. Proteins produced from TA481 cells harboring pTA402 were purified through Ni^{2+}-NTA agarose chromatography and analyzed by a 12% SDS-polyacrylamide gel electrophoresis followed by Coomassie Brilliant Blue staining (lane 3). Two DD-tagged proteins (bands II and III) and an untagged protein (band I) were detected. Lanes 1 and 2 represent protein markers and purified normal CRP protein, respectively. **B:** Mass spectrometry analysis of purified His_6-CRP chimeric proteins. About 1 μg of the purified proteins were subjected to mass spectrometry analysis. The schematic drawings of the His_6-CRP-LESG and His_6-CRP-LESG-DD-tag proteins deduced from mass spectrum peaks are shown in the right side along with the molecular weights estimated from amino acid sequences. **C:** Mass spectrometry analysis of His_6-CRP chimeric proteins after treatment with lysyl endopeptidase. Three bands (I, II, III) shown in **A** were cut out from the gel. The gel was treated with lysyl endopeptidase and subjected to mass spectrometry analysis. The data for bands I and III are shown. The signals that are expected to correspond to the C-terminal fragments are shown by arrows along with the observed mass. The peptide sequences of the C-terminal fragments of His_6-CRP chimeric proteins generated by lysyl endopeptidase digestion are shown in the right side along with the estimated molecular weights.

tagging (Fig. 7B, lane 4). In addition, certain amino acid substitutions at −3, −4, and −5 affected more or less the SsrA tagging depending on the amino acid substitution (data not shown). Taken together, we conclude that the amino acid sequence of the nascent pentapeptide is responsible for the stop codon-dependent SsrA tagging.

The (R)LESG sequence prior to the UGA codon induces translation readthrough

To test how the (R)LESG sequence upstream of he UGA stop codon affects the efficiency of translation termination/readthrough, we constructed fusion genes in which the IIAGlc ORF encoded by *crr* was fused to

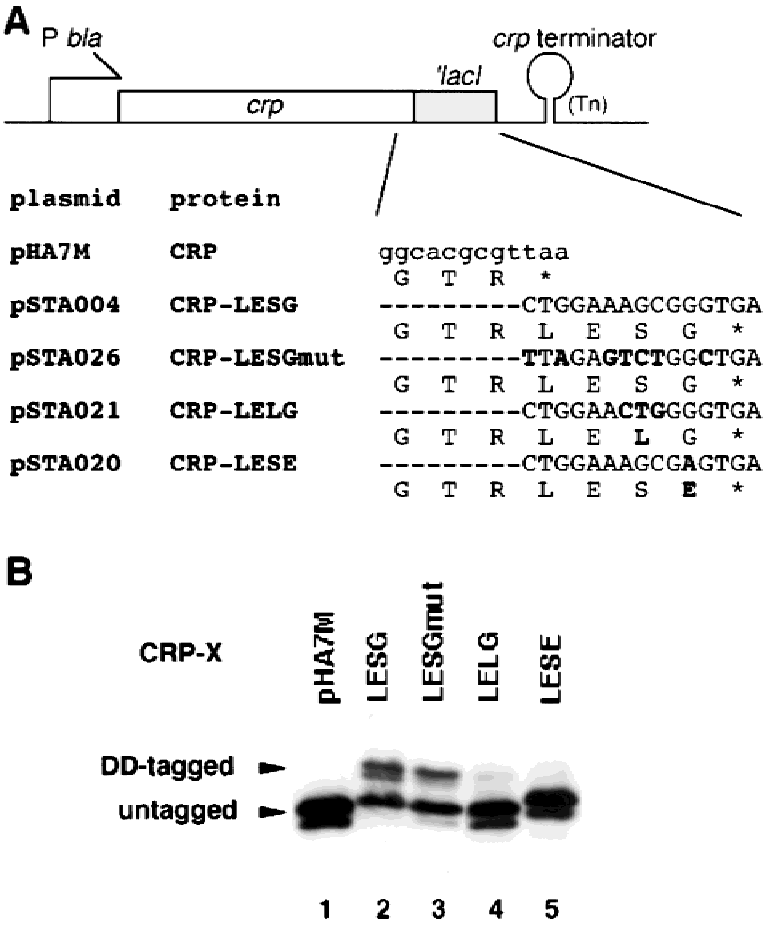

FIGURE 7. Effects of nucleotide and peptide sequences prior to stop codon on SsrA tagging. **A:** Schematic drawing and sequences of several variants of the *crp* gene used. Nucleotides and amino acid residues substitutions are indicated by bold letters. **B:** Western blot analysis of variant CRP chimeric proteins. TA481 (Δ*crp ssrA*^{DD}) cells harboring pHA7M (lane 1), pSTA004 (lane 2), pSTA026 (lane 3), pSTA021 (lane 4), and pSTA020 (lane 5) were grown in LB medium. Total extracts equivalent to OD_{600} = 0.01 were analyzed by western blotting using anti-CRP antibodies.

just after the UGA stop codon of the variant *crp* genes encoding either CRP-LESG or CRP-SG (Fig. 8A). This system can be used to monitor readthrough of the stop codon because each fusion construct gives a CRP-IIAGlc fusion protein as the result of translation readthrough. Only a little readthrough product was produced from the SG fusion construct in the absence of SsrA RNA (Fig. 8B, lane 2). The readthrough product markedly increased when the LESG fusion construct was expressed in the absence of SsrA RNA. As expected, a strong SsrA tagging of CRP was observed when the LESG fusion construct but not the SG fusion construct was expressed in the presence of SsrA^{DD} RNA (Fig. 8B, lanes 7 and 8). The presence of SsrA RNA^{DD} had only a small effect on the readthrough at the UGA stop codon. Thus, the SsrA tagging provoking C-terminal sequence apparently causes translation readthrough at the UGA codon. Similar experiments were performed with the fusion genes in which the UGA stop codon was replaced by UAA and UAG stop codons. The amount of readthrough product was maximum with UGA and minimum with UAA (Fig. 8B, lanes 1–12). These results are as expected because it is known that UAA gives the most efficient and UGA the least efficient translation termination (Tate & Mannering, 1996; Mottagui-Tabar & Isaksson, 1997). An important observation was that the replacement of the UGA stop codon with UAA or UAG again did not affect the efficiency of SsrA tagging of CRP-LESG (Fig. 8B, lanes 7–12). Thus, the (R)LESG sequence could lead to an efficient SsrA tagging even at strong termination codons where only a weak translational readthrough occurs.

It is known that translation readthrough at UGA varies depending on the last two amino acids of the C-terminal end of the nascent peptide (Mottagui-Tabar et al., 1994; Bjornsson et al., 1996; Mottagui-Tabar & Isaksson, 1997). For example, a strong translational readthrough occurs when the terminal dipeptide is Asp-Pro whereas the Arg-Pro dipeptide causes only a little readthrough (Mottagui-Tabar et al., 1994). It is interesting to test how the terminal Asp-Pro and Arg-Pro dipeptides influence the SsrA tagging. For this, we constructed fusion genes in which the IIAGlc ORF was fused to just after the UGA stop codon of the *crp* gene variants encoding either CRP-DP or CRP-RP (Fig. 8A). Extracts of cells carrying these fusion genes were analyzed by western blotting. As expected, the fusion gene encoding CRP-DP produced a high level of readthrough product (CRP-IIAGlc) in the absence of SsrA RNA whereas little CRP-IIAGlc was produced from the fu-

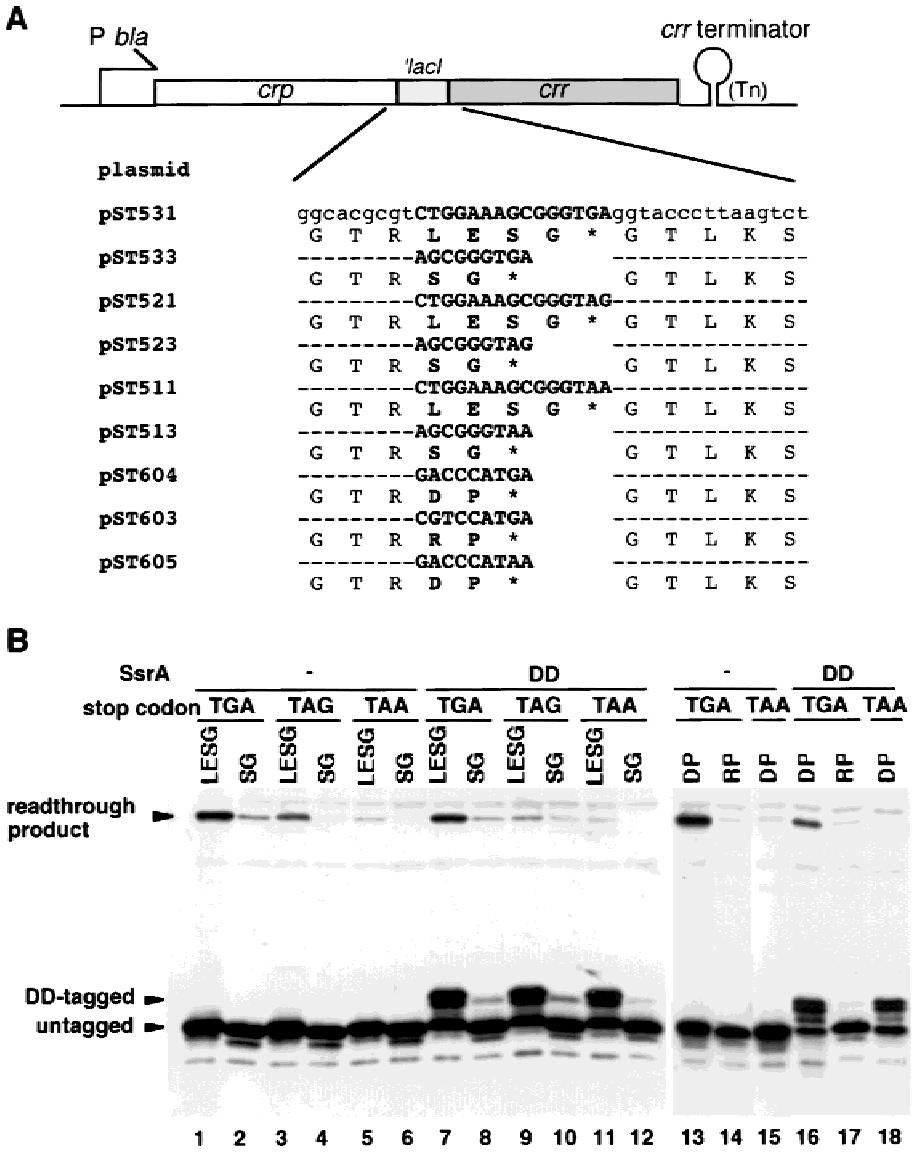

FIGURE 8. Translational readthrough and SsrA tagging at stop codons. **A:** Schematic drawing of *crp-'lacI-crr* fusion gene. Nucleotide and amino acid sequences of the junction region are shown. Translational readthrough at the *crp* stop codon results in the production of CRP-IIAGlc fusion protein. **B:** TA481 (Δ*crp*, *ssrA*^{DD}) cells harboring pST531 (lanes 1 and 7), pST533 (lanes 2 and 8), pST521 (lanes 3 and 9), pST523 (lanes 4 and 10), pST511 (lanes 5 and 11), pST513 (lanes 6 and 12), pST604 (lanes 13 and 16), pST603 (lanes 14 and 17), and pST605 (lanes 15 and 18) were grown in LB medium. Total extracts equivalent to OD_{600} = 0.01 were analyzed by western blotting using anti-CRP antibodies. CRP proteins with two or four amino-acid addition, their DD-tagged products, readthrough products are indicated by arrowheads.

sion gene encoding CRP-RP (Fig. 8B, lanes 13 and 14). Interestingly, CRP-DP was efficiently tagged in the presence of SsrA^{DD} RNA whereas no SsrA tagging of CRP-RP was observed (Fig. 8B, lanes 16 and 17). The presence of SsrA RNA^{DD} slightly reduced the readthrough at the UGA stop codon (Fig. 8B, lane 16). When the UGA stop codon of CRP-DP was replaced with UAA, the readthrough was dramatically reduced whereas SsrA-tagged CRP-DP was still efficiently produced. Taken together, we conclude that amino acid sequences prior to the stop codon that induce an efficient translation readthrough at the UGA codon could cause an efficient tagging of full-length protein not only at weak but also strong stop codons.

DISCUSSION

Several natural proteins including λ repressor, RbsK, and YbeL can be tagged at positions corresponding to the normal stop codons in *E. coli* cells (Roche & Sauer, 2001; Collier et al., 2002; Hayes et al., 2002a). It has been proposed that both an inefficient UGA termination codon and an arginine rare codon prior to the stop codon are responsible for the SsrA tagging at the stop

codons (Collier et al., 2002; Hayes et al., 2002a). We have found here that a variant LacI (LacIΔC1) lacking the last C-terminal amino acid residue is efficiently tagged at the stop codon by the SsrA system without the LacIΔC1–operators interaction. The major finding in the present study is that the C-terminal amino acid sequence of the nascent peptide is an important determinant for an efficient SsrA tagging at all stop codons. This conclusion is based on the following observations: (1) any one of the three stop codons could lead to the LacIΔC1 tagging even without any rare codon prior to the stop codon, (2) the C-terminal tetra- or pentapeptide of LacIΔC1 confers the ability to be tagged to an unrelated protein when fused to its C-terminus, and (3) alteration of the C-terminal amino acid but not nucleotide sequences eliminates the SsrA tagging.

Translation termination is mediated at the UAG stop codon by polypeptide release factor 1 (RF1), at UGA by release factor 2 (RF2), and at UAA by both RF1 and RF2 (Nakamura et al., 1996). When the ribosome reaches a termination codon in the mRNA, polypeptide release factors (RFs) recognize the stop codons in the ribosomal A-site to cause the release of the polypeptide by hydrolyzing the ester bond between the tRNA in the P-site and the polypeptide chain (Nakamura et al., 1996). The stop codons in the A-site can be also recognized by near-cognate or suppressor tRNAs resulting in translational readthrough. Thus, both RFs and near-cognate or suppressor tRNAs compete each other for the recognition of the stop codons at the A-site, which, in turn, determine the efficiency of translation termination. Among the three stop codons, UAA gives the most efficient and UGA the least efficient translation termination (Tate & Mannering, 1996; Mottagui-Tabar & Isaksson, 1997). It is also known that the efficiency of UGA readthrough is dependent on the base just after the stop codon (Poole et al., 1995). In addition, it has been demonstrated that the last two amino acids in the nascent peptide markedly influence translation readthrough (Mottagui-Tabar et al., 1994; Bjornsson et al., 1996; Mottagui-Tabar & Isaksson, 1997). Thus, the efficiency of translation termination or readthrough is affected by the nature of the stop codon and by mRNA and/or nascent peptide sequence context. The present study clearly indicates that SsrA RNA may compete with RFs for the occupation of the A-site, resulting in the SsrA tagging of full-length proteins depending on the amino acid sequence of the C-terminal end of the nascent peptide. In fact, the SsrA tagging at stop codons is increased in the *prf* mutants that express a partially defective RF1 or RF2, whereas it is reduced by the overexpression of wild-type RF1 or RF2 (our unpubl. results). Thus, three factors, RFs, tRNAs, and SsrA RNA, apparently compete for the occupation of the A-site. It should be noted, however, that the nature of the stop codon has little effect on the SsrA tagging mediated by the C-terminal RLESG sequence whereas it markedly influences the efficiency of translation readthrough. This is presumably because both RFs and tRNAs recognize the stop codons whereas SsrA RNA may recognize only the empty A-site.

It has been suggested that the nature of the last two amino acid residues of the nascent peptide influences translation readthrough at a stop codon, presumably by causing ribosome stalling (Bjornsson et al., 1996). Our studies suggest that certain C-terminal sequences could substantially influences both translation readthrough and SsrA tagging by causing ribosome stalling at stop codons. It is proposed that the C terminus of the nascent peptide is located around the peptidyltransferase center and the peptide passes through the tunnel of the 50S ribosomal subunit (Ban et al., 2000; Nissen et al., 2000). Therefore, certain nascent peptides may interact with ribosome itself or another component(s) of the translational machinery to prevent partially the action of RFs resulting in a transiently empty A-site. If this is the case, there would be more chance for both near-cognate tRNAs and SsrA RNA to enter the empty A-site. In fact, we observed that the C-terminal RLESG sequence that gives efficient SsrA tagging at stop codons could also markedly induce translation readthrough at the UGA codon. Inversely, the C-terminal dipeptides that induce an efficient readthrough could also allows a strong SsrA tagging. An alternative possibility regarding the mechanism of SsrA tagging at the stop codon would be that the ribosome stalling may induce endonucleolytic cleavages of an mRNA around the stop codon resulting in truncated mRNAs lacking a stop codon. In this regard, it is interesting to note that ribosome stalling appears to cause mRNA cleavages in several cases presumably by the action of either ribosome itself or a ribosome-associated factor (Loomis et al., 2001; Drider et al., 2002). We also observed mRNA fragments by northern blot analysis that presumably correspond to the cleavage products when SsrA tagging at the stop codons occurred in our system (our unpubl. results). Further studies are necessary to establish whether the ribosome stalling really induces endonucleolytic cleavages of an mRNA that lead to the SsrA tagging at a stop codon.

Several cases have been reported in which the nascent peptide can influence the translation termination or elongation, thereby affecting gene expression (Lovett & Rogers, 1996; Tenson & Ehrenberg, 2002). For example, it is known that the 24-residue product of the *tnaC* gene prevents the release of the peptide at the stop codon (Konan & Yanofsky, 1997, 1999). Another example is translation arrest in the chloramphenicol transacetylase gene (*cat*) in gram-positive bacteria by the nascent pentapeptide (MVKTD) that interacts with the ribosome in the presence of chloramphenicol (Lovett & Rogers, 1996). More recently, it has been shown that a specific nascent peptide in the *secM* gene induces the translation arrest, presumably by interacting with

the ribosomal exit tunnel (Nakatogawa & Ito, 2001, 2002). The present study provides an additional example of the control of ribosome function by the nascent peptide.

It is reported that highly expressed *E. coli* genes avoid an inefficient UGA stop codon and have a bias against C-terminal dipeptides, which give inefficient termination (Bjornsson et al., 1996). A search of the database revealed that no *E. coli* proteins possess either LESG or ESG sequences at their C terminus. In addition, only one *E. coli* protein, YabN, was found to retain the terminal DP dipeptide. Thus, C-terminal sequences that could induce an efficient SsrA tagging appear to have been avoided in *E. coli* genes during evolution, resulting in a bias in amino acid usage in the region immediately prior to the stop codon, at all three stop codons. One of the natural proteins, YbeL, has been shown to be tagged at a position corresponding to the strong UAA stop codon (Roche & Sauer, 2001). Interestingly, the C-terminal sequence of YbeL is EP that is known to also induce efficient translational readthrough at the UGA stop codon (Mottagui-Tabar et al., 1994). We also observed that the EP dipeptide confers the ability to be tagged to an unrelated protein when fused to its C terminus, although less efficiently compared to RLESG or DP sequences. At this moment, it is difficult to see any common feature between RLESG and DP. Although the SsrA tagging-evoking terminal sequences are expected to be avoided in general, they are apparently used in some cases such as YbeL. Studies on proteins and/or genes that retain the SsrA tagging-evoking terminal sequences would certainly be important to understand further the biological roles of SsrA RNA and *trans*-translation.

Just after the completion of this work, a paper was published by Hayes et al. (2002b) asserting that proline residues at the C terminus of nascent chains induce SsrA-tagging during translation termination. Our results are not only consistent with their finding but also reveal that the terminal sequences other than the C-terminal proline residue could induce the SsrA tagging at a stop codon. It is certainly interesting to study whether any other sequences could affect SsrA tagging as well as translational readthrough. Such study would be helpful to understand how particular C-terminal peptide sequences trigger SsrA tagging at stop codons.

MATERIALS AND METHODS

Media and growth conditions

Cells were grown aerobically at 37 °C in Luria-Bertani (LB) medium. Antibiotics were used at the following concentrations: ampicillin (50 μg/mL) and tetracycline (10 μg/mL). Bacterial growth was monitored by determining the optical density at 600 nm (OD_{600}).

Strains and plasmids

The *E. coli* K-12 strains used are MC4100Δ*ssrA* (Abo et al., 2000), TA411 (MC4100 $ssrA^{DD}$-FRT), TA481 (W3110Δ*crp* $ssrA^{DD}$-FRT), TA501 (W3110Δ*crp* Δ*ssrA*-FRT). The gene knock-out system of Datsenko and Wanner (2000) was used to construct the strains except MC4100Δ*ssrA*. The detailed procedures are described elsewhere (Abo et al., 2002).

Plasmid pIT613 contains *lacIZYA* in which the *lacI* is transcribed from the $lacI^q$ promoter on pBR322 (Abo et al., 2000). Plasmid pTA100 is a derivative of pIT613 in which *O3* is changed to the ideal operator($O3^{id}$; Abo et al., 2000). Plasmids pIT604, pIT605, pIT606, pIT607, pIT608, pST201, pST202, and pST203, derivatives of pIT613, were constructed by site-directed polymerase chain reaction (PCR) mutagenesis. Plasmids pHA7M (Abo et al., 2002), pHA7M0 (Abe & Aiba, 1996), and pHA7HA9 are derivatives of pHA7 carrying the *crp* gene under the control of the *bla* promoter (Aiba et al., 1982). To construct pHA7HA9, a *Mlu*I site was introduced near the last codon of *crp* of pHA7M0 by PCR mutagenesis. To construct pST010, derivatives of pHA7M, primers pBR H3-U (5′-GCAATTTAACTGTGATAAACTA-3′) and STA010-L (5′-AACGTACGGCACATCACCCACGCGTGCCGTAAAC-3′) were used to amplify the *crp* gene on pHA7M. The PCR-amplified fragment was digested with *Hin*dIII and *Bsi*WI(*Spl*I) and cloned into the *Hin*dIII-*Bsi*WI region of pHA7MHA9. Similarly, plasmids pSTA009, pSTA008, pSTA004, pSTA101, pSTA102, pSTA103, pSTA026, pSTA021, and pSTA020 were constructed by site-directed PCR mutagenesis by using appropriate primers. The plasmid pTA402 producing 6xHis-CRP-LESG was constructed as follows. Primers TA1 (5′-CAGGTTACGCGTCTGGAAAGCGGG-3′) and TA3 (5′-TCC AGCGTACGTTACGGCACCGCTTCT-3′) were used to amplify the *'lacIΔC1-lacPO-lacZ'* region of pIT604, by introducing *Mlu*I site within *lacI* ORF, and the TAA stop codon and *Bsi*WI site in *lacZ* ORF. The *Mlu*I- and *Bsi*WI-digested PCR fragment was cloned into the *Mlu*I-*Bsi*WI region of plasmid pHA7HA9 to obtain pTA202 producing a variant CRP (CRP-LESG) having the last four amino acid residues of LacIΔC1 fused at its C terminus. The region containing *crp-'lacIΔC1-lacPO-lacZ'* was amplified with primers *cro-crp-1* (Ueda et al., 2002) and pBR-RI (5′-GTATCACGAGGCCCTT-3′) to introduce a ATG-6x (CAT) sequence preceded by the SD sequence of lambda *cro* gene into *crp-'lacIΔC1*. The PCR fragment was digested with *Hin*dIII and *Eco*RI, and cloned into the *Hin*dIII-*Eco*RI region of the vector plasmid pBR322 to obtain pTA402. Plasmid pST511, pST521, pST523, pST531, pST533, pST603, pST604, and pST605 were constructed from pST513 (Abo et al., 2002) by PCR mutagenesis using appropriate primers.

Western blotting

Anti-CRP, anti-LacI, and anti-DD-tag polyclonal antibodies were described previously (Ishizuka et al., 1993; Abo et al., 2000). Bacterial cells grown in LB medium containing appropriate antibiotic were harvested at OD_{600} = 0.8 and suspended in 100 μL of SDS-PAGE loading buffer (62.5 mM Tris-HCl, pH 6.8, 2% SDS, 10% glycerol, 5% β-mercaptoethanol, 0.1% bromophenol blue). The sample was heated at 100 °C for 5 min. Total cellular extracts of the indicated amount were subjected to 12% acrylamide–0.1% SDS gel electro-

phoresis and transferred to Immobilon membrane (Millipore). The polypeptides detected by the antibodies were visualized by an ECL system (Pharmacia).

β-galactosidase assay

β-galactosidase activity was determined with permeabilized cells by the method of Miller (1992).

Purification of His-tagged protein

TA481 cells harboring pTA402 were grown in 20 mL of LB medium overnight, harvested, washed with 10 mM Tris-HCl, pH 8.0, 100 mM NaCl, 1 mM EDTA, and then suspended in buffer B (100 mM NaH_2PO_4, 10 mM Tris-HCl, 8 M urea, pH 8.0) containing 20 mM imidazole. The cells were lysed by sonication and centrifuged. The crude lysates were subjected to Ni^{2+}-NTA agarose (Qiagen) chromatography. The resin was washed three times with buffer C (100 mM NaH_2PO_4, 10 mM Tris-HCl, 8 M urea, pH 6.3) containing 20 mM imidazole. The His_6-fusion proteins were eluted from the resin with either 100 μL of 0.3% trifluoroacetic acid (TFA) or SDS-PAGE loading buffer.

Mass spectrometry

For mass spectrometry (MS) analysis of the untagged and DD-tagged His_6-fusion proteins, 1 μL of 0.3% TFA containing about 1 μg of purified proteins was mixed with 1 μL of 1% sinapinic acid in 30% acetnitrile and subjected to MALDI/TOF-MS. For fine mapping of the tagging site, His_6-fusion proteins were separated by 12% SDS-PAGE. The bands were cut out from the gel and a small piece of each band containing about 0.5 μg protein was treated with 0.1 μg of lysyl endopeptidase (Wako) in 20 μL of 100 mM Tris-HCl, pH 9.0, for 12 h at 37°C. The digested peptides were eluted with 300 μL of 50% acetnitrile, 5% formic acid, and concentrated to 20 μL. Then, the sample was desalted with a zip-tip reverse-phase column, mixed with 1% α-cyano-4-hydroxycinnamic acid in 70% acetnitrile, and subjected to MALDI/TOF-MS.

ACKNOWLEDGMENTS

We thank Hideaki Tagami (Nagoya University) for his help in mass spectrometry analysis. We are grateful to Kenneth Ritchie and Waleed El-Kazzaz for careful reading of this manuscript. This work was supported by Grants-in-Aid from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

Received July 19, 2002; returned for revision August 21, 2002; revised manuscript received September 3, 2002

REFERENCES

Abe H, Aiba H. 1996. Differential contributions of two elements of rho-independent terminator to transcription termination and mRNA stabilization. *Biochimie 78*:1035–1042.

Abo T, Inada T, Ogawa K, Aiba H. 2000. SsrA-mediated tagging and proteolysis of LacI and its role in the regulation of lac operon. *EMBO J 19*:3762–3769.

Abo T, Ueda K, Sunohara T, Ogawa K, Aiba H. 2002. SsrA-mediated protein tagging in the presence of miscoding drugs and its physiological role in *Escherichia coli*. *Genes Cells 7*:629–638.

Aiba H, Fujimoto S, Ozaki N. 1982. Molecular cloning and nucleotide sequencing of the gene for *E. coli* cAMP receptor protein. *Nucleic Acids Res 10*:1345–1361.

Ban N, Nissen P, Hansen J, Moore PB, Steitz TA. 2000. The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. *Science 289*:905–920.

Bjornsson A, Mottagui-Tabar S, Isaksson LA. 1996. Structure of the C-terminal end of the nascent peptide influences translation termination. *EMBO J 15*:1696–1704.

Collier J, Binet E, Bouloc P. 2002. Competition between SsrA tagging and translational termination at weak stop codons in *Escherichia coli*. *Mol Microbiol 45*:745–754.

Datsenko KA, Wanner BL. 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci USA 97*:6640–6645.

De La Cruz J, Vioque A. 2001. Increased sensitivity to protein synthesis inhibitors in cells lacking tmRNA. *RNA 7*: 1708–1716.

Drider D, DiChiara JM, Wei J, Sharp JS, Bechhofer DH. 2002 Endonuclease cleavage of messenger RNA in *Bacillus subtilis*. *Mol Microbiol 43*:1319–1329.

Gillet R, Felden B. 2001. Emerging views on tmRNA-mediated protein tagging and ribosome rescue. *Mol Microbiol 42*:879–885.

Gottesman S, Roche E, Zhou Y, Sauer, RT. 1998. The ClpXP and ClpAP proteases degrade proteins with carboxy-terminal peptide tails added by the SsrA-tagging system. *Genes Dev 12*:1338–1347.

Hayes CS, Bose B, Sauer RT. 2002a. Stop codons preceded by rare arginine codons are efficient determinants of SsrA tagging in *Escherichia coli*. *Proc Natl Acad Sci USA 99*:3440–3445.

Hayes CS, Bose B, Sauer RT. 2002b. Proline residues at the C-terminus of nascent chains induce SsrA-tagging during translation termination. *J Biol Chem 277*:33825–33832.

Herman C, Thevenet D, Bouloc P, Walker GC, D'Ari R. 1998. Degradation of carboxy-terminal-tagged cytoplasmic proteins by the *Escherichia coli* protease HflB (FtsH). *Genes Dev 12*:1348–1355.

Ishizuka H, Hanamura A, Kunimura T, Aiba H. 1993. A lowered concentration of cAMP receptor protein caused by glucose is an important determinant for catabolite repression in *Escherichia coli*. *Mol Microbiol 10*:341–350.

Karzai AW, Roche ED, Sauer RT. 2000. The SsrA-SmpB system for protein tagging, directed degradation and ribosome rescue. *Nat Struct Biol 7*:449–455.

Keiler KC, Waller PR, Sauer RT. 1996. Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. *Science 271*:990–993.

Konan KV, Yanofsky C. 1997. Regulation of the *Escherichia coli* tna operon: Nascent leader peptide control at the tnaC stop codon. *J Bacteriol 179*:1774–1779.

Konan KV, Yanofsky C. 1999. Role of ribosome release in regulation of tna operon expression in *Escherichia coli*. *J Bacteriol 181*:1530–1536.

Loomis WP, Koo JT, Cheung TP, Moseley SL. 2001. A tripeptide sequence within the nascent DaaP protein is required for mRNA processing of a fimbrial operon in *Escherichia coli*. *Mol Microbiol 39*:693–707.

Lovett PS, Rogers EJ. 1996. Ribosome regulation by the nascent peptide. *Microbiol Rev 60*:366–385.

Miller JH. 1992. *A short course in Bacterial Genetics: A laboratory manual for Escherichia coli and related Bacteria.* Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Mottagui-Tabar S, Bjornsson A, Isaksson LA. 1994. The second to last amino acid in the nascent peptide as a codon context determinant. *EMBO J 13*:249–257.

Mottagui-Tabar S, Isaksson LA. 1997. Only the last amino acids in the nascent peptide influence translation termination in *Escherichia coli* genes. *FEBS Lett 414*:165–170.

Muto A, Fujihara A, Ito KI, Matsuno J, Ushida C, Himeno H. 2000. Requirement of transfer-messenger RNA for the growth of *Bacillus subtilis* under stresses. *Genes Cells 5*:627–635.

Nakamura Y, Ito K, Isaksson LA. 1996. Emerging understanding of translation termination. *Cell 87*:147–150.

Nakatogawa H, Ito K. 2001. Secretion monitor, SecM, undergoes self-translation arrest in the cytosol. *Mol Cell 7*:185–192.

Nakatogawa H, Ito K. 2002. The ribosomal exit tunnel functions as a discriminating gate. *Cell 108*:629–636.

Nissen P, Hansen J, Ban N, Moore PB, Steitz TA. 2000. The structural basis of ribosome activity in peptide bond synthesis. *Science 289*:920–930.

Poole ES, Brown CM, Tate WP. 1995. The identity of the base following the stop codon determines the efficiency of in vivo translational termination in *Escherichia coli*. *EMBO J 14*:151–158.

Ranquet C, Geiselmann J, Toussaint A. 2001. The tRNA function of SsrA contributes to controlling repression of bacteriophage Mu prophage. *Proc Natl Acad Sci USA 98*:10220–10225.

Roche ED, Sauer RT. 1999. SsrA-mediated peptide tagging caused by rare codons and tRNA scarcity. *EMBO J 18*:4579–4589.

Roche ED, Sauer RT. 2001. Identification of endogenous SsrA-tagged proteins reveals tagging at positions corresponding to stop codons. *J Biol Chem 276*:28509–28515.

Tate WP, Mannering SA. 1996. Three, four or more: The translational stop signal at length. *Mol Microbiol 21*:213–219.

Tenson T, Ehrenberg M. 2002. Regulatory nascent peptides in the ribosomal tunnel. *Cell 108*:591–594.

Ueda K, Yamamoto Y, Ogawa K, Abo T, Inokuchi H, Aiba H. 2002. Bacterial SsrA system plays a role in coping with unwanted translational readthrough caused by suppressor tRNAs. *Genes Cells 7*:509–519.

Withey JH, Friedman DI. 2002. The biological roles of trans-translation. *Curr Opin Microbiol 5*:154–159.