RNA

The three-dimensional structure of the Moorella thermoacetica selenocysteine insertion sequence RNA hairpin and its interaction with the elongation factor SelB

Alexander V. Beribisky, Tony J. Tavares, Andrew N. Amborski, Mina Motamed, Anne E. Johnson, Tobi L. Mark and Philip E. Johnson

References	This article cites 33 articles, 8 of which can be accessed free at: http://www.rnajournal.org/cgi/content/full/13/11/1948#References
Email alerting service	Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here**

Notes

To subscribe to *RNA* go to:
http://www.rnajournal.org/subscriptions/

The three-dimensional structure of the *Moorella thermoacetica* selenocysteine insertion sequence RNA hairpin and its interaction with the elongation factor SelB

ALEXANDER V. BERIBISKY,[1,3] TONY J. TAVARES,[1,3] ANDREW N. AMBORSKI,[1] MINA MOTAMED,[1] ANNE E. JOHNSON,[2] TOBI L. MARK,[1] and PHILIP E. JOHNSON[1]

[1]Department of Chemistry, York University, Toronto, Ontario M3J 1P3, Canada
[2]Department of Chemistry and Biology, Ryerson University, Toronto, Ontario M5B 2K3, Canada

ABSTRACT

Incorporation of the amino acid selenocysteine into a growing protein chain involves the interaction between a hairpin in the mRNA termed the selenocysteine insertion sequence (SECIS) and the special elongation factor SelB. Here we present the structure of the SECIS from the thermophilic organism *Moorella thermoacetica* (SECIS-MT) determined using nuclear magnetic resonance (NMR) spectroscopy. The SECIS-MT hairpin structure contains a pentaloop with the first and fourth nucleotides of the loop forming a noncanonical GC base pair; the fifth loop nucleotide is bulged out and unstructured. The G and U in positions two and three are on opposite sides of the loop and solvent exposed. The backbone resonances of the SECIS-binding domain from the *M. thermoacetica* SelB protein were assigned, and the degree of chemical shift perturbations that occur upon SECIS binding were mapped onto the structure of the complex. We demonstrate that a region in the third winged-helix domain of SelB, not previously implicated in binding, is affected by SECIS binding.

Keywords: NMR; RNA structure; selenocysteine; SelB; protein–RNA interaction

INTRODUCTION

The genetic code is made up of codons, each encoding one of the 20 common amino acids or establishing the start or stop points for the translation of mRNA. In certain circumstances the function of the UGA stop codon is reassigned so that it codes for the rare amino acid selenocysteine (Atkins and Gesteland 2000). Selenocysteine is found in both eukaryotes and prokaryotes and is the main biological form of the essential trace element selenium. The selenol group has the ability to act as a powerful nucleophile, and it is often found in the active site of a variety of redox enzymes such as glycine reductase, formate dehydrogenase, glutathione peroxidase, and thioredoxin reductase (Stadtman 2002).

Selenocysteine, together with pyrrolysine (Atkins and Gesteland 2002), are unique in that they are the only

known rare amino acids that are genetically encoded and do not result from post-translational modification. There are two known mechanisms for selenocysteine incorporation into proteins (Krol 2002). The most well studied is that found in bacteria, where *Escherichia coli* serves as the model system for understanding the mechanism of selenocysteine incorporation in bacterial species. In this mechanism, the UGA selenocysteine-directing codon is immediately upstream of an mRNA hairpin structure known as the selenocysteine insertion sequence (SECIS; Fig. 1). The SECIS element is then bound by the special ribosomal elongation factor B (SelB). SelB is a multidomain protein with the N-terminal domain being similar to the conventional elongation factor EF-Tu and the C-terminal domain possessing a high binding specificity for the SECIS mRNA hairpin (Kromayer et al. 1996; Leibundgut et al. 2005). This protein–RNA pair then forms part of a complex that contains GTP and the activated tRNA for selenocysteine (tRNA^{Sec}). Once assembled, the ribosome can direct the incorporation of selenocysteine into the growing polypeptide chain.

The crystal structure of the SECIS-binding C-terminal domain of *Moorella thermoacetica* SelB has been determined

---
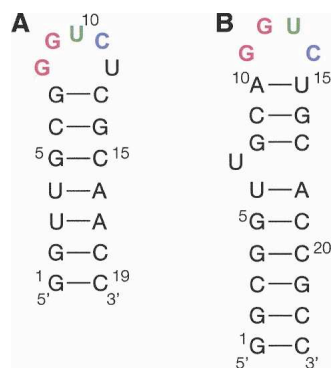
FIGURE 1. Secondary structure diagram of (*A*) the SECIS RNA from the *M. thermoacetica fdhA* gene (SECIS-MT) (Selmer and Su 2002) and (*B*) the SECIS from the *E. coli fdhF* gene. Loop nucleotides are colored the same in both structures.

both free and bound to a *M. thermoacetica* SECIS (SECIS-MT). It has been shown to contain four winged helix (WH) domains (Selmer and Su 2002; Yoshizawa et al. 2005; Ose et al. 2007). The winged helix fold is a well known class of DNA-binding proteins and is a subset of the helix-turn-helix family (Gajiwala and Burley 2000). Deletion mutation studies of *E. coli* SelB localized the SECIS binding function to the C-terminal half of the SelB C-terminal domain (Kromayer et al. 1996). This corresponds to the C-terminal pair of winged helix domains in the *M. thermoacetica* SelB structure, though only the last WH domain was shown to interact with the SECIS (Yoshizawa et al. 2005).

In *E. coli* a number of SECIS molecules have been identified, and their function has been the subject of extensive biochemical analysis (Huttenhofer et al. 1996; Klug et al. 1999; Li et al. 2000; Thanbichler et al. 2000). The best-studied SECIS RNA molecules are from the genes encoding the proteins formate dehydrogenase H and N, *fdhF* and *fdnG*, respectively. The minimal RNA region required for SelB binding to *fdhF* and *fdnG* has been defined to be the apical portion of a larger hairpin (Kromayer et al. 1996) and is shown in Figure 1. The nucleotides essential for high affinity binding are the G and U nucleotides in the loop as well as a bulged U near the base of the stem. These residues are conserved in all known *E. coli* SECIS molecules. The structure of the *fdhF* RNA has been determined (Fourmy et al. 2002) and shows that the residues essential for *E. coli* SelB binding, the bulged U and the GU loop sequence, both lie along in the major groove of the RNA and are recognized by SelB. The universal necessity of these conserved nucleotides in the bacterial selenocysteine incorporation mechanism remains unclear. In contrast to the *E. coli* SECIS, the *M. thermoacetica* SECIS retains the conserved G and U nucleotides in the loop, but in the context of a pentaloop sequence as opposed to a tetraloop. Additionally, the *M. thermoacetica* SECIS lacks the bulged U in the stem (Fig. 1). This may be the key

factor that lowers the affinity of the *M. thermoacetica* SelB for its SECIS by a factor of 1000 compared with the *E. coli* SelB/SECIS pair (Huttenhofer et al. 1996; Yoshizawa et al. 2005).

While providing important structural information about the nature of the SelB SECIS binding domain, the structures of the *M. thermoacetica* SelB SECIS complex and free *E. coli* SECIS leave some questions open. What is the basis for the greater affinity of the *E. coli* SECIS-SelB interaction compared to that for *M. thermoacetica*? Are other WH domains aside from the C-terminal one interacting with the SECIS?

In the structure of the *M. thermoacetica* SelB-SECIS complex, the *M. thermoacetica* SECIS shows large-scale structural changes compared to the unbound *E. coli* SECIS. For the bound *M. thermoacetica* SECIS, positions one and four in the loop form a Watson–Crick base pair, the uracil at position three is stacked on this base pair, and the nucleotide at position five is bulged out. The guanine at position two in the loop is the major site for protein recognition (Yoshizawa et al. 2005). In contrast, in the free *E. coli* SECIS, positions one and four form a non-Watson–Crick base pair, the guanine and uracil in positions two and three are on opposite sides of the stem, and there is no bulged out nucleotide in this tetraloop (Fourmy et al. 2002). Clearly there is a large amount of conformational plasticity in the SECIS loop region. However, it is not known how much of the structural differences observed arise from the SECIS molecules having different loop sequences or as a result of conformational change in the SECIS with SelB binding. Neither of the SECIS molecules has had its structure determined in both the free and bound form. Here we present the structure of the free *M. thermoacetica* SECIS determined using NMR methods. This SECIS-MT structure allows a direct comparison between a free and bound SECIS and enables us to see how its conformation changes with SelB binding.

RESULTS

SECIS-MT RNA forms a monomer under native conditions

To test whether the SECIS-MT construct used in this study is a monomeric hairpin or a duplex, a sample was run on a 20% acrylamide gel under native conditions (Supplemental Fig. 1). On this gel, the SECIS-MT sample in H_2O (Supplemental Fig. 1, lane 2) gave rise to one strong band migrating slightly slower than an 8 base-pair (bp) marker (Supplemental Fig. 1, lane 1). This is consistent with the 19 nucleotide (nt) SECIS-MT forming a hairpin conformation. As a control, a sample of SECIS-MT in 1 M NaCl was also analyzed (Supplemental Fig. 1, lane 3). The high salt concentration should promote duplex formation. In this sample, a mixture of two bands was observed. A minor

band migrated the same distance as the SECIS-MT sample in lane 2 (Supplemental Fig. 1), while a more intense band migrated more slowly. We attribute the slower migrating band to be the duplex form of SECIS-MT. We conclude that under the conditions used for the NMR studies, the SECIS-MT existed as a hairpin.

Aromatic anomeric resonance assignment

Using H8 or H6 to H1′ connectivities in a 2D ^{1}H, ^{1}H NOESY acquired in ^{2}H$_2$O, the non-exchangeable ^{1}H signals in the stem of SECIS-MT were successfully assigned (Fig. 2) using established NMR techniques (Wüthrich 1986; Varani et al. 1996; Wijmenaga and van Buuren 1998; Flinders and Dieckmann 2006). The strong H6/H5 cross-peaks in a ^{1}H, ^{1}H TOCSY were used to distinguish between the purine and pyrimidine spin systems. A ^{1}H, ^{13}C HSQC was used to assign the H5 pyrimidine residues as cytosine or uracil. The aromatic–anomeric "walk" assignment starts at G1 and continues to G8, where it is interrupted. It resumes at C13 and terminates at the last residue, C19. It is impossible to connect the H8 (or H6)/H1′ cross-peaks of the loop residues G9, U10, C11, and U12. All but two intranucleotide H8 (or H6) to H1′ cross-peaks are weak in the SECIS-MT, which is consistent with the formation of anti base conformations. The two exceptions are G9H8/G9H1′, which is strong and adopts a syn conformation, and G1H8/G1H1′, which is of a medium intensity. Also of note is the very sharp H6/H5

signal in both the NOESY and TOCSY for U12, suggesting a high degree of flexibility for this nucleotide.

Assignment of ribose residues

Complete ^{1}H and ^{13}C assignments of SECIS-MT were obtained using a set of 3D ^{1}H, ^{13}C HCCH-TOCSY, 3D ^{1}H, ^{13}C HCCH-COSY, and ^{13}C-HSQC-NOESY experiments. The 4′ and 5′ ^{1}H and ^{13}C atoms were assigned primarily using the HCCH-TOCSY and 2D ^{1}H, ^{13}C HSQC spectra, as these atoms are quite well dispersed in carbon chemical shift. To assign the 2′ and 3′ ^{1}H and ^{13}C atoms, the HCCH-COSY was used to connect the H2′ to the previously assigned H1′ atoms. We note that the C1′ assignments of the non-base-paired G9, U10, and U12 resonate significantly upfield from the C1′ atoms of the nucleotides that are base-paired. This indicates that for the SECIS-MT the C1′ chemical shift consistently reflects this nature of the nucleotide conformation.

The ribose sugar pucker conformations were determined from a 2D ^{1}H, ^{1}H TOCSY run with a short mixing time. H1′/H2′ and H1′/H3′ cross-peaks were observed at G9, U10, and U12, indicating these nucleotides adopt a C2′-endo conformation. For the other nucleotides in SECIS-MT no H1′/H2′ cross-peaks were observed, with the exception of G1. The sugar pucker conformation is also reflected by the ribose C1′, C4′, and C5′ chemical shift data (Furtig et al. 2003). Analysis of the assignments of SECIS-MT confirmed the TOCSY data that G9, U10, and U12 all adopt a C2′-endo conformation.
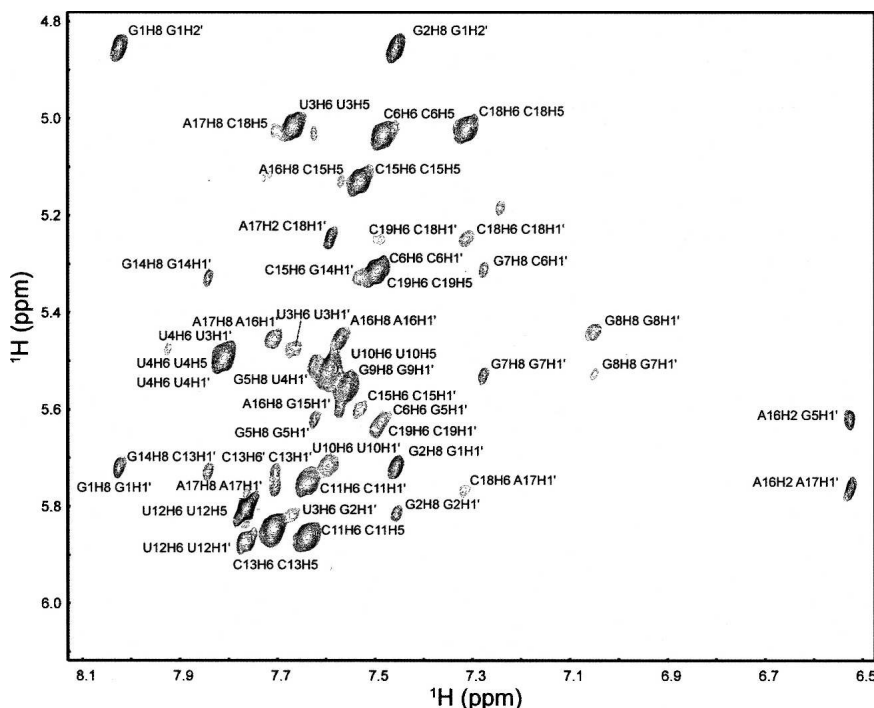

FIGURE 2. Aromatic–anomeric region of the 2D NOESY (τ_m = 200 msec) of SECIS-MT in ^{2}H$_2$O. The spectrum was recorded at 25°C. Assignments are labeled in the spectrum.

Assignment of SECIS-MT exchangeable resonances

A 2D ^{1}H, ^{1}H NOESY and a 3D ^{1}H, ^{15}N NOESY-HSQC with the RNA in H$_2$O were used to assign the exchangeable resonances in SECIS-MT. A series of imino–imino NOEs were observed connecting all the stem base pairs from G2 to G7 (Fig. 1). The terminal G1 imino was not observed, presumably due to its rapid exchange with the solvent. Cytosine amino (H41 and H42) peaks were observed and assigned for all nucleotides (C6, C13, C15, C18, C19) with the exception of C11 in the loop.

An unusually downfield-shifted guanine amino peak, assigned to G8, was observed at ^{15}N resonance of 95 ppm (Fig. 3). The downfield shift indicates that this atom is not involved in a hydrogen bond, but its appearance indicates that it is protected from solvent exchange. In a 3D ^{1}H, ^{15}N NOESY-HSQC
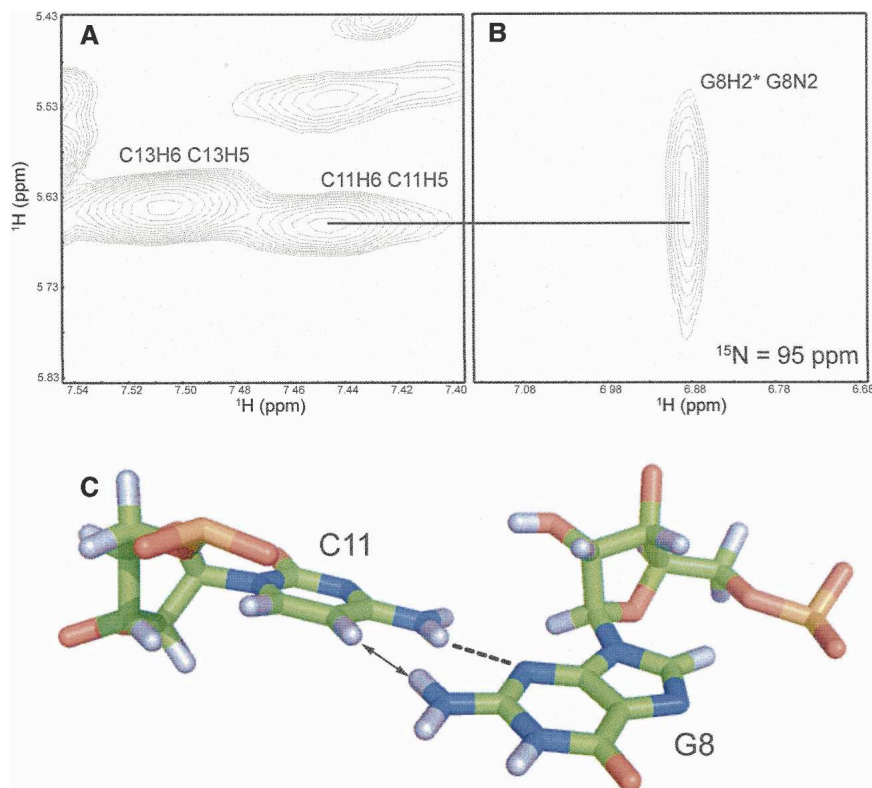
FIGURE 3. G8 and C11 in the loop of SECIS-MT form a noncanonical base pair. (*A*) A region of the 2D NOESY (τ_m = 200 msec) of the SECIS-MT in 90% H_2O, 10% ^2H_2O recorded at 5°C. (*B*) Plane of 3D ^1H, ^{15}N NOESY-HSQC (τ_m = 200 msec) recorded at 5°C showing the G8H21 to C11H5 NOE. (*C*) Depiction of the distances between G8H21 and C11H5 (arrow) and the predicted hydrogen bond between the C11H41 donor and the G8N3 acceptor atoms (dotted line) in the G8–C11 base pair.

this peak showed a strong NOE to C11H5. As there is no G7–G8 imino–imino NOE observed in the 2D H_2O NOESY, and the G8 imino resonance is not observed, a Watson–Crick base pair is not formed. Instead, the G8 and C11 nucleotides form a noncanonical base pair involving a single hydrogen bond from the C11 amino to the G8 N3 atom (Fig. 3).

Structure of the SECIS-MT

The aromatic–anomeric walk and the imino–imino walk in the 2D ^1H, ^1H NOESY acquired in H_2O clearly defines the SECIS-MT stem as being comprised of seven canonical Watson–Crick base pairs beginning at G1–C19 and continuing up to G7–C13. All the bases in the stem assume the *anti* conformation as indicated by weak H8 (or H6)/H1′ NOE cross-peaks, and the ribose sugars assume the *C3′-endo* conformation as indicated by the absence of H1′/H2′ cross-peaks in the short mixing time 2D ^1H, ^1H TOCSY. Both of these structural features are consistent with that of a regular A-form helix.

The structure of SECIS-MT was determined with 171 NOE restraints, 40 hydrogen bond restraints, and 95 dihedral angle restraints. From a starting ensemble of 30 structures, the 10 structures with the lowest energy were selected and are shown in Figure 4. The overall structure is very well defined with a RMSD of 0.25 ± 0.16 Å, with the loop nucleotides having a RMSD of 0.39 ± 0.20 Å (Table 1).

The SECIS-MT loop is best described as a pseudo-diloop. Of the five loop nucleotides (G8–U12; Fig. 1), G8 and C11 base-pair to cap off the stem and U12 is flipped out leaving G9 and U10 to form the turn. A very strong H8/H1′ intranucleotide NOE indicates that G9 is in a *syn* conformation while the other loop nucleotides (G8, U10, C11, and U12) are all in an *anti* conformation. The base of G9 is stacked on G8, with its Watson–Crick edge pointed to the minor groove (Fig. 5). The U10 base is on the opposite side of the loop than the G9 base and points out toward the minor groove. The structure of the backbone turn is defined by a set of NOEs including those between G9H8 and U10H2′, G9H8 to U10H3′, G9H8 to U10H4′, G9H1′ to U10H5″, G9H4′ to U10H6, G9H1′ to U10H2′, G9H1′ to U10H4′, and G9H1′ to U10H3′. There is only a single internucleotide NOE involving the U10 base (G9H4′ to U10H6) and it is consequently more poorly defined in our final ensemble than the other loop residues. A strong U10H6 to U10H2′ NOE helps to orient the U10 base on the minor groove side of the loop. Additionally, there are no internucleotide NOEs involving the base of U12. This residue is poorly structured, its base is extruded from the stem–loop, and it is conformationally flexible. The dynamic nature of this base is reflected in the very sharp H5/H6 correlations in both the NOESY and TOCSY spectra (Fig. 2) of SECIS-MT. The extruded nature of U12 is also indicated by internucleotide NOEs between C11 and C13 (C11H2′ and C13H6, C11H2′ and C13H5′) typical of adjacent nucleotides in an A-form helix.

NMR assignments and chemical shift perturbation in SelB-MT with SECIS-MT binding

In order to understand better how the SECIS-MT RNA hairpin interacts with the SelB protein, we assigned the backbone residues of the SECIS-binding region of SelB and observed the chemical shift changes upon RNA binding. The two C-terminal winged-helix-turn-helix domains (508–634) of the full-length SelB protein (SelB$_{508-634}$) that
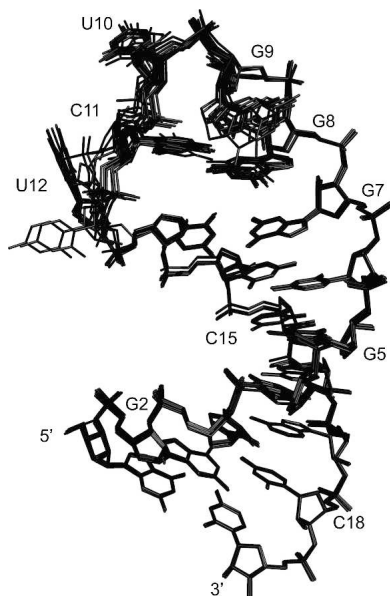
FIGURE 4. A superposition of the 10 final lowest energy SECIS-MT structures. The heavy atoms of all nucleotides, except the unstructured U12, were used for the superposition.

are responsible for RNA binding (Kromayer et al. 1996) were cloned and expressed, and standard NMR techniques were used to assign the backbone atoms. Near complete assignments for SelB$_{508-634}$ were achieved. Of the 121 nonproline backbone amide resonances, only the amides from residues Q575, A598, and R606 were not assigned (Supplemental Fig. 2; Supplemental Table 1).

The effect of RNA binding on the ^{1}H, ^{15}N HSQC of SelB$_{508-634}$ is shown in Figure 6. At first inspection, we can see that some peaks shift with RNA binding, while some are not affected. A more detailed analysis of the effect of RNA binding is shown in Figure 7. Here the chemical shift perturbations are measured and both graphed (Fig. 7A) and mapped out onto the structure of the SECIS-MT/ SelB$_{508-634}$ structure (Fig. 7B). In order to help quantify the chemical shift changes for both the ^{1}H and ^{15}N resonances, a mean chemical shift change was determined as well as a standard deviation. The residues whose amide ^{1}H or ^{15}N chemical shift experienced the most significant chemical shift changes (greater than two times the standard deviation from the mean chemical shift change) upon SECIS binding are Q532, G594, A601, S604, S605, and V630. Also showing large chemical shift changes (1.5–2 times the standard deviation from the mean chemical shift change) are E538, L602, D617, V625, and V631. Aside from these resonances that shift, a number of amide peaks could not be observed in the RNA-bound state. These are the amides from A596, E597, R599, K607, N608, V609, L613, R623, R624, and G626. Presumably these are close to or at the binding site and are exchange broadened due to contact with the RNA.

Chemical shift perturbation in SelB-MT with SECIS-Ecoli binding

Due to the different nature of the *E. coli* and *M. thermoacetica* SECIS sequences, the binding of the *E. coli fdnG* SECIS to SelB$_{508-634}$ was also monitored by amide chemical shift changes. For this pairing the binding constant is not known, but binding clearly does occur, as shown by changes in the ^{1}H, ^{15}N HSQC spectrum (Supplemental Fig. 3). In this case, chemical shift changes were not measurable, as all of the signals affected by RNA binding appear to be in intermediate exchange on the NMR timescale and disappear in the bound HSQC spectrum. The amide signals disappeared for residues Q532, G594, A596, E597, R599, D600, S604, S605, K607, N608, L610, L612, V619, R624, V625, G626, D627, R629, V630, and V631. These amino acid changes are very similar to those observed with SECIS-MT binding and indicate that SelB-MT binds both SECIS molecules at the same site on the protein.

DISCUSSION

In this study, we used NMR methods to determine the structure of the free SECIS-MT and to investigate the interaction between SelB-MT and SECIS molecules. The structure of the free SECIS-MT permits the direct comparison of the same SECIS molecule in both the free and SelB-bound conformation. Chemical shift perturbation data for

TABLE 1. Statistics for the free SECIS-MT RNA ensemble of structures

NMR distance and dihedral angle restraints	
Distance restraints	
Total NOE	172
Intraresidue	101
Interresidue	
Sequential ($|i - j| = 1$)	61
Nonsequential ($|i - j| > 1$)	10
Hydrogen bonds	40
Total dihedral angle restraints	
Sugar pucker	32
Backbone	63
Structure statistics	
Violations	
Distance restraints	0.087 ± 0.004 Å
Dihedral angle restraints	$0.23 \pm 0.07°$
Deviations from idealized geometry	
Bond lengths	0.0100 ± 0.0003 Å
Bond angles	$1.18 \pm 0.01°$
Impropers	$0.90 \pm 0.04°$
Average pairwise RMSD	
All heavy atoms (nucleotides 1–19)	0.25 ± 0.16 Å
Stem heavy atoms (nucleotides 1–7, 13–19)	0.09 ± 0.02 Å
Loop heavy atoms (nucleotides 8–11)	0.39 ± 0.20 Å

The ensemble consists of 10 structures with one NOE violation > 0.5 Å and no dihedral angle violations > 5°.
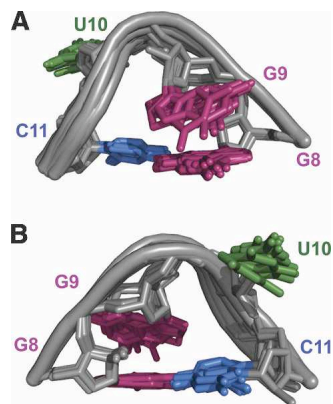
FIGURE 5. Structure of the SECIS-MT loop region. Guanine residues are colored in purple, uracil residues in green, and cytosine residue in blue. These are the same colors as in Figure 1.

SelB-MT with SECIS binding provides evidence for the involvement of the WH domain 3 in the binding of SelB to SECIS molecules.

## Structure of SECIS-MT and similarity to SECIS from *E. coli*

The structure of the free form of SECIS-MT is well defined by the NMR data presented here. SECIS-MT forms an A-form stem that is capped by a noncanonical base pair between G8 and C11 in the loop. The backbone turn is made by a s*yn*-G9 and U10. The fifth loop nucleotide, U12, is flipped out and disordered.

Despite the sequence differences between them, the SECIS-MT RNA described here shares a number of similar features with the *E. coli fdhF* SECIS structure determined by Fourmy et al. (2002). With the U12 bulged out in the SECIS-MT structure, the remaining loop nucleotides can be overlayed with an RMSD of 1.8 Å (Fig. 8). This RMSD drops to 1.0 Å if the poorly ordered U10 is not included in the overlay. In both SECIS structures, the guanine in the second loop position (G9/G12) adopts a *syn* conformation. Additionally, in both structures a noncanonical Watson–Crick loop base pair between guanine and cytosine residues caps off the stem. The apical uracil residue is poorly defined in both SECIS molecules, suggesting a large amount of structural flexibility for this residue.

The determination of this free SECIS-MT structure now permits a direct comparison between the free and RNA-bound SECIS-MT structure (Yoshizawa et al. 2005). It is clear that there is substantial rearrangement of the loop structure upon SelB-MT binding its SECIS partner. The noncanonical G8-C11 base pair in the free SECIS changes to a regular Watson–Crick base pair when the complex is formed. Additionally, the U10 base flips from the minor groove side of the loop to the major groove side, stacking on the top GC base pair. Perhaps the disordered nature of U10 in the free SECIS aids in this rearrangement. The nucleotide at the apex of the loop, G9, is a major recognition site for the protein and switches from a *syn* to *anti* conformation upon SelB binding. There is some similarity between the free and bound SECIS-MT structures; in both, position five of the loop (U12) is flipped out and unstructured. It is unclear what role, if any, U12 plays in binding, but its appearance is analogous to the nucleotide in loop position five of the Smaug recognition element (SRE) that is also flipped out and unstructured in both the free and protein-bound forms of the SRE
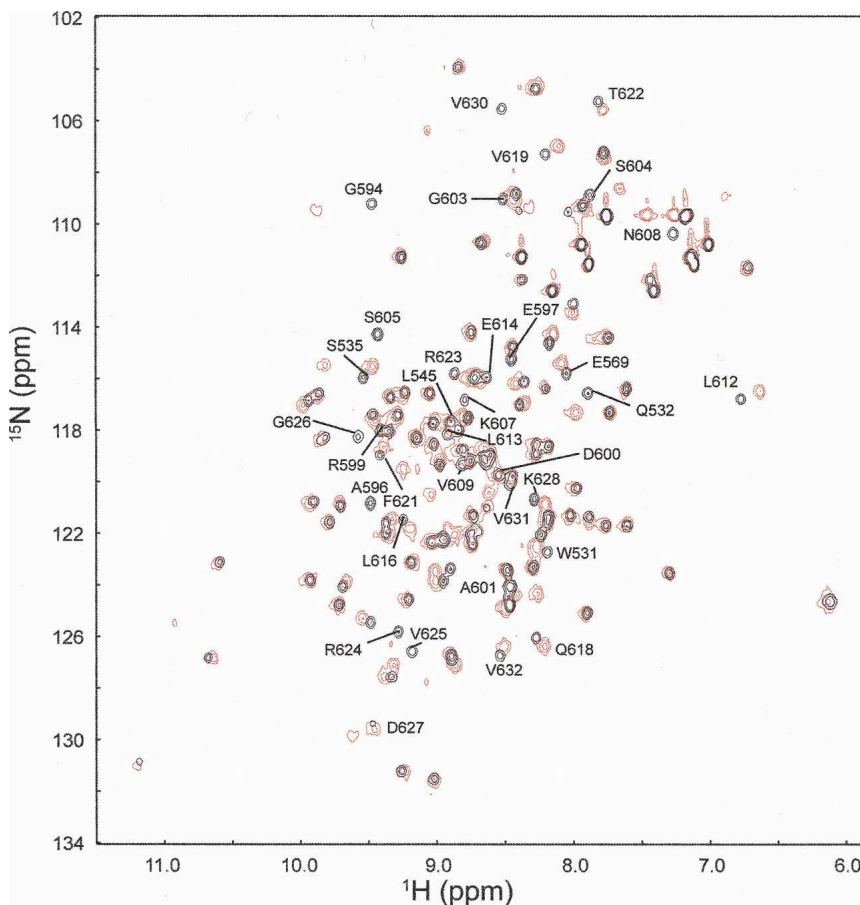
FIGURE 6. Spectral overlay of the SelB$_{508-634}$ ^{1}H, ^{15}N HSQC (black) and the SelB$_{508-634}$/SECIS-MT complex ^{1}H, ^{15}N HSQC (red). Spectra were recorded at 25°C. Peaks that exhibit the largest chemical shift changes (greater than two standard deviations from the average ^{15}N or ^{1}H chemical shift change) or disappear in the bound form are annotated.
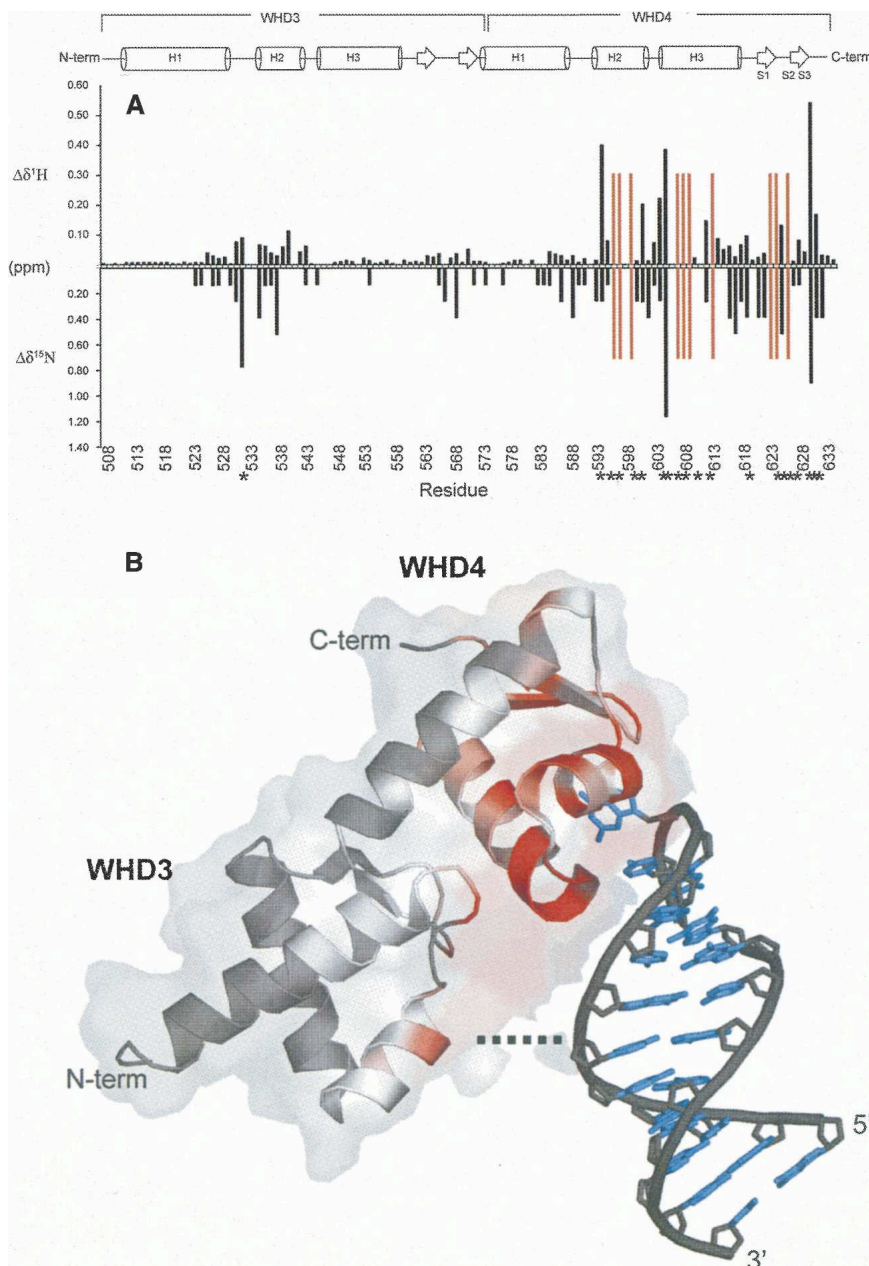
FIGURE 7. Chemical shift perturbation of SelB$_{508-634}$ with SECIS-MT binding. (A) Chemical shift perturbations of SelB$_{508-634}$ represented in a histogram in terms of $\Delta\delta^1H$ (positive y-axis) and $\Delta\delta^{15}N$ (negative y-axis) with residue number on the x-axis. The location of residues whose signals disappear with SECIS-MT binding are indicated with red bars. Asterisks along the x-axis indicate the location of residues in SelB$_{508-634}$ whose signals disappear with SECIS-Ecoli binding. The secondary structure elements of SelB$_{508-634}$ are drawn at the *top* of the diagram. (B) Chemical shift perturbations observed mapped onto the structure of the SelB$_{508-634}$/SECIS-MT complex (pdbid: 1WSU.pdb) (Yoshizawa et al. 2005). Coloring on the structure is the degree of change in chemical shift observed with SECIS-MT binding. Red indicates the greatest chemical shift change with SECIS-MT and white the least. The dotted line indicates the location where a bulged out U would be found in the stem of the SECIS-Ecoli structure.
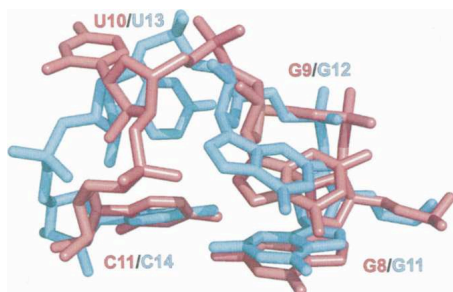
essential feature, as the free SECIS-Ecoli does not have this bulged nucleotide, and in the case of the SRE, the free and protein-bound RNA structures do not show the large-scale structural changes observed with the SECIS molecules.

The SECIS-SelB interaction

The chemical shift data presented here provide a basis for understanding why the *E. coli* SECIS binds 1000 time more tightly to its SelB partner than the *M. thermoacetica* SECIS to its SelB protein. The chemical shift perturbation data we observe in WH4 is consistent with the crystal structure. As seen in Figure 7, the regions of the protein in direct contact with the RNA experience the greatest chemical shift change. However, there are significant chemical shift perturbations readily apparent at the C-terminal end of helix 1, in the loops region between helix 1 and 2 and in the N-terminal region of helix 2 in WH3 of SelB-MT with both SECIS-MT and SECIS-Ecoli binding (Fig. 7). Direct involvement of the third WH domain in RNA binding was not previously seen in the SelB-MT/SECIS-MT structure (Yoshizawa et al. 2005). Presumably, what we observe are short-term interactions that may arise from the natural dynamics of the protein, dynamics at the protein–RNA interface, or from hinge motions at the junction of WH3 and WH4. In a recent study of the SelB-MT/SECIS-MT complex involving all four WH domains Ose et al. (2007) observed a conformational change at the WH2/WH3 hinge with SECIS binding but not at the WH3/WH4 hinge. However, this lack of conformational change does not preclude dynamics occurring in this region of the protein leading to the observed chemical shift perturbations.

We note that this region in WH3 is perfectly positioned to make contact with the bulged U in the stem of SECIS-Ecoli (Fig. 7B). This bulged U is the major structural difference between the SECIS-MT and SECIS-Ecoli and has been demonstrated to be essential for high affinity binding (Liu et al. 1998). Presumably in the SelB-Ecoli/SECIS-Ecoli complex, the third WH domain

(Johnson and Donaldson 2006). It could be proposed that this unstructured nucleotide provides a flexible transition site to enable the structural change that occurs between the free and bound SECIS. However, this is unlikely to be an

FIGURE 8. Overlay of the SECIS-MT and SECIS-Ecoli loop region. The SECIS-MT is colored in pink and the SECIS *E. coli* in cyan. The heavy atoms in the loop superimpose with a RMSD of 1.8 Å. If the poorly ordered U10 is not included the RMSD is 1.0 Å.

is more involved in RNA binding and makes significant contact with SECIS-Ecoli. These new interactions could result in a significantly higher affinity for the SelB-MT/SECIS-Ecoli complex and other SECIS elements that contain a bulged U in the stem.

MATERIALS AND METHODS

RNA synthesis and purification of SECIS-MT

Unlabeled as well as $^{13}C/^{15}N$ labeled SECIS-MT and SECIS-Ecoli RNA were synthesized using the in vitro transcription method (Milligan et al. 1987; Milligan and Uhlenbeck 1989) using a synthetic DNA template designed using Assembly PCR Oligo Maker (Rydzanicz et al. 2005). The RNA product was purified on a denaturing (7 M urea) polyacrylamide gel and subsequently electroeluted. The final yield of RNA samples was typically 3–5 mg from a 40 mL reaction. Prior to analysis, the RNA sample was heated to 95°C and then rapidly chilled. To investigate its oligomeric state the SECIS-MT RNA was run on a 20% acrylamide gel under native conditions.

NMR spectroscopy of SECIS-MT

NMR spectra of the unlabeled SECIS-MT RNA were acquired using a Bruker Avance 600 MHz NMR spectrometer. Sample concentrations varied between 0.5 and 1.0 mM. NOESY spectra in both H_2O and ^2H_2O were acquired with mixing times of 50 msec and 200 msec. 2D TOCSY (Griesinger et al. 1988) with long (73 msec) as well as short (30 msec) mixing time were also acquired. NMR spectra acquired using a 0.7 mM ^{13}C, ^{15}N-labeled SECIS-MT sample included a 3D HCCH-TOCSY (Kay et al. 1993), 2D ^1H, ^{13}C HSQC spectra (Vuister and Bax 1992), and a 3D ^1H, ^{13}C-NOESY-HSQC (τ_m = 100 msec) (Davis et al. 1992) optimized for the aromatic region of the RNA. Additional ^1H, ^{13}C 3D NOESY-HSQC spectra optimized for both aromatic and ribose resonances in the RNA with mixing times of 200 msec and 50 msec and an amino optimized ^1H, ^{13}N 3D NOESY-HSQC were acquired using the Varian 800 MHz NMR spectrometer equipped with a cryogenically cooled probe at the Quebec/Eastern Canada High Field NMR Facility at McGill University. A 3D HCCH-COSY (Bax et al. 1990) was acquired on a Varian 500 MHz NMR spectrometer at McGill University. Data were processed and analyzed using

nmrPipe (Delaglio et al. 1995) and NMRView software (Johnson and Blevins 1994).

Structure calculations

NOE restraints were binned according to their intensity. Strong cross-peaks were assigned a distance between 2 and 2.5 Å, medium cross-peaks were assigned a distance between 2 and 3.5 Å, weak cross-peaks were assigned a distance between 2 and 5 Å, and very weak peaks were assigned a distance between 2 and 6 Å. RNA $\alpha/\beta/\gamma/\varepsilon/\zeta$ dihedral angles were assigned to standard A-form helix values for the stem (nucleotides 1–7,13–19). Dihedral angle restraints were used to constrain the ribose sugars of all stem residues and G8 to the *C3′-endo* conformation, while G9 was constrained to the *C2′-endo* conformation on the basis of a strong H1′ to H2′/H3′ transfer in a short mixing time TOCSY spectrum. The sugar pucker conformations of U10 and U12 were not constrained due to their large mobility. From the observation of a very strong G9H8/G9H1′ NOE, the glycosidic angle for G9 was constrained to a *syn* conformation (60 ± 30°). Hydrogen bond restraints were used to constrain all base pairs in the stem as well as the noncanonical G8–C11 base pair. Structure calculations were performed using XPLOR-NIH (Schwieters et al. 2003). Database potentials were employed to constrain the SECIS-MT stem to A-form geometry (Clore and Kuszewski 2003). From a starting total of 30 structures, the 10 structures with lowest energy were selected to represent SECIS-MT.

$^{13}C/^{15}N$ SelB-MT sample preparation and purification

The DNA sequence coding for an N-terminal His-tagged SelB-MT protein (N-His$_6$-resi-508–634) was inserted into the multiple cloning site of a *lac* repressor controlled expression plasmid, p719THIVB. This plasmid was transformed into a BL21 (DE3) *E. coli* strain. The cells were grown, harvested, and resuspended in a minimal media solution containing ^{13}C glucose (5 g/L) and $^{15}NH_4Cl$ (1 g/L). The cells were induced with 1 mM IPTG at an OD_{600} of 0.7. The solution was centrifuged and the supernatant discarded. The cells were resuspended in lysis buffer (30 mM Tris at pH 8.0, 150 mM NaCl) and lysed with a French pressure cell. The lysate was centrifuged and the supernatant was passed through a 0.22 µm filter and loaded onto a 10 mL nickel affinity column (Amersham). The bound His-tagged doubly labeled SelB-MT protein was eluted from the column using a linear gradient of lysis buffer versus elution buffer (lysis buffer with 400 mM imidazole). $^{13}C/^{15}N$ SelB-MT protein came off the nickel column at a 20% concentration of the elution buffer. SelB-MT was further purified by rerunning the nickel column under denaturing conditions. The eluted fractions were concentrated to 1.5 mM in 0.5 mL of NMR buffer (100 mM phosphate at pH 7.2, 150 mM NaCl, 5 mM MgCl$_2$, 0.5 mM EDTA, 10% ^2H_2O).

NMR assignment of SelB-MT and analysis of complexes with RNA

The backbone of SelB-MT was assigned by use of a series of conventional experiments (^1H, ^{15}N HSQC, ^1H, ^{13}C HSQC, HNCA, CBCA(CO)NH, and HNCACB) recorded at 20°C on the Bruker Avance 600 MHz NMR spectrometer. To prepare the samples of the protein–RNA complexes, an equimolar amount of

either unlabeled SECIS-MT or SECIS-Ecoli in 50 mL of NMR buffer and a dilute ^{15}N-labeled SelB-MT solution, also in 50 mL of NMR buffer, were combined. This solution was then concentrated down to 0.5 mL to make a 0.5–0.7 mM solution of SelB-MT/SECIS-MT or SelB-MT/SECIS-Ecoli. Chemical shift perturbations due to RNA binding were measured using ^{1}H, ^{15}N HSQC experiments.

COORDINATE DEPOSITION

Coordinates for the 10 lowest energy structures of the SECIS-MT have been deposited in the Protein Data Bank (PDB ID 2RLU).

SUPPLEMENTAL DATA

All supplemental material can be found at http://www.yorku.ca/pjohnson/suppmaterial.html.

ACKNOWLEDGMENTS

REFERENCES

Atkins, J.F. and Gesteland, R.F. 2000. The twenty-first amino acid. *Nature* **407:** 463–465.

Atkins, J.F. and Gesteland, R. 2002. The 22nd amino acid. *Science* **296:** 1409–1410.

Bax, A., Clore, G.M., and Gronenborn, A.M. 1990. ^{1}H-^{1}H correlation via isotropic mixing of ^{13}C magnetization, a new three-dimensional approach for assigning ^{1}H and ^{13}C spectra of ^{13}C-enriched proteins. *J. Magn. Reson.* **88:** 425–431.

Clore, G.M. and Kuszewski, J. 2003. Improving the accuracy of NMR structures of RNA by means of conformational database potentials of mean force as assessed by complete dipolar coupling cross-validation. *J. Am. Chem. Soc.* **125:** 1518–1525.

Davis, A.L., Keeler, J., Laue, E.D., and Moskau, D. 1992. Experiments for recording pure-absorption heteronuclear correlation spectra using pulsed field gradients. *J. Magn. Reson.* **98:** 207–216.

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. 1995. NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6:** 277–293.

Flinders, J. and Dieckmann, T. 2006. NMR spectroscopy of ribonucleic acids. *Prog NMR Spec.* **48:** 137–159.

Fourmy, D., Guittet, E., and Yoshizawa, S. 2002. Structure of prokaryotic SECIS mRNA hairpin and its interaction with elongation factor SelB. *J. Mol. Biol.* **324:** 137–150.

Furtig, B., Richter, C., Wohnert, J., and Schwalbe, H. 2003. NMR spectroscopy of RNA. *ChemBioChem* **4:** 936–962.

Gajiwala, K.S. and Burley, S. 2000. Winged helix proteins. *Curr. Opin. Struct. Biol.* **10:** 110–116.

Griesinger, C., Otting, G., Wüthrich, K., and Ernst, R.R. 1988. Clean TOCSY for ^{1}H spin system identification in macromolecules. *J. Am. Chem. Soc.* **110:** 7870–7872.

Huttenhofer, A., Westhof, E., and Böck, A. 1996. Solution structure of mRNA hairpins promoting selenocysteine incorporation in *Escherichia coli* and their base-specific interaction with special elongation factor SelB. *RNA* **2:** 354–366.

Johnson, B.A. and Blevins, R.A. 1994. NMRView: A computer program for the visualization and analysis of NMR data. *J. Biomol. NMR* **4:** 603–614.

Johnson, P.E. and Donaldson, L.W. 2006. RNA recognition by the Vts1 SAM domain. *Nat. Struct. Mol. Biol.* **13:** 177–178.

Kay, L., Xu, G.-Y., Singer, A., Muhnadiram, D., and Forman-Kay, J. 1993. A gradient-enhanced HCCH-TOCSY experiment for recording side-chains ^{1}H and ^{13}C correlations in H$_2$O samples of proteins. *J. Magn. Reson. B.* **101:** 333–337.

Klug, S.J., Huttenhofer, A., and Famulok, M. 1999. In vitro selection of RNA aptamers that bind special elongation factor SelB, a protein with multiple RNA-binding sites, reveals one major interaction domain at the carboxyl terminus. *RNA* **5:** 1180–1190.

Krol, A. 2002. Evolutionarily different RNA motifs and RNA–protein complexes to achieve selenoprotein synthesis. *Biochimie* **84:** 765–774.

Kromayer, M., Wilting, R., Tormay, P., and Böck, A. 1996. Domain structure of the prokaryotic selenocysteine-specific elongation factor SelB. *J. Mol. Biol.* **262:** 413–420.

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

Li, C., Reches, M., and Engelberg-Kulka, H. 2000. The bulged nucleotide in the *Escherichia coli* minimal selenocysteine insertion sequence participates in interaction with SelB: A genetic approach. *J. Bacteriol.* **182:** 6302–6307.

Liu, Z., Reches, M., Groisman, I., and Engelberg-Kulka, H. 1998. The nature of the minimal ''selenocysteine insertion sequence'' (SECIS) in *Escherichia coli*. *Nucleic Acids Res.* **26:** 896–902.

Milligan, J.F. and Uhlenbeck, O.C. 1989. Synthesis of small RNAs using T7 RNA polymerase. *Methods Enzymol.* **180:** 51–62.

Milligan, J.F., Groebe, D.R., Witherell, G.W., and Uhlenbeck, O.C. 1987. Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Res.* **15:** 8783–8798. doi: 10.1093/nar/15.21.8783.

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.

Rydzanicz, R., Zhao, X.S., and Johnson, P.E. 2005. Assembly PCR Oligo Maker: A tool for designing oligodeoxynucleotides for constructing long DNA molecules for RNA production. *Nucleic Acids Res.* **33:** W521–W525.

Schwieters, C.D., Kuszewski, J.J., Tjandra, N., and Clore, G.M. 2003. The XPLOR-NIH molecular structure determination package. *J. Magn. Reson.* **160:** 65–73.

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.

Stadtman, T.C. 2002. Discoveries of vitamin B12 and selenium enzymes. *Annu. Rev. Biochem.* **71:** 1–16.

Thanbichler, M., Böck, A., and Goody, R.S. 2000. Kinetics of the interaction of translation factor SelB from *Escherichia coli* with guanosine nucleotides and selenocysteine insertion sequence RNA. *J. Biol. Chem.* **275:** 20458–20466.

Varani, G., Aboul-ela, F., and Allain, F.H.T. 1996. NMR investigation of RNA structure. *Prog NMR Spec.* **29:** 51–127.

Vuister, G.W. and Bax, A. 1992. Resolution enhancement and spectral editing of uniformly ^{13}C-enriched proteins by homonuclear broadband ^{13}C decoupling. *J. Magn. Reson.* **98:** 428–435.

Wijmenaga, S.S. and van Buuren, B.M.N. 1998. The use of NMR methods for conformational studies of nucleic acids. *Prog NMR Spec.* **32:** 287–387.

Wüthrich, K. 1986. *NMR of proteins and nucleic acids*. John Wiley & Sons, New York.

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.