Crystal structure of the full-length bacterial selenocysteine-specific elongation factor SelB

Yuzuru Itoh[1,2], Shun-ichi Sekine[1,2,3] and Shigeyuki Yokoyama[1,2,4,*]

[1]RIKEN Systems and Structural Biology Center, 1-7-22 Suehiro-cho, Tsurumi, Yokohama 230-0045, Japan,
[2]Department of Biophysics and Biochemistry, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan, [3]Division of Structural and Synthetic Biology, RIKEN Center for Life Science Technologies, 1-7-22 Suehiro-cho, Tsurumi, Yokohama 230-0045, Japan and [4]RIKEN Structural Biology Laboratory, 1-7-22 Suehiro-cho, Tsurumi, Yokohama 230-0045, Japan

ABSTRACT

Selenocysteine (Sec), the 21st amino acid in translation, uses its specific tRNA (tRNA^Sec) to recognize the UGA codon. The Sec-specific elongation factor SelB brings the selenocysteinyl-tRNA^Sec (Sec-tRNA^Sec) to the ribosome, dependent on both an in-frame UGA and a Sec-insertion sequence (SECIS) in the mRNA. The bacterial SelB binds mRNA through its C-terminal region, for which crystal structures have been reported. In this study, we determined the crystal structure of the full-length SelB from the bacterium *Aquifex aeolicus,* in complex with a GTP analog, at 3.2-Å resolution. SelB consists of three EF-Tu-like domains (D1–3), followed by four winged-helix domains (WHD1–4). The spacer region, connecting the N- and C-terminal halves, fixes the position of WHD1 relative to D3. The binding site for the Sec moiety of Sec-tRNA^Sec is located on the interface between D1 and D2, where a cysteine molecule from the crystallization solution is coordinated by Arg residues, which may mimic Sec binding. The Sec-binding site is smaller and more exposed than the corresponding site of EF-Tu. Complex models of Sec-tRNA^Sec, SECIS RNA, and the 70S ribosome suggest that the unique secondary structure of tRNA^Sec allows SelB to specifically recognize tRNA^Sec and characteristically place it at the ribosomal A-site.

INTRODUCTION

Selenocysteine is widely distributed in all three domains of life: Bacteria, Archaea and Eukarya, including humans. Selenocysteine is known as the 21st amino acid that is translationally incorporated into proteins, and its three-letter code is Sec (1). The chemical structure of selenocysteine is similar to that of cysteine, where the selenium atom in selenocysteine corresponds to the sulfur atom in cysteine. The selenol group in selenocysteine is more nucleophilic than the thiol group in cysteine, with pK_a values of ∼5.2 and ∼8.5, respectively (2). Therefore, the Sec residue is mainly found within the catalytic centers of enzymes performing oxidation-reduction reactions. Thus, selenium is an essential nutrient for human health.

The selenocysteine-specific tRNA (tRNA^Sec) is the largest tRNA, and its anticodon is complementary to UGA, a stop codon (3). Selenocysteine is synthesized as the tRNA^Sec-ligated form via a multi-step process, and not as the free amino acid (1). tRNA^Sec is ligated first with serine by seryl-tRNA synthetase (SerRS), to form seryl-tRNA^Sec (Ser-tRNA^Sec) (3). The selenocysteine synthase SelA converts Ser-tRNA^Sec directly to selenocysteinyl-tRNA^Sec (Sec-tRNA^Sec) in Bacteria (4). In Eukarya and Archaea, however, the Ser moiety of Ser-tRNA^Sec is phosphorylated by phosphoseryl-tRNA kinase (PSTK) (5), and the phosphate group of phosphoseryl-tRNA (Sep-tRNA^Sec) is then converted to the selenol group by Sep-tRNA^Sec Sec-tRNA^Sec synthase (SepSecS) (6,7). The translational incorporation of Sec into proteins utilizes the Sec-specific elongation factor (SelB or EF-Sec), rather than the general elongation factor (EF-Tu or EF1A) (8–10). SelB brings Sec-tRNA^Sec into the A site of the ribosome, depending not only on the UGA codon but also on the Sec insertion sequence (SECIS) in the mRNA, and Sec is then incorporated into the nascent, growing polypeptide (9,11).

The bacterial SECIS is located just downstream of the UGA codon within the coding region (12,13), while the eukaryal/archaeal SECIS is located in the untranslated region (14,15). The bacterial SelB directly interacts with the SECIS RNA (11,16,17), whereas the eukaryal/archaeal SelB requires an additional factor for SECIS recognition (10,18). The bacterial SelB uniquely possesses RNA-binding winged-helix domains (WHDs) for the SECIS recognition. The crystal structures of the full-length ar-

Table 1. Data collection and refinement statistics

Data collection	
Beam line	Spring-8 BL41XU
Wavelength (Å)	0.97895
Space group	$P4_32_12$
Cell parameters	$a = b = 114.2, c = 204.6$ Å
Resolution (Å)	50.0–3.19 (3.38–3.19)[a]
Unique reflections	42,987 (6,909)[a]
Completeness (%)	99.8 (99.4)[a]
Redundancy	15.48 (15.46)[a]
$CC_{1/2}$	99.9 (50.9)[a]
$I/\sigma(I)$	16.87 (1.26)[a]
Phasing	
No. of identified Se sites	8
Phasing power	0.718
FOM (before DM/after DM)	0.198/0.864
Structure refinement	
Working-set reflections	40,775
Test-set reflections	2179
Resolution (Å)	40.4–3.19
Number of protein atoms	4663
Number of ligands	1 (GMPPMP), 1 (L-cysteine)
Number of ions	1 (Mg^{2+}), 3 ($SO4^{2-}$)
Number of solvent molecules	0
R_{work} / R_{free}	0.207/0.261
Average B factor (Å^2)	
Overall	131.7
Protein	131.5
Ligand	125.4
Ion	181.7
RMSD bond lengths (Å)	0.010
RMSD bond angles (°)	1.39

[a]The statistics in the highest-resolution shell are given in parentheses.

chaeal SelB from the archaeon *Methanococcus maripaludi* and those of the C-terminal fragment of bacterial SelB from *Moorella thermoacetica* and *Escherichia coli* were reported (16,17,19–21). However, the full-length structure of the bacterial SelB was unknown.

In this study, we solved the crystal structure of the full-length SelB from the bacterium *Aquifex aeolicus*. This full-length structure clarified how the N- and C-terminal halves, which bind Sec-tRNA^{Sec} and the SECIS RNA, respectively, are connected. The structure of the aminoacyl-binding site revealed the mechanism of the discrimination of Sec from Ser by SelB. Furthermore, we discuss the UGA-decoding mechanism, based on the complex model of the ribosome, SelB, Sec-tRNA^{Sec} and mRNA.

MATERIALS AND METHODS

Protein preparation

The gene encoding *A. aeolicus* SelB was amplified by PCR from the genomic DNA, using forward and reverse primers containing NdeI and SalI restriction sites, respectively. The PCR product was digested with the restriction enzymes and ligated between the NdeI and SalI restriction sites of the pET25b vector (Novagen). The coding region in the plasmid is exactly the same as that of the native gene, and thus the expressed protein has no N- or C-terminal affinity or cloning tag.

The *E. coli* strain Rosetta 2(DE3) (Stratagene) was transformed with the expression plasmid, and the protein was overexpressed in LB medium at 37°C. The harvested *E. coli* cells were resuspended in 10 mM Tris–HCl buffer (pH 8.1), containing 0.25 M NaCl, 0.25 M $MgCl_2$ and 5 mM 2-mercaptoethanol (2-ME). The cells were disrupted with an ultrasonic homogenizer, and the supernatant of the lysate was incubated at 74°C for 20 min to denature and remove most of the *E. coli* proteins. The supernatant was diluted three-fold with buffer A [20 mM Tris–HCl buffer (pH 8.1) containing 10 mM 2-ME] and loaded on an SP Sepharose Fast Flow column (GE Healthcare) equilibrated with buffer A. The protein was eluted by a linear gradient of 0.21–1.2 M NaCl. The fractions containing SelB were pooled and dialyzed against buffer B [20 mM Tris–HCl buffer (pH 8.1), containing 250 mM NaCl and 10 mM 2-ME]. The sample was then loaded on a Mono S column (GE Healthcare), and eluted by a linear gradient of 0.24–0.96 M NaCl. The purified protein was dialyzed against buffer B, and concentrated to 10.0 mg/ml.

The selenomethionine (SeMet)-substituted SelB was overexpressed in the same strain as the native protein, grown in M9 minimal medium at 37°C. When the OD$_{600}$ reached 0.3, final concentrations of 60 mg/l L-selenomethionine, 100 mg/l each of L-threonine, L-lysine hydrochloride and L-phenylalanine, and 50 mg/l each of L-leucine, L-isoleucine and L-valine were added, to prevent methionine production. After 15 min, induction was performed by adding a final concentration of 0.5 mM isopropyl β-D-1-thiogalactopyranoside, and the cells were cultured overnight at 37°C.
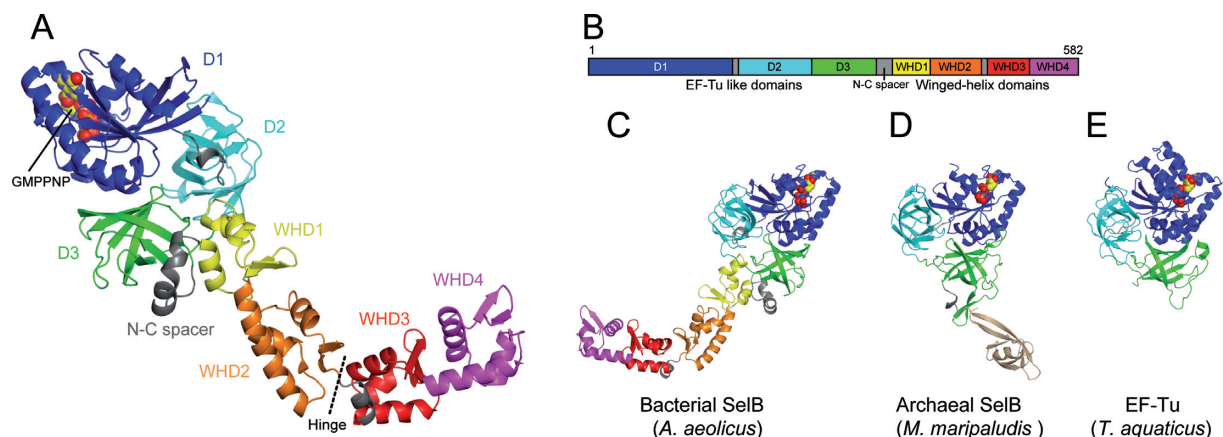
Figure 1. Overall structure of the full-length bacterial SelB. (**A**, **B**) Overall crystal structure of *A. aeolicus* SelB (A) and its domain composition (B). The GTP analog, GMPPNP, is shown as a sphere model. (**C–E**) Structure comparisons of the bacterial and archaeal SelBs and EF-Tu. The crystal structures of *A. aeolicus* SelB (C), *M. maripaludis* SelB (D) (PDB ID: 4ACB) (19), and *T. aquaticus* EF-Tu (PDB ID: 1EFT) (33) are shown in the same orientation. The C-terminal halves of the bacterial and archaeal SelBs have different orientations.

The SeMet-substituted protein was purified in a similar manner to that of the native protein. In detail, the harvested *E. coli* cells were resuspended in 20 mM Tris–HCl buffer (pH 8.1), containing 0.50 M NaCl, 0.50 M MgCl$_2$, and 10 mM 2-ME, and disrupted, and the lysate was incubated at 75°C for 10 min. The supernatant was diluted 10-fold with buffer A, and loaded on an SP Sepharose Fast Flow column. The protein was eluted by a linear gradient of 0.24–1.2 M NaCl. The fractions containing SelB were pooled, and diluted 2.3-fold with buffer A. The protein was further purified by Mono S column chromatography, with a linear gradient of 0.24–1.2 M NaCl. The purified protein was dialyzed against buffer B, and concentrated to 10.9 mg/ml.

Crystallization and X-ray diffraction data collection

The SelB protein, in complexes with 5′-guanylyl imidodiphosphate (GMPPNP) and with GDP, was prepared by adding final concentrations of 10 mM Mg$_2$SO$_4$ and 3.0 mM of each ligand. Crystallization of the SelB•GMPPNP and SelB•GDP complexes and the ligand-free SelB was performed at 20°C by the sitting-drop vapor diffusion method, by mixing the samples and the reservoir solutions. The Crystal Screen 2, Natrix (Hampton Research), Wizard I, and Wizard II (Emerald Bio-Science) screening kits were used for the initial screening of the crystallization conditions. SelB•GMPPNP was crystallized in Wizard I condition #8 [100 mM trisodium citrate–citric acid buffer (pH 5.5) containing 2.0 M (NH$_4$)$_2$SO$_4$], whereas SelB•GDP and the ligand-free SelB were not crystallized. Optimization was performed using the SeMet-substituted SelB•GMPPNP, and the final reservoir conditions were 100 mM sodium acetate–HCl buffer (pH 5.1) containing 1.7 M (NH$_4$)$_2$SO$_4$ and 20 mM L-cysteine.

The crystals thus obtained were transferred into the cryo-protective solution, 100 mM trisodium citrate–HCl buffer (pH 5.2) containing 42% D-glucose, 2.0 M (NH$_4$)$_2$SO$_4$, 20 mM L-cysteine, and 20 mM MgCl$_2$, and flash-cooled in a cryo-N$_2$ stream or liquid N$_2$. X-ray diffraction data were collected at 90 K, by using synchrotron radiation at the BL41XU of SPring-8 (Harima, Japan). The data were processed with the XDS program (22) (Table 1).

Structure determination and refinement

The SelB structure was solved by the single-wavelength anomalous dispersion method, using the SeMet-substituted SelB. The autoSHARP program (23) was used for experimental phasing, and the DM program (24,25) was used for the subsequent density modification. The Coot program (26) was used for manual fitting of the models to the electron density map. The structure was refined against the diffraction data by using the Phenix program (27), with iterative cycles of positional and temperature-factor refinements. The structure determination and refinement statistics are provided in Table 1, and the electron density map of a representative region is shown in Supplementary Figure S1.

Model building of the complex structures

The structural model of SelB in complex with Sec-tRNA^{Sec} was built, based on the crystal structure of the EF-Tu•Cys-tRNA^{Cys} complex (PDB ID: 1B23) (28). The EF-Tu in the complex was replaced with the present SelB crystal structure, by superimposing its EF-Tu-like domains on the corresponding region. The tRNA^{Cys} structure was then replaced by the *A. aeolicus* tRNA^{Sec} crystal structure reported previously (PDB ID: 3W3S) (29). The back-bone phosphorus atoms of the acceptor and T stems were used for the superimposition (nucleotides 1–7, 49–52 and 62–72 in tRNA^{Sec} versus nucleotides 1–7, 49–53 and 61–72 in tRNA^{Cys}). The CCA end of tRNA^{Sec} and the ligated Sec moiety were modeled, based on the L-cysteine molecule bound to the aminoacyl binding pocket of SelB, followed by slight moving of the acceptor stem and G73 to connect its CCA end.

The structural model of the SelB•Sec-tRNA^{Sec}•SECIS RNA complex was created by integrating the SECIS RNA into the SelB•Sec-tRNA^{Sec} complex model. The crystal structure of SECIS RNA in complex with *M. thermoacetica* SelB WHD3–4 (PDB ID: 1WSU) (17) was bound by superimposing the WHD3–4 structures.
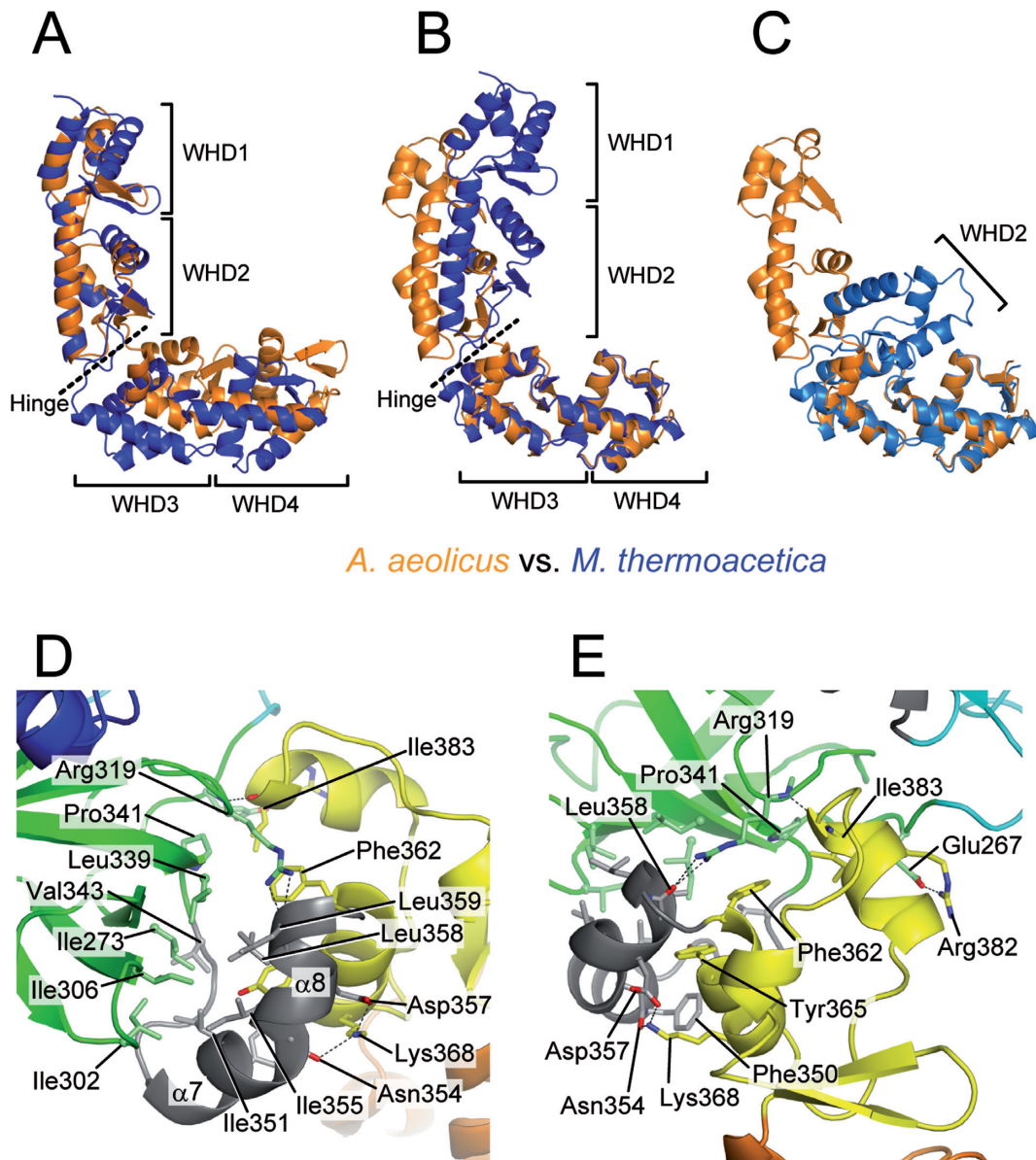
Figure 2. Structures of the C-terminal half and the spacer region of SelB. (**A–C**) Structural flexibility of the WHDs. The C-terminal half from the crystal structure of *A. aeolicus* SelB is superimposed on the crystal structure from the C-terminal half of *M. thermoacetica* SelB (PDB ID: 1LVA) (16), by fitting their WHD1–2 (A) and WHD3–4 (B), respectively. The C-terminal half of *M. thermoacetica* in complex with the SECIS RNA (PDB ID: 2UWM) (20) is also superimposed by fitting WHD3–4 (C), where WHD1 is disordered and the SECIS RNA is not shown. (**D**, **E**) Interactions of the N-C spacer between the N- and C-terminal halves. The spacer, colored gray, interacts with D3 and WHD1. Each region is colored as in Figure 1.

The modeled complex structure of the ribosome, SelB, Sec-tRNA^{Sec} and the SECIS RNA was constructed, based on the crystal structure of the 70S ribosome in complex with EF-Tu•Trp-tRNA^{Trp}, mRNA and P- and E-site tRNA^{Phe}s (PDB ID: 4V5L) (30). The EF-Tu•Trp-tRNA^{Trp} in the complex was replaced by the modeled complex of SelB, Sec-tRNA^{Sec} and the SECIS RNA, by superimposing D3 of SelB on the corresponding region. The orientation of WHDs 2–4 relative to WHD1 was changed by about 38°, to avoid clashing with the 16S rRNA (Supplementary Figure S2). To clarify the connection between the A-site codon and the SECIS RNA, the mRNA was replaced by the longer form from the crystal structure of the 70S ribosome in com-

plex with mRNA, and the A-, P- and E-site tRNA^{Phe}s (PDB ID: 4V6F) (31). The position of Sec-tRNA^{Sec} was adjusted to that of the Trp-tRNA^{Trp} by using the T stem for the superimposition (nucleotides 7, 49–52 and 62–66 in tRNA^{Sec} versus nucleotides 49–53 and 61–65 in tRNA^{Trp}). The coordinates of the distal half of the tRNA^{Sec} anticodon arm (nucleotides 30–41) were modified to generate the codon-anticodon base pairing.
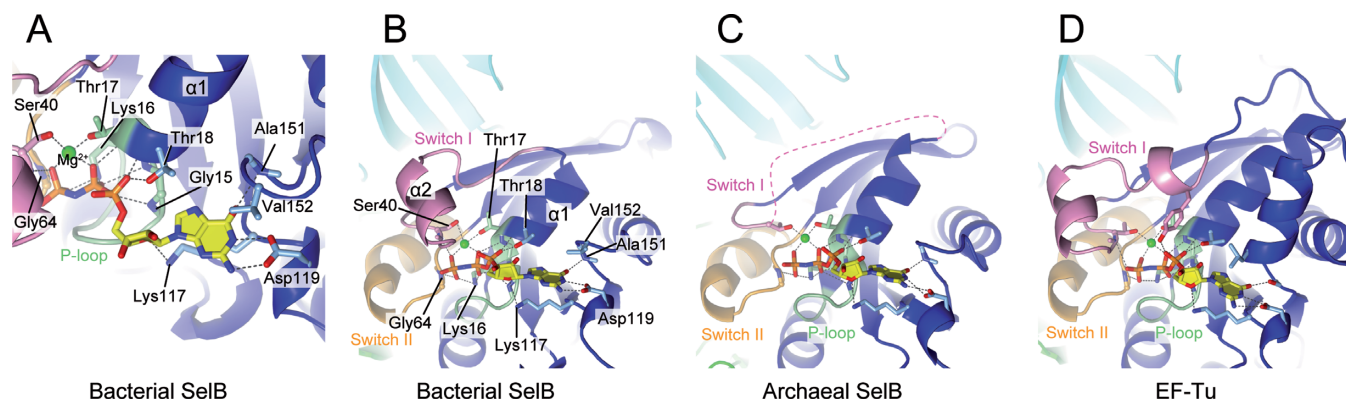
Figure 3. GMPNP interaction and the switch regions of SelB. (**A**) Interaction with GMPPNP in the crystal structure. GMPPMP and Mg^{2+} are shown as a stick and a green sphere, respectively. The P-loop motif (residues 10–17), switch I (residues 26–42), and switch II (residues 62–79) in D1 are colored pale green, pink and light orange, respectively. The rest of D1 is colored blue, and the other domains are colored as in Figure 1. (**B–D**) Comparison with archaeal SelB and bacterial EF-Tu. The GMPPNP-binding site and the switch regions of *A. aeolicus* SelB (B), *M. maripaludis* SelB (C) (PDB ID: 4ACB) (19), and *T. aquaticus* EF-Tu (D) (PDB ID: 1EFT) (33) are shown in the same orientation. Most of the switch I region in the archaeal SelB is disordered.

RESULTS AND DISCUSSION

Crystal structure of *A. aeolicus* SelB

We determined the crystal structure of the full-length bacterial SelB from *A. aeolicus* in complex with a GTP analog, 5′-guanylyl imidodiphosphate (GMPPNP), at 3.2-Å resolution (Figure 1A). The N-terminal half consists of the three EF-Tu-like domains, D1 (residues 1–171), D2 (residues 179–265) and D3 (residues 266–342), and the spacer connecting D1 and D2 (D1–2 linker, residues 172–178) (Figure 1A and B). The C-terminal half consists of the four WHDs, WHD1 (residues 362–406), WHD2 (residues 407–467), WHD3 (residues 475–524), and WHD4 (residues 525–582), and the linker connecting WHD2 and WHD3 (WHD2–3 linker, residues 468–474). The N- and C-terminal halves are connected by a spacer region (the N-C spacer, residues 343–361).

The structure of the N-terminal half consists of the GMPPNP-bound D1, the barrel-like D2 and the β-barrel D3, which are similar to those of the GMPPNP-bound archaeal SelB and the bacterial EF-Tu (Figure 1C–E) (19,32,33). The four WHDs are arranged in tandem in the unique C-terminal half (Figure 1A). The WHDs 2–4 share the same fold, α-β-α-α-β-β, where the β strands form an antiparallel β-sheet. In contrast, WHD1 has an irregular α-α-η-β-β fold, where one β-strand is imperfect and the third helix is a 3$_{10}$ helix. The structures of WHDs 2–4 are similar to those in the reported C-terminal-fragment structure of SelB from the bacterium *Moorella thermoacetica* (16,20) (Figure 2A and B). However, the WHD1 from *M. thermoacetica* has the standard α-β-α-α-β-β fold, and the difference may be due to the interactions with D3 and the N-C spacer.

The eukaryal/archaeal SelB contains another component in the C-terminal region in addition to the EF-Tu-like domains, and it is quite different from that in the bacterial SelB (Figure 1C and D). The C-terminal region of the eukaryal/archaeal SelB consists of only one domain. The orientation of the C-terminal half relative to the EF-Tu-like domains in the present bacterial SelB structure is different from that of the archaeal SelB C-terminal domain (19). This

difference may be related to the putatively distinct SECIS recognition mechanism.

The structural flexibility

The N-C spacer contains two α helices (α7 and α8), which are directly connected at an angle of 56°. The N-C spacer interacts with both D3 and WHD1 (Figure 2D and E). The Ile273, Ile302, Ile306, and Leu339 residues from D3 form a hydrophobic surface that interacts with Val343, Ile351, Ile355, Leu358 and Leu359 from the N-C spacer, while the Arg319 side chain from D3 hydrogen bonds with the main-chain carbonyl group of Leu358 from the N-C spacer. In addition, Asn354 and Asp357 from the N-C spacer form hydrogen bonds with Lys368 from WHD1, and Phe350 and Leu358 from the N-C spacer form hydrophobic interactions with Tyr365 from WHD1.

Direct interactions also exist between D3 and WHD1 (Figure 2D and E). The Glu267 (side chain) and Arg319 (main chain) residues from D3 form hydrogen bonds with Arg382 (side chain) and Ile383 (main chain) from WHD1, respectively. In addition, the Arg319 side chain stacks on Phe362 from WHD1 and Pro341 from D3 forms a hydrophobic interaction with Ile383 from WHD1. These interactions clearly indicate that the position of WHD1 relative to D3 is tightly fixed.

In contrast, a mobile hinge separates the C-terminal half into two regions, WHD1–2 and WHD3–4 (Figure 1A). WHD3–4 is stabilized by extensive inter-domain interactions, while WHD1–2 is much more flexible. However, a drastic movement between WHD1 and WHD2 is unlikely, since they are directly connected with no linker. This flexibility is consistent with the reported crystal structures (16,20) (Figure 2A–C).

The GTP-binding site and the switch regions

In the crystal structure, the GTP analog GMPPNP is bound to D1. The guanine ring is sandwiched between the side chains of Val152 and Lys117 (Figure 3A and B). The Asp119 side chain and the Ala151 main chain form hydrogen bonds with the guanine ring, to achieve the specificity
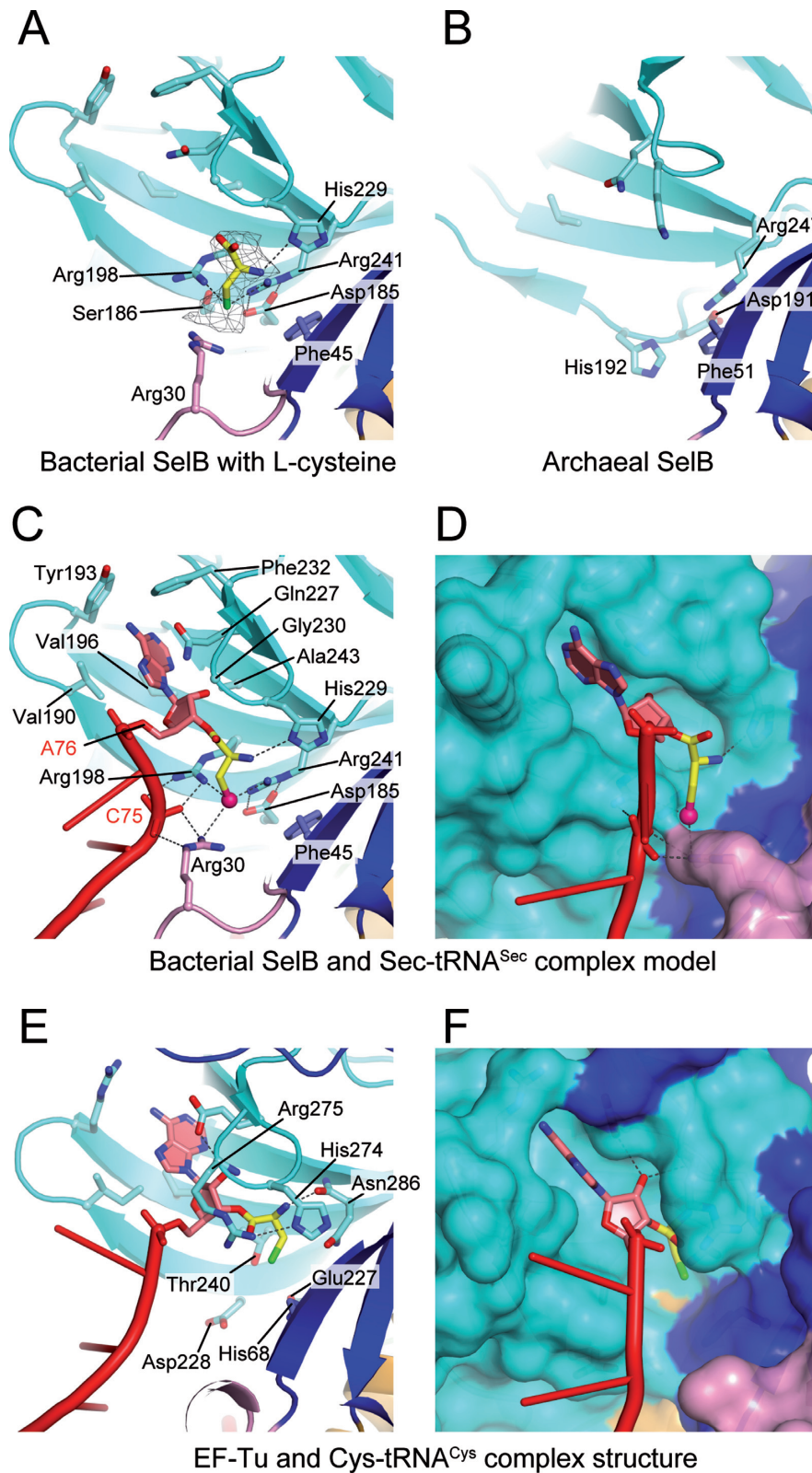
Figure 4. The binding site for the aminoacyl moiety on SelB. (**A**) The L-cysteine molecule (stick model) bound in the putative Sec-binding site. The gray mesh shows the omit $F_O − F_C$ electron density map for the L-cysteine at the 3.0 σ level. Each region of SelB is colored as in Figure 3. (**B**) The Sec binding site of the archaeal SelB (PDB ID: 4ACB), shown in the same orientation as in (A). Asp191 and Arg247 do not form a salt bridge, possibly due to the use of soaking techniques. (**C, D**) Modeled complex structure of bacterial SelB and Sec-tRNA^{Sec}. The Sec moiety, A76, and the phosphate group of C75 are shown as stick models. The selenium atom is shown as a magenta sphere. The solvent accessible surface is shown in (D). (**E, F**) The complex of EF-Tu and Cys-tRNA^{Cys} from the reported crystal structure (PDB ID: 1B23) (28) is shown for comparison. The binding site for the aminoacyl moiety is less exposed than that of SelB.
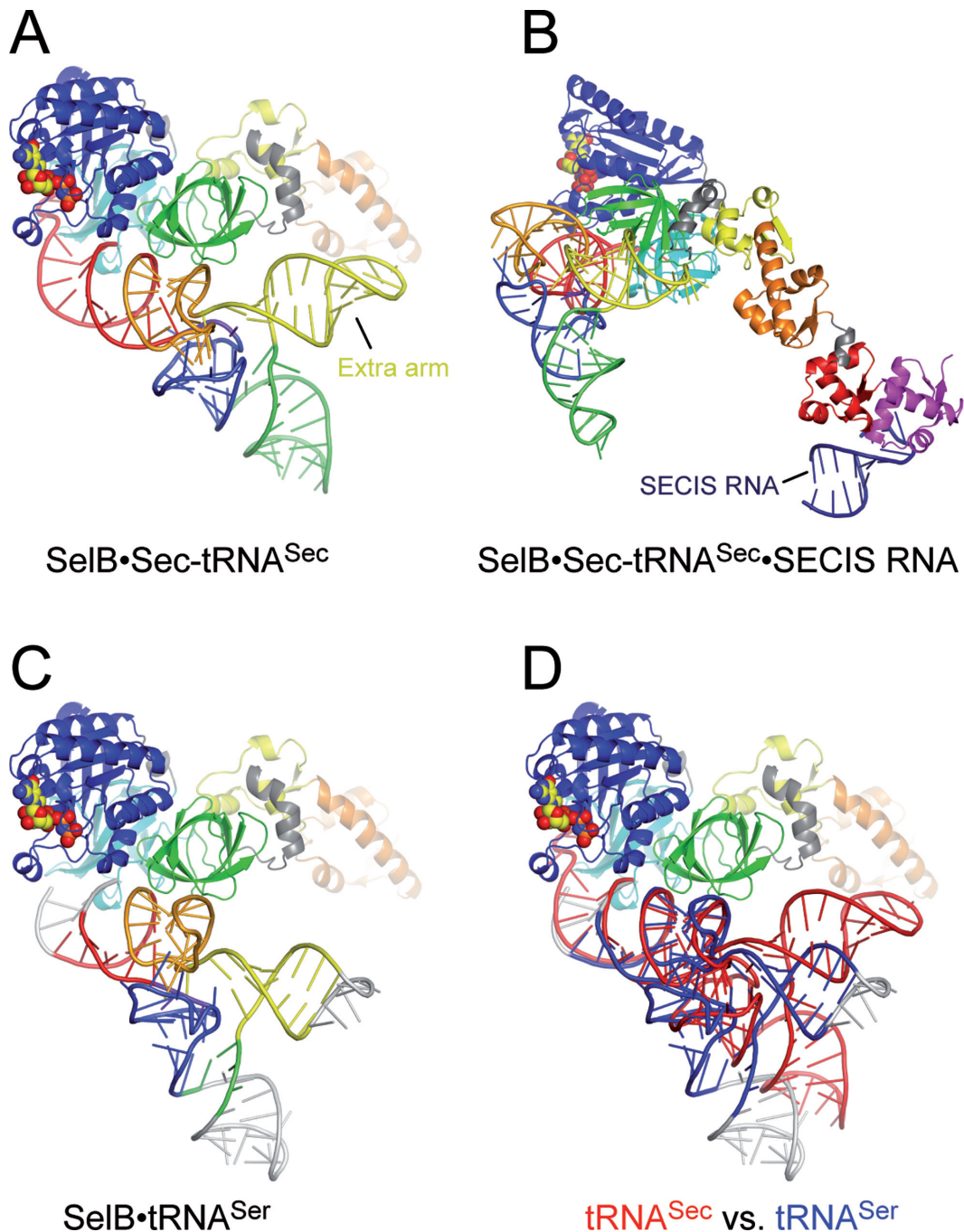
Figure 5. Modeled complex of SelB with Sec-tRNA^{Sec} and SECIS RNA. (**A**) Overall view of the modeled complex of SelB and Sec-tRNA^{Sec}. The acceptor, D, anticodon, extra and T arms of tRNA^{Sec} are colored red, blue, green, yellow, and orange, respectively. The linker between the acceptor and T stems is violet. Each region of SelB is colored as in Figure 1. (**B**) Modeled complex of SelB, Sec-tRNA^{Sec} and SECIS RNA. The hairpin of SECIS RNA is colored deep blue. (**C**) Modeled complex of SelB and tRNA^{Ser}, shown in the same orientation as in (A). The tRNA^{Ser} structure is from the crystal structure of the tRNA^{Ser}•SerRS complex (PDB ID: 1SER) (39). The disordered regions of tRNA^{Ser} in the crystal structure are colored light gray, and the other regions are colored as in tRNA^{Sec}. (**D**) Merged view of SelB with tRNA^{Sec} (red) and tRNA^{Ser} (blue), in the same orientation as in (A) and (C). Due to the difference in the rotation angle, the extra arm of tRNA^{Sec} contacts D3 of SelB, whereas that of tRNA^{Ser} is located away from SelB.
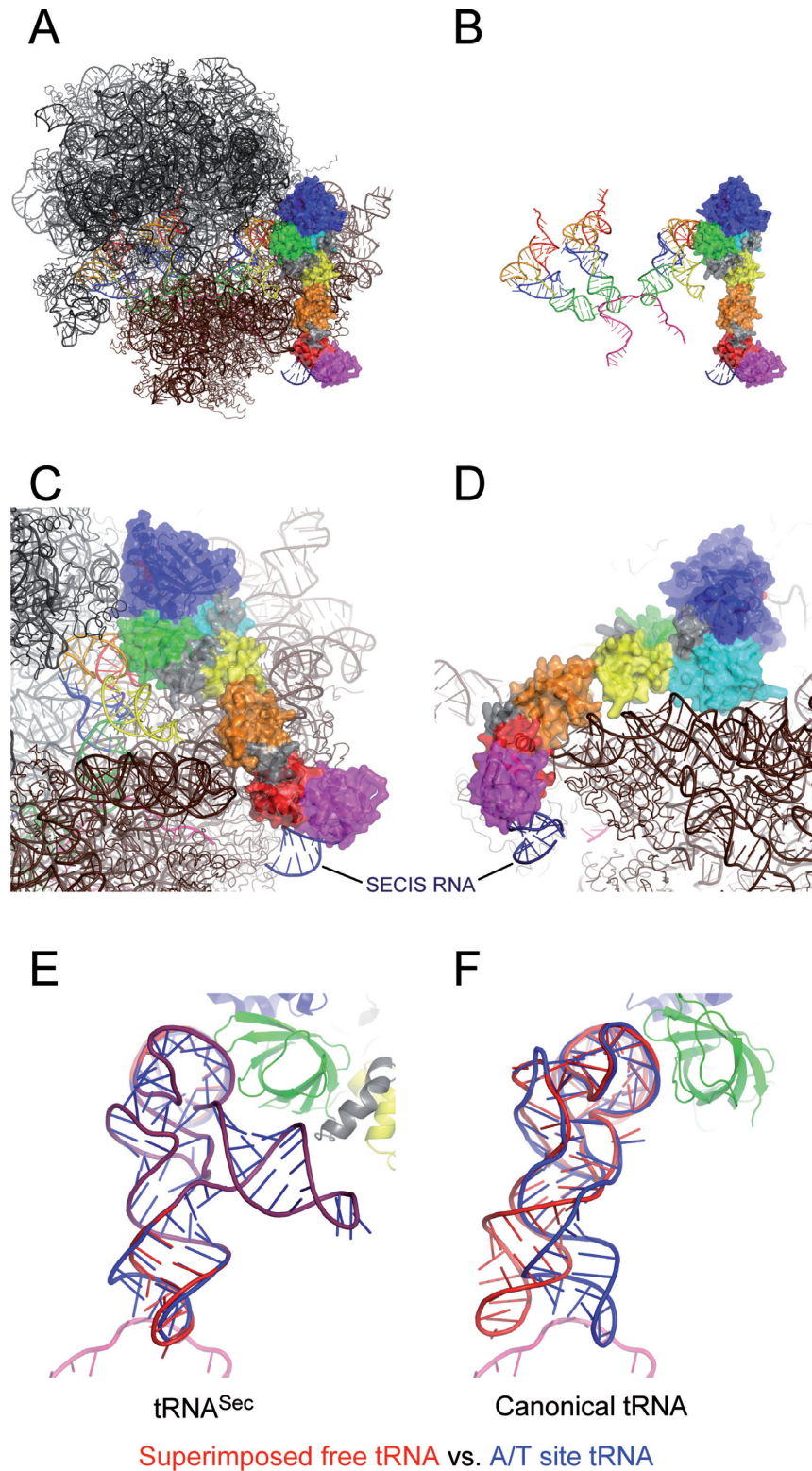
Figure 6. Modeled complex of SelB with the ribosome. (**A**, **B**) The modeled complex of SelB, Sec-tRNA^Sec, and SECIS RNA was further docked onto the 70S ribosome in complex with mRNA and the P- and E-site tRNAs. The 30S and 50S ribosomal particles and the mRNA are colored brown, black, and magenta, respectively. Surface model of SelB, colored as in Figure 1. The ribosome is not shown in (B). (**C**, **D**) Close-up views of SelB in the model. WHD1 and WHD2 contact the ribosome. (**E**) Structure comparison of tRNA^Sec in the A/T state (blue) and that in the ribosome-free form (red). The crystal structure of *A. aeolicus* tRNA^Sec, from the complex with archaeal SerRS (PDB ID: 3W3S) (29), was superimposed on tRNA^Sec in the model. In the crystal, only the extra and T arms of tRNA^Sec interact with SerRS, and the other regions can be regarded as the 'free' forms. (**F**) Structure comparison of tRNA^Trp in the A/T state (blue) in the crystal structure (PDB ID: 4V5L) (30) and the superimposed free tRNA^Phe (red) (PDB ID: 4TNA) (40). The orientation is the same as in (E).

for guanine nucleotides. The α- and β-phosphate groups interact with the continual main-chain NHs from the P-loop motif (residues 10–17) and the α1 helix. The Lys16 and Thr18 side chains also participate in the phosphate group binding.

The γ-phosphate group interacts with the Lys16 side chain, the Ser40 main chain and the Gly64 main chain (Figure 3A and B). The Ser40 and Gly64 residues are located on switches I and II, respectively. The Ser40 side chain also interacts with the γ-phosphate group, in a Mg^{2+}-mediated manner. The Mg^{2+} is coordinated by the β- and γ-phosphate groups and the Thr17 and Ser40 side chains. Since GDP lacks the γ-phosphate group, the interactions of the phosphate group with switches I and II are unique to the GTP-bound form.

Switch I consists of α2 and the linker connecting α1 and α2 (Figure 3B). In the reported *M. maripaludis* SelB crystal structure, switch I is mostly disordered (Figure 3C), possibly due to the method used to solve the GTP-bound form structure: back-soaking of the SelB•GDP co-crystal, followed by soaking with GMPPNP (19). Therefore, the 'GTP-bound form' of the archaeal SelB may be an intermediate between the GTP- and GDP-bound forms. The α1 helix and its following linker in the bacterial SelB are shorter than the corresponding regions in EF-Tu (32,33) (Figure 3B and D). In the EF-Tu structure, the corresponding linker region has a 3_{10} helix that interacts with the GTP analog (Figure 3D). On the other hand, the switch II structure is quite similar to that of EF-Tu (Figure 3A–C).

The Sec-binding site

As in EF-Tu, the aminoacyl-binding pocket for Sec-tRNA^{Sec} is located on the interface of D1 and D2 in the bacterial SelB. SelB in this study was crystallized under conditions containing L-cysteine, as an analog of selenocysteine. The aminoacyl-binding pocket contains an electron density corresponding to L-cysteine (Figure 4A), and it probably mimics the binding of the Sec moiety of Sec-tRNA^{Sec}. The thiol group of the bound L-cysteine interacts with Arg198 and Arg241 in D2, while the amino group interacts with His229. The Arg241 side chain is fixed, by forming a salt bridge with the Asp185 side chain and further stacking with the Phe45 side chain from D1.

The importance of Arg241 was confirmed by a mutational analysis, monitoring the *in vivo* 'recoding' activity in *E. coli* (Arg236 in *E. coli* numbering) (19). In contrast, Arg198 is unique to *A. aeolicus* SelB and is not conserved in other bacterial SelBs (Supplementary Figure S3). At the position of Ser186 in *A. aeolicus* SelB, the other SelBs have a basic residue (Arg181 in *E. coli*), and its importance has also been confirmed for *E. coli* (19). In the crystal structure, Arg198 and Ser186 are located next to each other on the β sheet (Figure 4A). In *A. aeolicus* SelB, Arg198 is likely to compensate for the loss of the basic residue at position 186.

The residues in the Sec-binding pocket are also conserved in the archaeal and eukaryal SelBs. In *M. maripaludis* SelB, Phe51, Asp191 and Arg247 correspond to Phe45, Asp185 and Arg241, respectively, of *A. aeolicus* SelB (19) (Figure 4A and B). Furthermore, His192 is located in the position corresponding to Ser186 of *A. aeolicus* SelB. However, the

Sec-binding site is not properly formed in the crystal structure of the archaeal SelB, where Asp191 and Arg247 do not form a salt bridge and Phe51 has no stacking interactions (Figure 4B). Phe51 is located next to switch I, which is disordered in the crystal structure of the archaeal SelB. The different orientations of the residue in the Sec-binding site probably occurred because of the GMPPNP soaking method (19). Therefore, the Sec interaction mechanism is likely to be conserved between Bacteria and Archaea.

Based on the bound L-cysteine in our crystal structure, we modeled the Sec-ligated CCA terminus of tRNA^{Sec} (Figure 4C and D). The binding pocket for the adenine ring of the tRNA terminal A76 is formed by Val190, Tyr193, Val196, Gln227, Gly230, Phe232 and Ala243. The adenine-ring binding site is located at the corresponding position in the EF-Tu•Cys-tRNA^{Cys} complex structure (28) (Figure 4E and F). In addition to Arg198 and Arg241 in the bacterial SelB, Arg30 located in switch I may also participate in the Sec interaction. The side chain of Arg30 is close to the thiol group of the bound L-cysteine in the crystal structure (Figure 4A). Since the selenol group of Sec is deprotonated at neutral pH, three Arg residues can simultaneously interact with the selenol group, which may contribute to the specificity for Sec-tRNA^{Sec}, rather than Ser-tRNA^{Sec}. Arg30 is well conserved as Arg/Lys in both Bacteria and Archaea/Eukarya, although some bacterial SelBs lack it (Supplementary Figure S3). In our model, Arg30 may also interact with the phosphate group of C75 in tRNA^{Sec} (Figure 4C).

The binding affinity between Sec-tRNA^{Sec} and SelB•GTP is reportedly much stronger than that between deacylated tRNA^{Sec} and SelB•GTP, with pK_d values of 0.2 pM and 190 nM, respectively (34). It is estimated that three ionic interactions are formed upon the binding of the Sec moiety of Sec-tRNA^{Sec} (34). This is consistent with the ionic interactions between the deprotonated selenol group and the Arg residues in our modeled structure.

In the EF-Tu•aminoacyl-tRNA complex structures, His68, Glu227, Asp228, Thr240, His274, Arg275 and Asn286 form a space to accommodate the side chain of the aminoacyl moiety, without specificity for any of the twenty amino acids (28,35) (Figure 4E and F). Interestingly, the corresponding space in SelB is occupied by the salt bridge Asp185•Arg241 (Figure 4A), and the Sec binding site is more exposed to the surface than the buried binding site in EF-Tu (Figure 4D and F). Using this shallow pocket with the Arg residues, SelB achieves the Sec-specific interaction, which is essential for SelB.

Structural model of the complex of SelB with tRNA^{Sec} and mRNA SECIS

In addition to the CCA terminus, the rest of the tRNA^{Sec} molecule was also modeled, based on the comparison to the EF-Tu•Cys-tRNA^{Cys} complex structure (PDB ID: 1B23) (28) (Figure 5A) (See methods). The SECIS RNA structure, generated from the SelB WHD3–4•SECIS RNA complex structure (PDB ID: 1WSU) (17), was also bound by superimposing the WHD3–4 structures (Figure 5B).

This model suggested the preferred positions of tRNA^{Sec} and the SECIS RNA on SelB. The tRNA^{Sec} and the SE-

CIS RNA are located on the N- and C-terminal halves of SelB, respectively. D1 and D2 interact with the acceptor arm of tRNA[Sec], while D3 interacts with the T and extra arms. WHD3 and WHD4 interact with the SECIS RNA. Although it was speculated that WHD2 interacted with the anticodon arm of tRNA[Sec] (20), this interaction does not seem to be possible, even if the flexibility of WHD1 and 2 is taken into account.

The interaction between D3 of SelB and the long extra arm of tRNA[Sec] could contribute to the specificity for tRNA[Sec], rather than the canonical tRNA species, such as tRNA[Ser] and tRNA[Cys]. The long extra arm, composed of a stem-and-loop, is unique to tRNA[Sec], tRNA[Ser], tRNA[Leu] and tRNA[Tyr]. Furthermore, the stem regions of the acceptor and T arms of tRNA[Sec] are quite different from those of the canonical tRNAs, including tRNA[Ser], tRNA[Leu] and tRNA[Tyr]. In the canonical tRNAs, the 7-base-pair acceptor stem and the 5-base-pair T stem form the 12-base-pair acceptor-T helix (the 7–5 secondary structure). In contrast, tRNA[Sec] has a 13-base-pair acceptor-T helix (the 8–5 and 9–4 secondary structures in Bacteria and Archaea/Eukarya, respectively) (29,36,37). The longer acceptor-T helix changes the orientation of the extra arm, because of the rotation of 33° per base pair along the helix axis. Therefore, even though tRNA[Ser] has a long extra arm, it is located away from D3 when tRNA[Ser] (39) is modeled on SelB (Figure 5C and D).

Recoding on the ribosome

To understand the recoding mechanism, the modeled complex of SelB, Sec-tRNA[Sec], and the SECIS RNA was further docked onto the 70S ribosome, by comparison with EF-Tu•Trp-tRNA[Trp] in complex with the 70S ribosome (PDB ID: 4V5L) (30) (Figure 6A–D). In the process of the modeling, the positions of WHDs 2–4 relative to the other domains were changed, to avoid clashing with the 16S rRNA (Supplementary Figure S2). In the resulting model, WHD1 and WHD2 contact helix 16 of the 16S rRNA (Figure 6C and D).

The SECIS RNA, bound to WHD3 and WHD4, is located in the vicinity of the mRNA-entrance channel of the ribosome (Figure 6B–D). In general, the SECIS RNA hairpin forms a 6–8 base-pair stem, and its loop has the conserved G at ~20 nucleotides downstream from the UGA codon (11). In our model, the distance between the A-site codon and the hairpin loop is ~57 Å, which can be connected by ~13 nucleotides. The hinge between WHD2 and WHD3 can accommodate the differences in the number of nucleotides derived from the diversity among the Sec-encoding mRNAs.

In first step of decoding, the canonical aminoacyl-tRNA is in the so-called A/T state, where EF-Tu binds to the acceptor arm of the tRNA and the anticodon forms base pairs with the codon at the A site. In this state, the anticodon and D arms form a bent helix (Figure 6F) (30,40), which causes repulsive force. This force can only be withstood by a matched codon-anticodon pair, and thus it accomplishes the high specificity (30,38). In contrast, Sec-tRNA[Sec] in the A/T state requires only a small conformation change in its anticodon loop, without helix bending, in the model (Figure 6E). This is due to the difference in the orientation (33°) of the anticodon arm, caused by the longer acceptor-T helix.

Thus, in this study, we identified the specific Sec-binding site formed by D1 and D2 of SelB. Moreover, the full-length SelB structure revealed that the N-C spacer restricts the arrangement of the C-terminal half, containing WHD1–4, relative to the N-terminal half, with D1–3, and facilitated model building to examine the recoding mechanism involving the C-terminal half and the bound SECIS RNA.

ACCESSION NUMBER

The atomic coordinates and structure factors have been deposited in the Protein Data Bank (ID: 4ZU9).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

FUNDING

REFERENCES

1. Böck,A., Forchhammer,K., Heider,J., Leinfelder,W., Sawers,G., Veprek,B. and Zinoni,F. (1991) Selenocysteine: the 21st amino acid. *Mol. Microbiol.*, **5**, 515–520.
2. Huber,R.E. and Criddle,R.S. (1967) Comparison of the chemical properties of selenocysteine and selenocystine with their sulfur analogs. *Arch. Biochem. Biophys.*, **122**, 164–173.
3. Leinfelder,W., Zehelein,E., Mandrand-Berthelot,M.A. and Böck,A. (1988) Gene for a novel tRNA species that accepts L-serine and cotranslationally inserts selenocysteine. *Nature*, **331**, 723–725.
4. Forchhammer,K. and Böck,A. (1991) Selenocysteine synthase from *Escherichia coli*. Analysis of the reaction sequence. *J. Biol. Chem.*, **266**, 6324–6328.
5. Carlson,B.A., Xu,X.M., Kryukov,G.V., Rao,M., Berry,M.J., Gladyshev,V.N. and Hatfield,D.L. (2004) Identification and characterization of phosphoseryl-tRNA[Ser]Sec kinase. *Proc. Natl. Acad. Sci. U.S.A.*, **101**, 12848–12853.
6. Yuan,J., Palioura,S., Salazar,J.C., Su,D., O'Donoghue,P., Hohn,M.J., Cardoso,A.M., Whitman,W.B. and Söll,D. (2006) RNA-dependent conversion of phosphoserine forms selenocysteine in eukaryotes and archaea. *Proc. Natl Acad. Sci. USA*, **103**, 18923–18927.

7. Xu,X.M., Carlson,B.A., Mix,H., Zhang,Y., Saira,K., Glass,R.S., Berry,M.J., Gladyshev,V.N. and Hatfield,D.L. (2007) Biosynthesis of selenocysteine on its tRNA in eukaryotes. *PLoS Biol.*, **5**, e4.

8. Forchhammer,K., Leinfelder,W. and Böck,A. (1989) Identification of a novel translation factor necessary for the incorporation of selenocysteine into protein. *Nature*, **342**, 453–456.

9. Fagegaltier,D., Hubert,N., Yamada,K., Mizutani,T., Carbon,P. and Krol,A. (2000) Characterization of mSelB, a novel mammalian elongation factor for selenoprotein translation. *EMBO J.*, **19**, 4796–4805.

10. Rother,M., Wilting,R., Commans,S. and Böck,A. (2000) Identification and characterisation of the selenocysteine-specific translation factor SelB from the archaeon *Methanococcus jannaschii*. *J. Mol. Biol.*, **299**, 351–358.

11. Baron,C., Heider,J. and Böck,A. (1993) Interaction of translation factor SELB with the formate dehydrogenase H selenopolypeptide mRNA. *Proc. Natl. Acad. Sci. U.S.A.*, **90**, 4181–4185.

12. Zinoni,F., Heider,J. and Böck,A. (1990) Features of the formate dehydrogenase mRNA necessary for decoding of the UGA codon as selenocysteine. *Proc. Natl. Acad. Sci. U.S.A.*, **87**, 4660–4664.

13. Berg,B.L., Baron,C. and Stewart,V. (1991) Nitrate-inducible formate dehydrogenase in *Escherichia coli* K-12. II. Evidence that a mRNA stem-loop structure is essential for decoding opal (UGA) as selenocysteine. *J. Biol. Chem.*, **266**, 22386–22391.

14. Berry,M.J., Banu,L., Chen,Y.Y., Mandel,S.J., Kieffer,J.D., Harney,J.W. and Larsen,P.R. (1991) Recognition of UGA as a selenocysteine codon in type I deiodinase requires sequences in the 3′ untranslated region. *Nature*, **353**, 273–276.

15. Wilting,R., Schorling,S., Persson,B.C. and Böck,A. (1997) Selenoprotein synthesis in archaea: identification of an mRNA element of *Methanococcus jannaschii* probably directing selenocysteine insertion. *J. Mol. Biol.*, **266**, 637–641.

16. Selmer,M. and Su,X.D. (2002) Crystal structure of an mRNA-binding fragment of *Moorella thermoacetica* elongation factor SelB. *EMBO J.*, **21**, 4145–4153.

17. Yoshizawa,S., Rasubala,L., Ose,T., Kohda,D., Fourmy,D. and Maenaka,K. (2005) Structural basis for mRNA recognition by elongation factor SelB. *Nat. Struct. Mol. Biol.*, **12**, 198–203.

18. Lesoon,A., Mehta,A., Singh,R., Chisolm,G.M. and Driscoll,D.M. (1997) An RNA-binding protein recognizes a mammalian selenocysteine insertion sequence element required for cotranslational incorporation of selenocysteine. *Mol. Cell. Biol.*, **17**, 1977–1985.

19. Leibundgut,M., Frick,C., Thanbichler,M., Böck,A. and Ban,N. (2005) Selenocysteine tRNA-specific elongation factor SelB is a structural chimaera of elongation and initiation factors. *EMBO J.*, **24**, 11–22.

20. Ose,T., Soler,N., Rasubala,L., Kuroki,K., Kohda,D., Fourmy,D., Yoshizawa,S. and Maenaka,K. (2007) Structural basis for dynamic interdomain movement and RNA recognition of the selenocysteine-specific elongation factor SelB. *Structure*, **15**, 577–586.

21. Soler,N., Fourmy,D. and Yoshizawa,S. (2007) Structural insight into a molecular switch in tandem winged-helix motifs from elongation factor SelB. *J. Mol. Biol.*, **370**, 728–741.

22. Kabsch,W. (2010) Xds. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 125–132.

23. Vonrhein,C., Blanc,E., Roversi,P. and Bricogne,G. (2007) Automated structure solution with autoSHARP. *Methods Mol. Biol.*, **364**, 215–230.

24. Winn,M.D., Ballard,C.C., Cowtan,K.D., Dodson,E.J., Emsley,P., Evans,P.R., Keegan,R.M., Krissinel,E.B., Leslie,A.G., McCoy,A. *et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.*, **67**, 235–242.

25. Cowtan,K. (1999) Error estimation and bias correction in phase-improvement calculations. *Acta Crystallogr. D Biol. Crystallogr.*, **55**, 1555–1567.

26. Emsley,P., Lohkamp,B., Scott,W.G. and Cowtan,K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 486–501.

27. Afonine,P.V., Grosse-Kunstleve,R.W., Echols,N., Headd,J.J., Moriarty,N.W., Mustyakimov,M., Terwilliger,T.C., Urzhumtsev,A., Zwart,P.H. and Adams,P.D. (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.*, **68**, 352–367.

28. Nissen,P., Thirup,S., Kjeldgaard,M. and Nyborg,J. (1999) The crystal structure of Cys-tRNA[Cys]-EF-Tu-GDPNP reveals general and specific features in the ternary complex and in tRNA. *Structure*, **7**, 143–156.

29. Itoh,Y., Sekine,S., Suetsugu,S. and Yokoyama,S. (2013) Tertiary structure of bacterial selenocysteine tRNA. *Nucleic Acids Res.*, **41**, 6729–6738.

30. Voorhees,R.M., Schmeing,T.M., Kelley,A.C. and Ramakrishnan,V. (2010) The mechanism for activation of GTP hydrolysis on the ribosome. *Science*, **330**, 835–838.

31. Jenner,L.B., Demeshkina,N., Yusupova,G. and Yusupov,M. (2010) Structural aspects of messenger RNA reading frame maintenance by the ribosome. *Nat. Struct. Mol. Biol.*, **17**, 555–560.

32. Berchtold,H., Reshetnikova,L., Reiser,C.O., Schirmer,N.K., Sprinzl,M. and Hilgenfeld,R. (1993) Crystal structure of active elongation factor Tu reveals major domain rearrangements. *Nature*, **365**, 126–132.

33. Kjeldgaard,M., Nissen,P., Thirup,S. and Nyborg,J. (1993) The crystal structure of elongation factor EF-Tu from *Thermus aquaticus* in the GTP conformation. *Structure*, **1**, 35–50.

34. Paleskava,A., Konevega,A.L. and Rodnina,M.V. (2010) Thermodynamic and kinetic framework of selenocysteyl-tRNA[Sec] recognition by elongation factor SelB. *J. Biol. Chem.*, **285**, 3014–3020.

35. Nissen,P., Kjeldgaard,M., Thirup,S., Polekhina,G., Reshetnikova,L., Clark,B.F. and Nyborg,J. (1995) Crystal structure of the ternary complex of Phe-tRNA[Phe], EF-Tu, and a GTP analog. *Science*, **270**, 1464–1472.

36. Itoh,Y., Chiba,S., Sekine,S. and Yokoyama,S. (2009) Crystal structure of human selenocysteine tRNA. *Nucleic Acids Res.*, **37**, 6259–6268.

37. Chiba,S., Itoh,Y., Sekine,S. and Yokoyama,S. (2010) Structural basis for the major role of O-phosphoseryl-tRNA kinase in the UGA-specific encoding of selenocysteine. *Mol. Cell*, **39**, 410–420.

38. Schmeing,T.M., Voorhees,R.M., Kelley,A.C., Gao,Y.G., Murphy,F.V., Weir,J.R. and Ramakrishnan,V. (2009) The crystal structure of the ribosome bound to EF-Tu and aminoacyl-tRNA. *Science*, **326**, 688–694.

39. Biou,V., Yaremchuk,A., Tukalo,M. and Cusack,S. (1994) The 2.9 Å crystal structure of *T. thermophilus* seryl-tRNA synthetase complexed with tRNA[Ser]. *Science*, **263**, 1404–1410.

40. Hingerty,B., Brown,R.S. and Jack,A. (1978) Further refinement of the structure of yeast tRNA[Phe]. *J. Mol. Biol.*, **124**, 523–534.