RNA (1996), 2:354–366. Cambridge University Press. Printed in the USA.

Solution structure of mRNA hairpins promoting selenocysteine incorporation in *Escherichia coli* and their base-specific interaction with special elongation factor SELB

ALEXANDER HÜTTENHOFER,[1] ERIC WESTHOF,[2] and AUGUST BÖCK[1]

[1] Lehrstuhl für Mikrobiologie der Universität München, Maria Ward Str. 1a, 80638 München, Germany
[2] IBMC du CNRS, 15 rue René Descartes, 67084-Strasbourg CEDEX, France

ABSTRACT

On the basis of chemical probing data, the solution structures of RNA hairpins within *fdhF* and *fdnG* mRNAs in *Escherichia coli*, which both promote selenocysteine incorporation at UGA codons, were derived with the help of computer modeling. We find that these mRNA hairpins contain two separate structural domains that possibly also exert two different functions. The first domain is comprised of the UGA codon, which is included within a complex and distorted double-stranded region. Thereby, release factor 2 might be prevented from binding to the UGA codon to terminate protein synthesis. The second domain is located within the apical loop of the mRNA hairpin structures. This loop region exhibits a defined tertiary structure in which no base is involved in Watson–Crick interactions. The structure of the loop is such that, following a sharp turn after G22 (A22 in *fdnG* mRNA), bases G23 and U24 are exposed to the solvent on the deep groove side of the supporting helix. Residues C25 and U26 close the loop with a possible single H-bonding interaction between the first and last residues of the loop, O4(U26) and N6(A21). The bulge residues U17 and U18 (in *fdhF* mRNA), or U17 only in *fdnG* mRNA, point their Watson–Crick positions in the same direction as loop residues G23 and U24 do, and at the same time open up the deep groove at the top of the hairpin helix. Chemical probing data demonstrate that bases G23 and U24 in both mRNA hairpins, as well as residues U17 and U17/U18 (for *fdhF* mRNA) located in a bulge 5′ to the loop, are involved directly in binding to special elongation factor SELB in both mRNAs. Therefore, SELB recognizes identical bases within both mRNA hairpins despite differences in their primary sequence, consistent with the derived 3D models for these mRNAs, which exhibit similar tertiary structures. Binding of SELB to the *fdhF* mRNA hairpin was estimated to proceed with an apparent K_d of 30 nM.

Keywords: chemical probing; mRNA stem loop; protein synthesis; 21st amino acid

INTRODUCTION

Selenocysteine, a cysteine in which the thiol group (SH) is replaced by the selenol group (SeH), represents the 21st amino acid in the genetic code (Böck et al., 1991). In eukaryotic as well as in prokaryotic organisms, several proteins, mostly oxido-reductases, are found that contain selenocysteine residues (Chambers et al., 1986; Zinoni et al., 1986). These selenocysteine residues have been shown to play an essential role in the activity of the respective enzymes, thereby demonstrating the crucial role for the selenocysteine incorporation pathway (Zinoni et al., 1987; Axley et al., 1991). In *Escherichia coli*, selenocysteine residues are co-translationally incorporated into formate dehydrogenase H, N, and O, directed by a specific UGA codon present within the reading frame of formate dehydrogenase mRNAs (for a review see Böck et al., 1991). This incorporation mechanism has also been suggested for the eukaryotic system (Berry et al., 1991). The decoding of UGA as a selenocysteine codon rather than a stop codon requires the presence of an RNA stem-loop structure 3′ adjacent to the UGA codon within the *E. coli* formate dehydrogenase mRNAs (Zinoni et al., 1990). The requirement of the mRNA secondary structure has been shown to be highly specific, including the primary sequence in the loop and stem region as well

Reprint requests to: A. Hüttenhofer, Lehrstuhl für Mikrobiologie der Universität München, Maria-Ward-Str. 1a, 80638 München, Germany; e-mail: uj44221@sunmail.lrz-muenchen.de.

as the correct folding of the RNA hairpin structure (Heider et al., 1992). By this mechanism, the UGA within the reading frame of formate dehydrogenase mRNAs is "recoded" as a sense codon for selenocysteine incorporation rather than serving as a termination signal in protein synthesis.

Furthermore, a specific selenocysteine-tRNA ($tRNA^{Sec}$) containing a UCA anticodon decodes UGA codons as selenocysteine (Leinfelder et al., 1988), prerequisite to the presence of the mRNA stem loop structure. This $tRNA^{Sec}$ is recognized—unlike all other canonical tRNAs—by a special elongation factor, designated as SELB (Baron et al., 1993). SELB has been shown to bind specifically to $tRNA^{Sec}$ as well as to the mRNA stem loop structure 3′ adjacent to the UGA codon in vitro (Baron et al., 1993). In addition, it has been demonstrated that SELB is able to bind GTP (Forchhammer et al., 1989). This formation of a quaternary complex consisting of SELB-GTP-$tRNA^{Sec}$ and the mRNA stem loop is unprecedented in protein synthesis, because all other canonical tRNAs exhibit a ternary complex consisting of EF-Tu-GTP and tRNA, only. It is believed that the tethering of $tRNA^{Sec}$ to the UGA codon, achieved by simultaneous binding of SELB to $tRNA^{Sec}$ and to the mRNA stem loop, promotes the incorporation of selenocysteine into a polypeptide at the ribosome.

Because the presence of the mRNA hairpin is the only element specifying whether a UGA codon is read as selenocysteine or as a signal for termination of protein synthesis, we wanted to elucidate its unique structural features. Therefore, we investigated the solution structure of SELB-binding mRNA hairpins in *E. coli* by chemical and enzymatic probing. To test the generality of our findings, we compared two different RNA hairpin structures within *fdhF* and *fdnG* mRNAs (coding for formate dehydrogenase H and N) to each other and also determined their base-specific interaction with SELB. The results of this study show that these mRNA hairpins exhibit unique structural features at their UGA selenocysteine codon, as well as within their loop region. Despite differences in their primary sequence, identical bases within the upper part of the loop and stem of both mRNA hairpins interact specifically with special elongation factor SELB, indicative of a common tertiary structure.

RESULTS

We investigated the secondary and tertiary structure of the selenocysteine-incorporating hairpins in *fdhF* and *fdnG* mRNAs by chemical and enzymatic probing. For each of the mRNA hairpins, a secondary structure model had been proposed previously using computer folding programs (Zinoni et al., 1990; Berg et al., 1991a). We also probed the flanking regions 5′ and 3′ to the *fdhF* or *fdnG* hairpins to investigate the influence of the sequence context on the structure of the mRNA hairpins. In previous experiments, synthetic sequences ("U-runs") were introduced at these sites for cloning of DNA fragments coding for mRNA hairpins into the *lacZ* gene (Heider et al., 1992).

Chemical probing of *fdhF* and *fdnG* mRNA hairpins with 1-cyclohexyl-3-(2-morpholinoethyl)-carbodiimide metho-p-toluene sulfonate (CMCT), dimethyl sulfate (DMS), and kethoxal (KE)

The three classical chemical probes, CMCT (for N-1 in G and N-3 of U), DMS (for N-1 in A and N-3 of G), and KE (for N-1 and N-2 of G), were employed to map the Watson–Crick positions within *fdhF* or *fdnG* stem loop structures, respectively (Fig. 1A,B). In general, the secondary structure predictions of the *fdhF* and *fdnG* mRNA hairpins based on assessment by computer folding programs could be confirmed; however, there were two significant exceptions.

First, although the UGA codon as well as sequences up and downstream from the mRNA hairpins were predicted not to be involved in Watson–Crick base pairing, we find that these sequences are neither accessible to base-specific probes (CMCT, DMS, or KE) nor to the single-strand-specific nuclease S1 (see below). Because chemical probing was performed at 37 °C and at presumably physiological ionic conditions (see the Materials and methods), we infer that these regions might be base paired in vivo also.

In particular, the UGA codons of both the *fdhF* and *fdnG* mRNAs show identical reactivity toward chemical probes. While the U and the G bases are not accessible to chemical probes CMCT or KE under native conditions, the A of the UGA codon appears to be hyperreactive toward DMS in both hairpins, indicating a bulged region. Also, bases 5′ to the UGA as well as 3′ to the predicted mRNA hairpin structure, proposed to be single-stranded, are not accessible to chemical as well as to enzymatic probes (see below) under native conditions (Figs. 1 A,B, 4).

The second difference to the predicted secondary structure concerns the loop region of the RNA hairpins. This loop region has been shown to be involved in binding to special elongation factor SELB (Baron et al., 1993). A loop consisting of six bases was suggested by folding programs for the *fdhF* hairpin, whereas a loop consisting of five bases was predicted for the *fdnG* hairpin structure. In contrast, only two bases in each loop region of the *fdhF* and *fdnG* hairpin are highly reactive toward chemical probes, namely G23 and U24. The other bases within the loop regions exhibit only moderate reactivity toward KE, DMS, or CMCT. In both mRNA hairpins, the bulged U17 next to the loop is also highly reactive toward modification by CMCT, as is, in addition, U18 in *fdhF* mRNA (Fig. 1A,B).

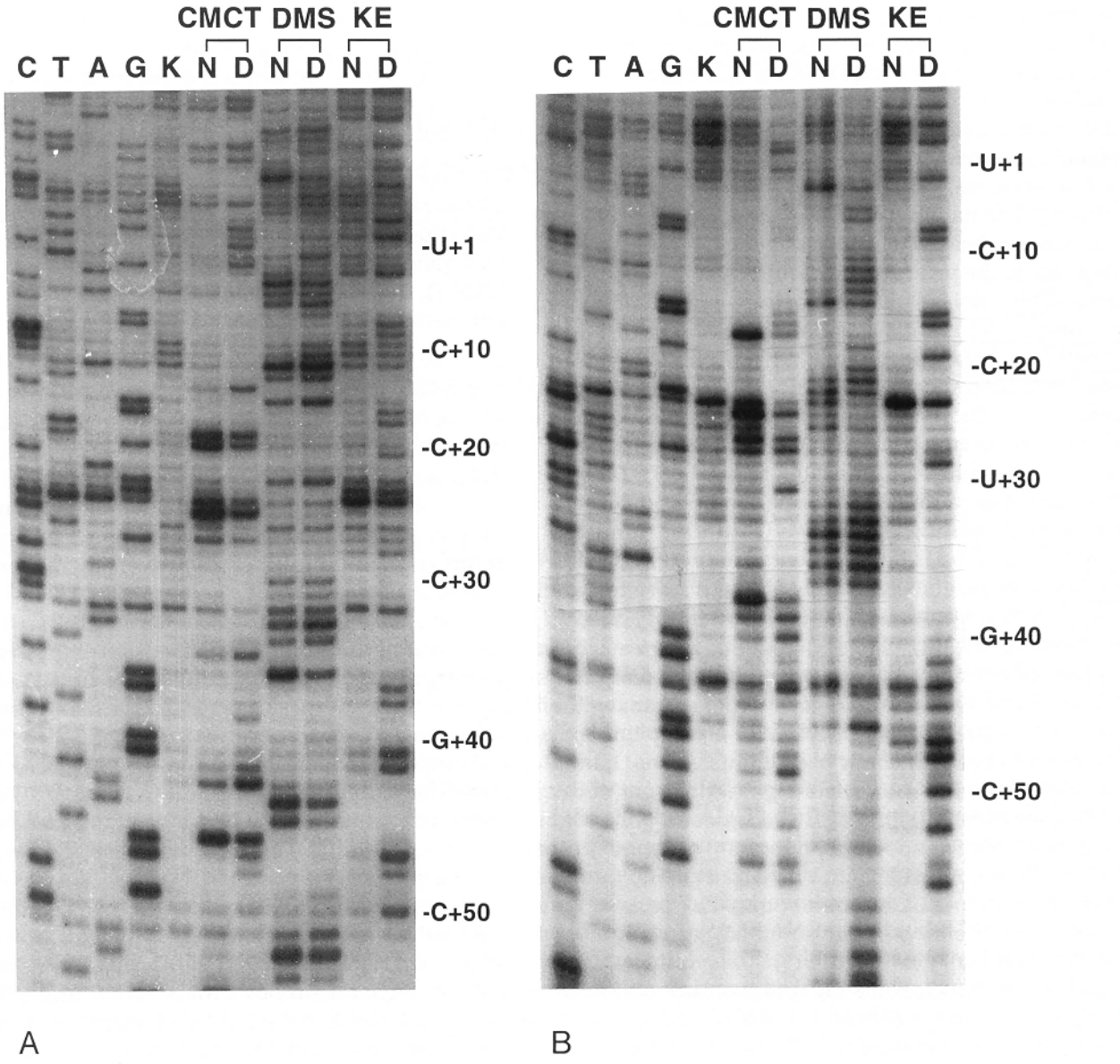

FIGURE 1. Chemical probing of bases of (**A**) the *fdhF* mRNA hairpin or (**B**) the *fdnG* mRNA hairpin with CMCT, DMS, and KE. C, T, A, G, sequencing lanes; K, control lane, no chemical probes added; N, probing under native conditions (37 °C); D, probing under denaturing conditions (60 °C).

We will show below that all highly reactive bases in the upper part of the *fdhF* and *fdnG* stem loop structures, e.g., U17, U18, G23, and U24, are involved in binding to special elongation factor SELB.

Enzymatic probing of *fdhF* and *fdnG* mRNA hairpins with CVN or S1 nuclease

To further support the proposed secondary structure model of both *fdhF* and *fdnG* mRNA hairpins, we performed an S1 and CVN nuclease mapping. In both mRNA hairpin structures, a strong S1 cleavage site between bases G23 and U24 is observed, as well as a weaker cleavage site between U24 and C25 (Figs. 2A,B, 4). G23 and U24 have been shown above to be highly reactive toward chemical probes. This is indicative of an asymmetrical loop structure, because a symmetrical loop structure would also exhibit S1 cleavage between G22/G23 (in *fdhF* mRNA) or A22/G23 (in *fdnG* mRNA). Even at the highest concentration of S1 nuclease, no further cleavage within the RNA fragments outside the proposed hairpin is observed, indicative of base paired regions present within the flanking regions. End-labeling of the *fdhF* mRNA fragment and subsequent treatment with S1 nuclease (data not shown) showed no accessibility of the RNA to digestion with S1 nuclease from about 50 bases upstream of the UGA to at least 30 bases downstream from the stem region. This

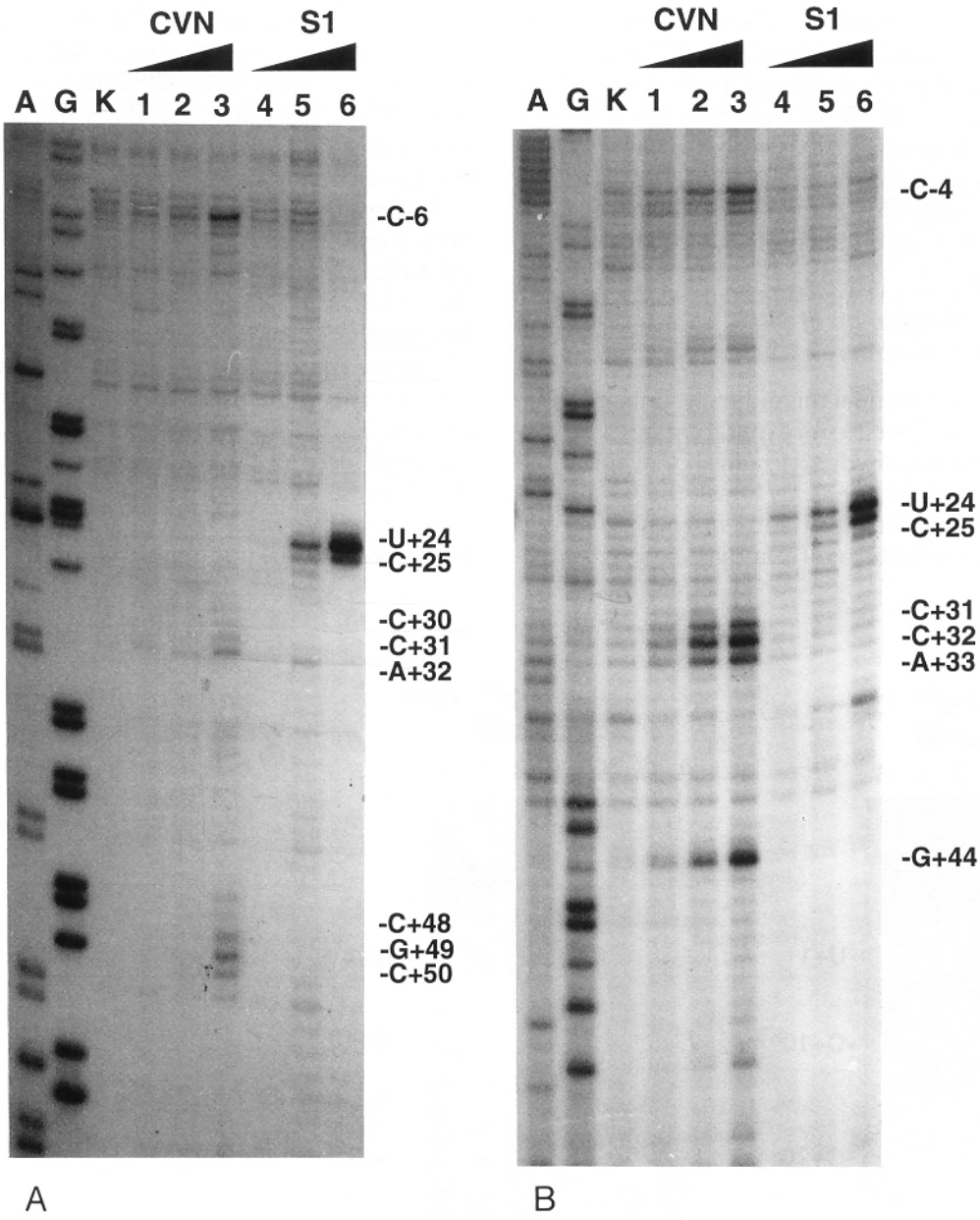

FIGURE 2. Nuclease probing of (**A**) the *fdhF* mRNA hairpin or (**B**) the *fdnG* mRNA hairpin with CVN nuclease (lanes 1–3) or S1 nuclease (lanes 4–6). A, G, sequencing lanes; K, control lane, without nuclease; 1, plus 0.005 U CVN; 2, plus 0.02 U CVN; 3, plus 0.04 U CVN; 4, plus 1U S1; 5, plus 5 U S1; 6, plus 10 U S1.

region of the *fdhF* mRNA, including the mRNA hairpin, which is comprised of about 120 bases, seems, therefore, to be highly structured.

Cleavage by CVN nuclease is observed at the CCA sequence (positions 30–32 or 31–33, respectively) in both hairpins. It is noteworthy that, despite differences in the primary and secondary structure, both mRNA hairpins exhibit a very similar reactivity toward chemical or enzymatic probes within the upper part of the stem loop structures. Differences are more pronounced in the lower part of the hairpins. Although a strong CVN cleavage site is observed between C43 and G44 in *fdnG* mRNA, this cleavage site is missing in *fdhF* mRNA. Instead, in *fdhF* mRNA, CVN cleavage is observed at positions C48, G49, and C50 (Figs. 2A,B, 5).

Pb^{2+} cleavage of *fdhF* or *fdnG* mRNA hairpins

In a number of cases, Pb^{2+} cleavage has been demonstrated to be a useful probe to investigate the tertiary structure of RNA molecules (Krzyzosiak et al., 1988 and references cited therein). Both the *fdhF* and *fdnG* mRNA hairpins exhibit distinct lead cleavage sites. Despite some differences in the intensities of the observed

lead cleavage sites, in both mRNA hairpins, the lead cuts occur near bulged or unpaired regions around the UGA codon, as well as in the upper part of the stem loop structures (Fig. 3A,B). In *fdnG* mRNA, the AACAU-bulged region (positions 33–37) seems to be especially susceptible to lead cleavage (Fig. 3B), whereas the strongest lead cut in *fdhF* mRNA is located opposite to the UGA codon in the 3′ half of the stem (Figs. 3A, 5).

Hydroxyl radical cleavage of *fdhF* and *fdnG* mRNA hairpins

In order to gather potential three-dimensional information on the selenocysteine-incorporating mRNA hairpins, we performed a hydroxyl radical cleavage of the free mRNAs. Hydroxyl radicals attack the C1 and C4 position of riboses and lead subsequently to strand cleavage. Riboses protected from hydroxyl radical attack are thought to be located at the "inside" of the tertiary structure of an RNA molecule, whereas riboses that are susceptible to hydroxyl radical cleavage are located at the "outside" (Latham & Cech, 1989). Therefore, hydroxyl radicals are a useful probe to study the tertiary structure of RNA molecules.

Surprisingly, although no tertiary specific interactions were expected within the mRNA stem loops, distinct riboses were found to be protected within both *fdhF* and *fdnG* mRNAs; protections were determined by quantitation of intensity of bands with a laser densitometer compared to hydroxyl radical cleavage of RNAs under denaturing conditions (data not shown, and Fig. 3A,B). Riboses protected from hydroxyl radical cleavage were located adjacent to phosphate groups, which were identified to be less accessible to iodine cleavage in thiophosphate-substituted *fdhF* mRNA (C. Baron & A. Böck, unpubl. results). As observed with

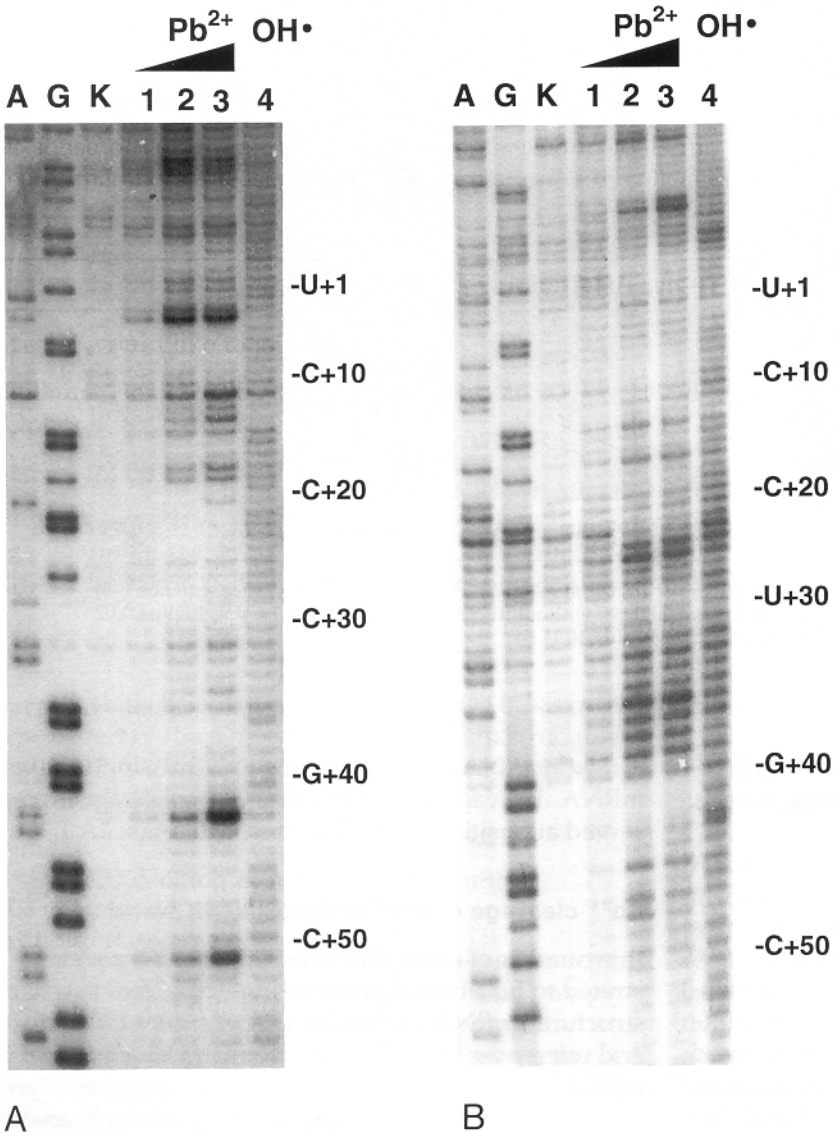

FIGURE 3. Probing of (**A**) the *fdhF* mRNA hairpin or (**B**) the *fdnG* mRNA hairpin with Pb^{2+} (lanes 1–3) or hydroxyl radicals (lane 4). A, G, sequencing lanes; K, control lane, no Pb^{2+} or hydroxyl radicals added. 1, plus 1 mM Pb^{2+}; 2, plus 5 mM Pb^{2+}; 3, plus 20 mM Pb^{2+}; 4, plus hydroxyl radicals (see the Materials and methods).

cleavage by Pb^{2+}, the protected regions are located mainly at and opposite to the UGA codon, as well as in the upper part of the stem loop structures, indicative of two structural domains present within the mRNA hairpins (Fig. 5).

Computer modeling of *fdhF* or *fdnG* mRNA hairpins

The secondary structures deduced from chemical probing for both hairpins are shown in Figures 4 and 5. The data do not give evidence for nonclassical base pairings, whereas the high reactivities of several bases together with the lead cleavages indicate that there are several bulged bases. In some instances, especially in the UGA region of the *fdhF* mRNA, some hydroxyl radical footprints could be rationalized by protections due to those bulged bases. However, this is not always the case. For example, in the central region of the *fdhF* mRNA hairpin, the two A-U pairs flanked by a C···C opposition and a C···A opposition display a mixture of ribose protections and base reactivities that we cannot rationalize. The large bulge in the central part of the

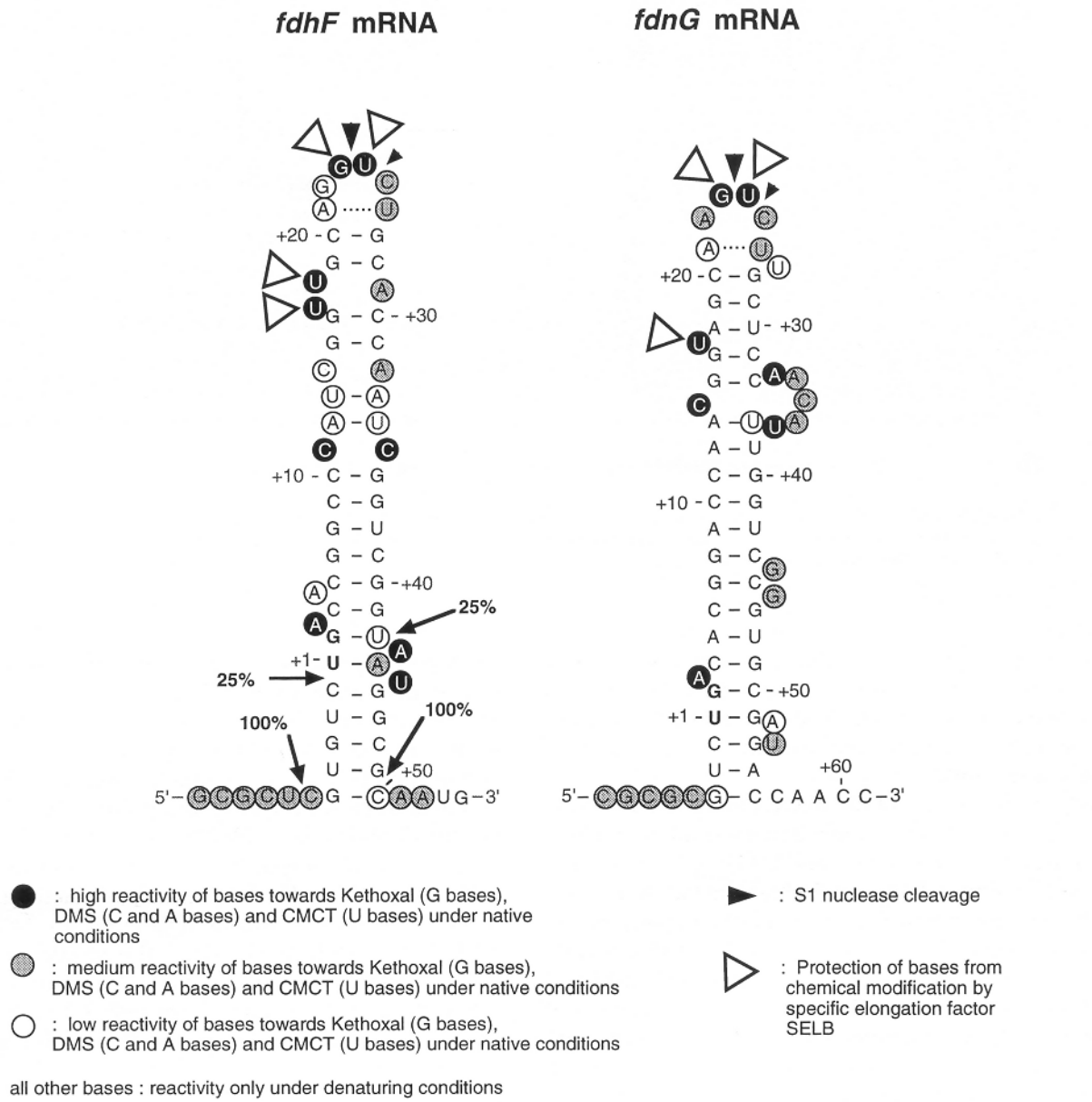

FIGURE 4. Probing the secondary structure of *fdhF* or *fdnG* mRNA hairpins with chemical probes or S1 nuclease and protection of bases from modification by SELB. Reactivities of bases toward chemical probes are indicated by circles. S1 nuclease cleavage is indicated by arrow heads. Protection of bases from chemical modification by SELB is shown by triangles. UGA readthrough and selenocysteine incorporation in various deletion construct of the *fdhF* mRNA sequence is indicated by arrows (Zinoni et al., 1990); level of readthrough within deletion mutants compared to the wild-type construct is given in percentage.

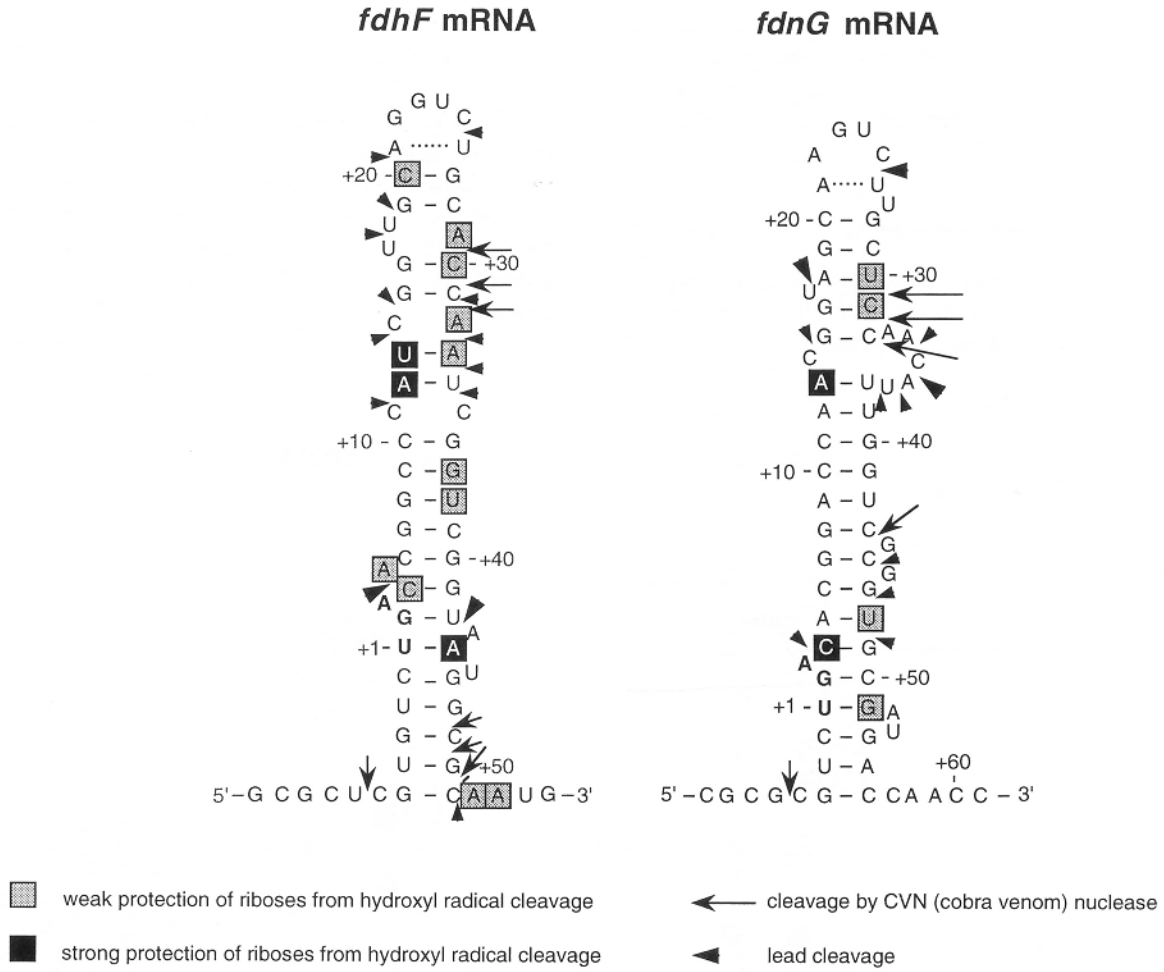

FIGURE 5. Probing the tertiary structure of *fdhF* or *fdnG* mRNA hairpins with Pb^{2+}, hydroxyl radicals, or CVN nuclease. Protection of riboses from hydroxyl radical cleavage is indicated by squares. Cleavage by Pb^{2+} or CVN nuclease is indicated by arrow heads or arrows, respectively.

fdnG mRNA was built so that the bases are reactive with exposition of the phosphates to allow for lead cleavage to occur. This induces a bend of the mRNA, but we cannot be very precise on the extent and direction of the bend.

For the apical part of the mRNA hairpins, after several trials, we settled for the following solution. Following the two G-C pairs, U17 and U18 of *fdhF* bulge out, whereas in *fdnG*, only U17 bulges out, in agreement with their high reactivities. After that asymmetric bulge, two base pairs follow in *fdhF* and three in *fdnG*. In *fdhF*, a weak U18-A29 base pair could have been considered and this would have maintained the same number of base pairs between a bulged residue (U17) and the apical loop. But the lead cleavages between U17 and U18 as well as between U18 and G19, together with the medium reactivity of A29, exclude this possibility. For the loop itself, because of the high reactivities, one can only propose an overall topology. A weak H-bonding interaction between N6(A21) and O4(U26) in *fdhF* and *fdnG* is suggested from the protection data; but a full Watson–Crick pair is probably not present. After residue 22, a sharp turn, similar to the U33-turn of tRNA, was built. The next two residues do not stack on residue 22, as in the anticodon loop of tRNA, but instead point up and perpendicularly to residue 22. Residue 25 in *fdhF* and in *fdnG* closes the loop, but points toward the opposite side and to the shallow groove. This conformation leads to a widening of the deep groove of the RNA, which is amplified by the presence of the bulged residues at positions 17 and 18 in *fdhF* mRNA (or only 17 in *fdnG* mRNA). In *fdnG* mRNA, the additional residue U27 is stacked between the last base pair of the stem, C20-G28, and the pseudo base pair A21-U26. Schematic views of both hairpins are shown in Figure 6.

Base-specific interaction of special elongation factor SELB with *fdhF* or *fdnG* mRNAs hairpins

Previous gelshift experiments, as well as iodine footprinting of SELB-*fdhF* mRNA complexes, pointed to a specific interaction of SELB with the mRNA hairpin ad-

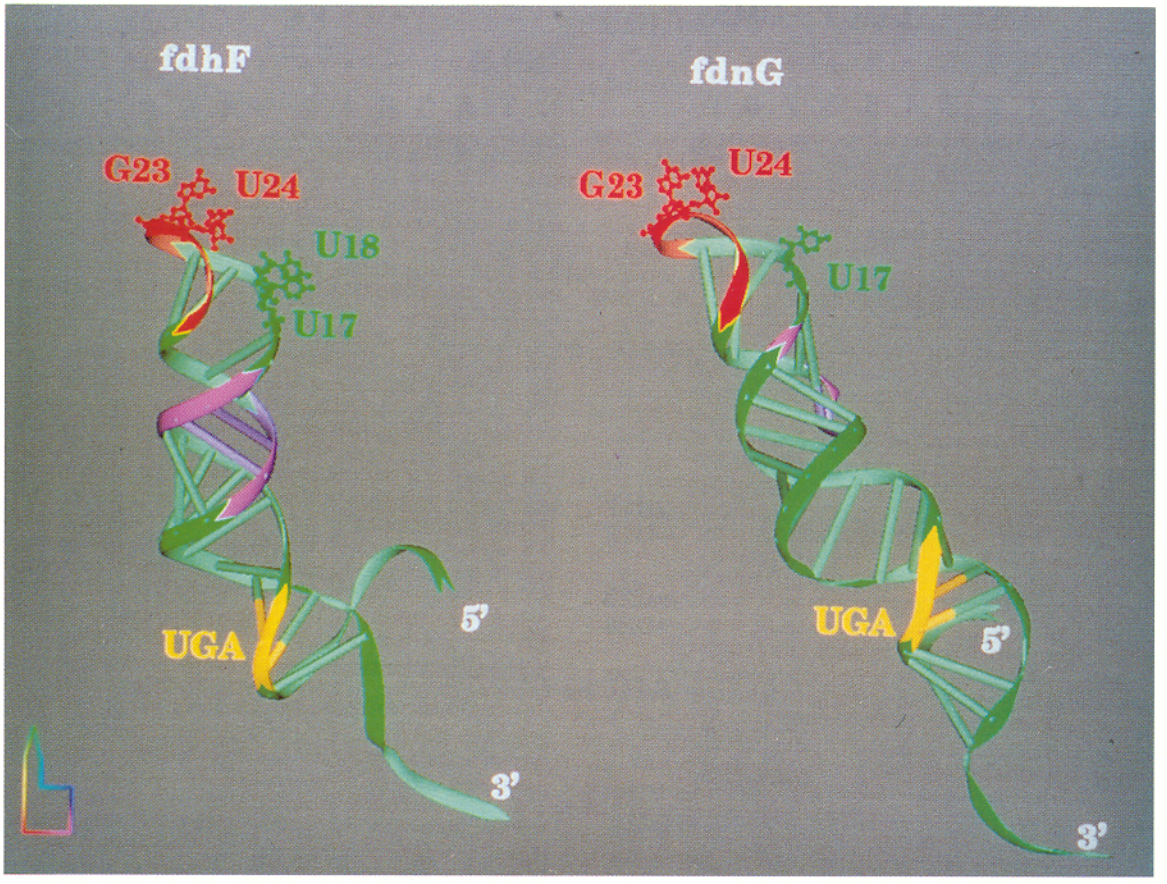

FIGURE 6. Schematic views of the two mRNA hairpins as drawn using the program DRAWNA (Massire et al., 1994). The UGA, which is involved in an intricate structure, is shown in yellow. The apical loop backbone and the residues G23 and U24 protected by SELB are shown in red. The bulged residues U17 and U18, also protected by SELB, are shown in green. The central region of *fdhF* (positions 11–14 and 32–35) and the main central asymmetric bulge in *fdnG* (positions 14 and 33–37) are shown in purple.

jacent to the UGA selenocysteine codon (Baron et al., 1993). Therefore, we wanted to probe the base-specific interaction of the special elongation factor SELB with the *fdhF* or *fdnG* mRNA hairpins. The objective of this study was to get more detailed information on the sequence requirements of the mRNA stem loop structures for binding to SELB, because site-directed mutagenesis of bases in the loop region of mRNA hairpins identified all bases in the loop region as being crucial for selenocysteine incorporation (Heider et al., 1992).

For probing SELB-*fdhF* or SELB-*fdnG* interactions, the same mRNA hairpin constructs were used as for probing the free mRNAs. SELB was added to *fdhF* or *fdnG* mRNAs in the presence of 0.5 mM GTP (see the Materials and methods) and complexes were probed subsequently with KE, DMS, or CMCT (Fig. 7). Footprinting of SELB-*fdhF* or SELB-*fdnG* complexes showed identical bases being protected from chemical modification by SELB in both mRNAs. As shown in Figure 7, G23 and U24, which are located within the loop region of the mRNA hairpins, are both protected strongly from chemical modification by SELB within *fdhF* or *fdnG* mRNAs. Also, U17, which is located in a bulge near the loop region, is strongly protected from modification with CMCT in both mRNAs as is, in addition, U18 within *fdhF* mRNA. As noted above, all bases being protected from chemical modification by SELB are also the ones that exhibit the highest reactivity toward chemical probes in the free mRNA (Fig. 4 and see above).

To assess further the backbone-specific interaction of SELB with the stem loop structure in *fdhF* or *fdnG* mRNAs, we also performed a hydroxyl radical footprinting of SELB–*fdhF* complexes. Surprisingly, despite clear protection of distinct bases within the *fdhF* or *fdnG* mRNA hairpins, only very weak protection of riboses from hydroxyl radical cleavage could be observed (data not shown). The protected region extends from about position G15 to position U26, thereby including the bases protected from chemical modification by SELB.

Affinity of SELB to *fdhF* mRNA as determined by chemical probing

To determine the binding affinity of SELB to the *fdhF* mRNA hairpin, we added increasing amounts of SELB to the *fdhF* mRNA and quantitated the protection of

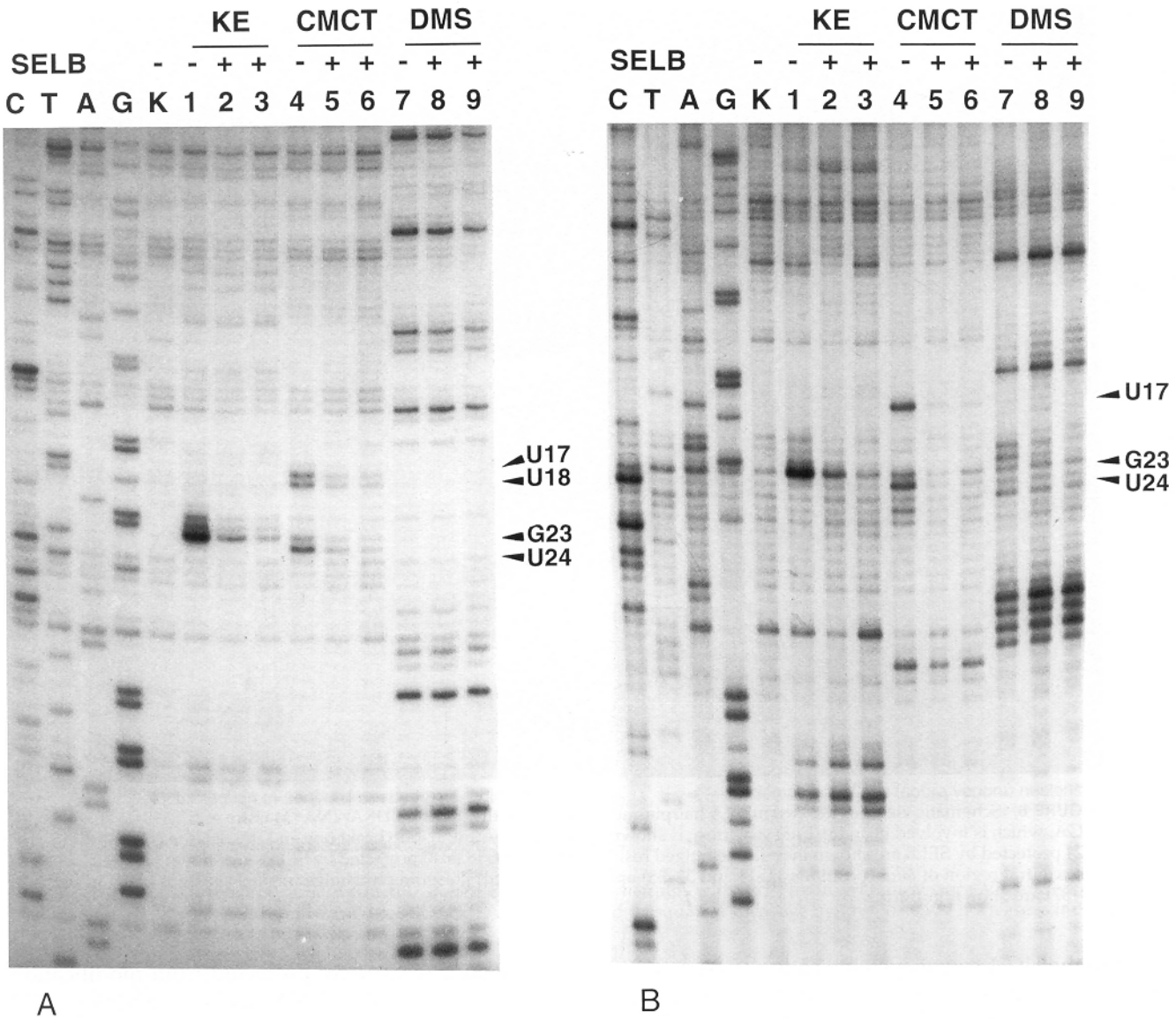

FIGURE 7. Protection of bases from chemical modification by special elongation factor SELB in (**A**) *fdhF* mRNA or (**B**) *fdnG* mRNA. Lanes 1-3, probing with KE; lanes 4-6, probing with CMCT; lanes 7-9, probing with DMS. C, T, A, G, sequencing lanes; K, control lane, no chemical probes added; lanes 1, 4, 7, no SELB; lanes 2, 5, 8, twofold excess of SELB over mRNA; lanes 3, 6, 9, fourfold excess of SELB over mRNA. Arrows indicate bases protected from chemical modification by SELB.

base G23 from modification with KE by SELB. Base G23, located in the loop of the *fdhF* hairpin, was shown to be highly reactive toward chemical modification with KE in its free state, as well as becoming strongly protected from chemical modification in the presence of SELB (see above). The strength of protection of base G23 at SELB concentrations between 16 nM and 320 nM was measured by quantifying the radioactive band corresponding to G23 with a densitometer. From Figure 8, which shows the graph of the peak intensity versus SELB concentration, it can be seen that the peak does not further level off at concentrations higher than 60 nM SELB. Samples showing a linear relationship between peak intensity and SELB concentration were used for a regression analysis to estimate the dissociation constant. From several experiments, the K_d of SELB to the *fdhF* mRNA hairpin was estimated to be 30 ± 3 nM.

DISCUSSION

In this study, we investigated the solution structure of *fdhF* and *fdnG* mRNA hairpins that promote selenocysteine incorporation in *E. coli* as well as their interactions with the special elongation factor SELB. Two distinct domains could be identified and refined structurally: the UGA region and the loop structure of the mRNA hairpins. These domains exhibit a similar structure

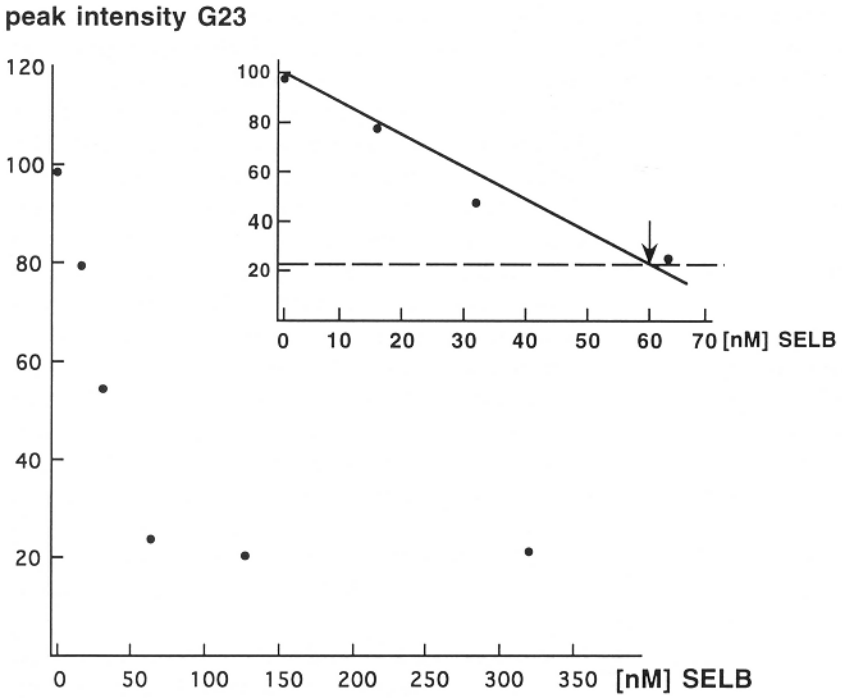

FIGURE 8. Quantification of protection against chemical modification of G23 in the loop of the *fdhF* mRNA hairpin by SELB. The peak intensity of G23 by KE reaches background at 50 nM SELB. Inset shows the same graph with an expanded X-axis and a linear regression reaching the saturation level (dotted line) at ~60 nM (see arrow). The K_d of binding of SELB to the *fdhF* mRNA hairpin was inferred to be 30 ± 3 nM.

within both mRNAs, supporting the generality of our results.

First, we explored more closely the secondary structure of the UGA region by investigating the accessibility of bases to chemical modification with KE, DMS, or CMCT in both mRNA hairpins. We observe that the U and G base of the selenocysteine codon as well as bases 5′ to the UGA are not accessible to chemical probes at physiological ionic and temperature conditions. The same results were observed for bases opposite to the UGA codon 3′ from the stem region. Further evidence for the presence of a double-stranded region comes from the absence of any S1 nuclease cleavage sites within these putatively single-stranded regions. Because we also observe distinct lead cleavage sites at bulged bases around the UGA codon in both mRNAs, this region of the stem loop might exhibit a defined secondary/tertiary structure, because lead cleavage is known to to occur at highly structured sites. Therefore, we suggest that an extended helix is present within the mRNA hairpins in which the U and the G of the UGA codon is base paired, whereas the A is located in a bulge (Figs. 4, 5).

Evidence for an in vivo function of the extended stem structure comes from an in-frame deletion analysis of *fdhF* mRNA, where sequences 5′ to the UGA and 3′ to the stem were deleted and UGA readthrough as well as selenocysteine incorporation was determined (Zinoni et al., 1990). These data show that in-frame deletions of sequences up to position −7 had no effect on UGA readthrough, whereas further deletion of sequences up to position −1 (just up to the UGA codon) reduced UGA readthrough to about 25% (Fig. 4). Similar results were observed for sequences 3′ from the hairpin, which were previously thought to be single-stranded. Deletion of these putative single-stranded sequences from positions +75 up to +50 had no effect on UGA readthrough, whereas deletion up to position +42 reduced the UGA readthrough to about 25% (Fig. 4). Therefore, the disruption of the proposed extended helix at either the 5′ or 3′ half by deleting the complementary sequence reduced UGA readthrough by about 75%. These observations strengthen the significance of our secondary structure model of *fdhF* and *fdnG* mRNA hairpins.

What could be the role of the extended helix in *fdhF* and *fdnG* mRNA hairpins? One function could be that the extended stem facilitates folding of the mRNA hairpin by enhancing the stability of the hairpin structure. Alternatively, enclosing the UGA selenocysteine codon within a stem structure prevents release factor 2 (RF2) from binding to the UGA codon. RF2 is believed to recognize and bind directly to single-stranded UGAN sequences, thereby terminating protein synthesis (Poole et al., 1995). We envision that the folded stem loop structures within *fdhF* or *fdnG* mRNAs are present 3′ next to the ribosomal A-site. Unfolding of the UGA codon in the ribosomal A-site and subsequent immediate decoding by $tRNA^{Sec}$—which is tethered to the UGA codon by SELB binding to the mRNA hairpin and to the tRNA—could efficiently prevent RF2 from binding to the 4-nt termination signal. Interestingly, UGAC, the sequence present in both *fdhF* or *fdnG* mRNA hairpins, was shown to work least efficiently in termina-

tion (Poole et al., 1995). Therefore, the entrapment of the UGA codon in a secondary structure, together with the least efficient termination signal—UGAC—might allow the selenocysteine incorporating system to compete effectively with termination of protein synthesis at these UGA codons.

The second structural domain is located in the loop of the selenocysteine-incorporating mRNA hairpins. We find that most bases in the loop of both mRNA hairpins are only moderately accessible to chemical probes, with the exception of G23 and U24. These bases are the only ones in the loop to exhibit a high reactivity toward modification with KE or CMCT, respectively. These data favor a model where only G23 and U24 are unpaired and exposed, whereas the other bases within the loop are tucked in. The topology of the loop is such that the deep groove is opened up, whereas residues U17 (and U18 in *fdhF* mRNA) and residues G23 and U24 line its borders (Fig. 6).

We finally investigated the base-specific interaction of special elongation factor SELB with *fdhF* or *fdnG* mRNAs. It was shown previously by gelshift assays that the special elongation factor SELB binds to the *fdhF* mRNA hairpin in vitro. Additionally, by iodine footprinting of thiophosphate-substituted *fdhF* mRNA-SELB complexes, it could be shown that the thiophosphate between G23 and U24 is protected from iodine cleavage in the presence of SELB, whereas the thiophosphate between U24 and C25 becomes more reactive to cleavage by iodine (Baron et al., 1993).

We find that bases that are highly reactive in the free mRNA, namely G23 and U24 in the loop and U17 and U18 (for *fdhF* mRNA only) in the bulge, become strongly protected from chemical modification by SELB. G23 and U24 are linked to the phosphate group shown previously to be protected by SELB in *fdhF* mRNA (Baron et al., 1993). Hydroxyl radical footprinting of SELB-mRNA complexes revealed only weak protections within the upper part of the stem loop (data not shown). Therefore, extensive backbone-specific contacts of the mRNA hairpins with SELB seem to be reduced to one phosphate group between G23 and U24. In contrast, strong protections from chemical modification by SELB can be observed in both hairpins at identical bases. This might be indicative of SELB preferentially recognizing distinct bases of the RNA hairpin that are presented to SELB by the three-dimensional structure of the stem loop. Because by an extensive site-directed mutagenesis approach all bases in the loop have been implicated in being crucial for selenocysteine incorporation (Heider et al., 1992), our study demonstrates that, in fact, only a few of them are involved directly in SELB binding, whereas the others are presumably needed for maintaining a defined tertiary structure of the loop. The residues protected by SELB are highly exposed in the derived model (Fig. 6) and line in the deep grove at the top of the hairpin. A similar situation occurs in the BIV Tat-TAR complex (Puglisi et al., 1995), where a widened deep groove at the top of a distorted hairpin binds the peptide.

The dissociation constant of the SELB-*fdhF* mRNA hairpin complex was determined to be 30 ± 3 nM. By this measurement, the binding affinity of SELB to the mRNA secondary structure is one to two orders of magnitude lower than the well-defined interaction between aminoacyl-tRNAs and EF-Tu GTP, and about in the same range as between aminoacyl-RNAs an their cognate synthetases (Schimmel & Söll, 1979; Ott et al., 1989). Preliminary data (A. Hüttenhofer & A. Böck, unpubl.) point to a twofold increase in the affinity of SELB to *fdhF* mRNA hairpin in the presence of selenocysteine-tRNA^{Sec}. The higher affinity of SELB for the mRNA stem loop in the presence of selenocysteine-tRNA^{Sec} might ensure that no unproductive mRNA-SELB complexes without tRNA are present in vivo, which would increase the possibility of chain termination by RF2.

In conclusion, the *fdhF* and *fdnG* mRNA hairpins exhibit two distinct structural domains within their stem loop region that might also serve two different functions. The UGA region was shown to be base paired, thereby possibly preventing RF2 from binding and terminating at the UGA codon. Therefore, this region might serve as an antideterminant for RF2 recognition. In contrast, distinct bases within the structured loop, as well as a base within a bulged region next to the loop, serve as determinants for SELB recognition in both hairpins. By this mechanism, an efficient readthrough of the UGA codon as selenocysteine rather than as a stop codon is accomplished.

MATERIALS AND METHODS

Materials

All reagents were purchased from Boehringer (Germany), except for AMV reverse transcriptase (Appligene, France), DMS and CMCT (Fluka AG, Switzerland), and KE (Upjohn, England).

Cloning of *fdhF* and *fdnG* sequences coding for mRNA hairpins into pT7-5

DNA fragments encoding mRNA hairpin structures in *fdhF* and *fdnG* mRNAs, respectively, were amplified by PCR (Mullis & Faloona, 1987). Plasmids used for PCR amplification were pFM30 for the *fdhF* fragment (Zinoni et al., 1986) and pVJS101 for the *fdnG* fragment (Berg et al., 1991b). Oligonucleotides used for PCR amplification of the *fdhF* mRNA hairpin sequence were:

AM1: 5′-GGCAGGCCTCGCGTGGTACGGGTAACGAA ACC-3′ (5′ primer);
AM2: 5′-GGCGGATCCCGGTATTATCAATTTCGTTAATAG C-3′ (3′ primer).

Oligonucleotides used for PCR amplification of the *fdnG* mRNA hairpin sequence were:

5-*fdnG*: 5′-CGGGAATTCGGGATGCTGGCGGTAGACAAC (5′ primer);
3-*fdnG*: 5′-CGCGGATCCTAGCGTTTTTGATATCCACCC AG-3′ (3′ primer).

PCR-DNA fragments were digested with *Bam*H I and *Sma* I nucleases (for *fdhF*) or *Bam*H I and *Eco*R I nucleases (for *fdnG*) and cloned into vector pT7-5 using standard recombinant DNA methods as described by Maniatis et al. (1989). The *fdhF* hairpin sequence-containing vector was subsequently designated as pAF1, the *fdnG* hairpin sequence-containing vector as pAG1.

T7 transcription of pAF1 or pAG1 vectors

pAF1 or pAG1 DNA vectors were linearized by digestion with *Eco*R I nuclease. T7 transcription reactions were performed in a reaction volume of 200 μL (Milligan et al., 1987). RNA transcripts were purified by electrophoresis on 8% (w/v) polyacrylamide–7 M urea gels. Full-length transcripts were identified by UV shadowing, excised from the gel, passively eluted into 0.3 M sodium acetate, pH 6.2, and precipitated with ethanol. Subsequently, RNAs were dissolved in H_2O and stored at −20 °C.

Binding of mRNAs to the special elongation factor SELB

RNAs (10 pmol) were renatured in reaction buffer A (50 mM sodium cacodylate, pH 7.2, 10 mM $MgCl_2$, and 50 mM KCl) at 60 °C for 1 min and then cooled slowly to 37 °C. SELB (a generous gift from K. Forchhammer and A. Herzog), at concentrations as indicated, was incubated with 0.5 mM GTP at 37 °C for 20 min in buffer A, including 2 mM DTT prior to addition of mRNA. Binding of the *fdhF* or *fdnG* mRNA hairpins to SELB was performed in buffer A in a total volume of 50 μL at 37 °C for 15 min; subsequently, reactions were placed at 4 °C for 5 min.

Chemical probing

Chemical probing (Stern et al., 1988) of free RNA or RNA-SELB complexes was performed by addition of 1.5 μL DMS (1:10 dilution in 95% ethanol), 1.5 μL KE (1:10 dilution of a 37 mg/mL stock solution in H_2O), or 10 μL CMCT (84 mg/mL in H_2O) to a 50-μL reaction mixture in buffer A by incubation at 37 °C for 10 min. For probing RNAs under denaturing conditions, reactions were performed in 50 mM sodium cacodylate, pH 7.2, (for probing with KE or DMS) or 50 mM sodium borate, pH 8.3 (for probing with CMCT), and 1 mM EDTA by incubation at 60 °C for 1.5 min. All reactions were stopped by the addition of 300 μL ethanol followed by rapid mixing. For DMS reactions, 25 μL DMS stop solution (1 M Tris-HCl, pH 7.5, 1 M 2-mercaptoethanol, 0.1 M EDTA) were added prior to addition of ethanol. KE-modified samples were adjusted to 25 mM sodium borate, pH 7.2. The pellets were resuspended in 200 μL 0.3 M sodium acetate, pH 6.2, and extracted once with phenol and chloroform. After precipitation, pellets were resuspended in 5 μL H_2O (for DMS and CMCT samples) or in 5 μL 25 mM sodium borate, pH 7.2 (KE-modified samples).

Hydroxyl radical probing

Hydroxyl radical probing was done according to Hüttenhofer and Noller (1992). Hydroxyl radical cleavage was performed in a total reaction volume of 20 μL at 4 °C for 20 min. After precipitation, samples were dissolved in 5 μL H_2O.

Enzymatic probing with S1 or cobra venom (CVN) nuclease

Digestion of 5 pmol of mRNAs with nuclease S1 (concentration as specified) or CVN nuclease (concentration as specified) was performed in 50 mM sodium cacodylate, pH 6.5, or 50 mM Tris-HCl, pH 7.2, respectively, and 10 mM $MgCl_2$ in a total reaction volume of 10 μL at 20 °C for 5 min. Reactions were terminated by the addition of 2.5 μg carrier tRNA, 100 μL 0.3 M sodium acetate, pH 6.2, followed by two phenol and one chloroform extraction; subsequently, samples were ethanol precipitated and resuspended in 5 μL H_2O.

Lead cleavage

Cleavage of RNAs (10 pmol) with Pb^{2+} was performed as described previously (Krzyzosiak et al., 1988). Reactions were performed in a 20-μL reaction volume at 20 °C for 20 min by addition of lead acetate in concentrations as specified. Reactions were stopped by addition of 2.5 μg carrier tRNA, 5 μL of 0.5 M EDTA, pH 8.2, and 50 μL of 0.3 M sodium acetate, pH 6.2, and precipitated with 300 μL ethanol; subsequently, pellets were dissolved in 5 μL H_2O.

Primer extension reactions of RNAs modified by chemical or enzymatic probing

Primer extension reactions were performed as described by Stern et al. (1988), using 5′ ^{32}P-end-labeled primer AM2 (for *fdhF* mRNA) or 3′ *fdnG* (for *fdnG* mRNA), which are described above. Samples were loaded onto 6% (w/v) polyacrylamide–7 M urea gels. Electrophoresis was performed at 2,000 [V], 25 [mA] for 1.5 h.

Computer modeling

The constructions of the hairpin models were performed on an Evans & Sutherland PS300 display system following the method described by Westhof (1993). The coordinates of the model were subjected to restrained least-squares refinement according to Konnert and Hendrickson (1980) with the programs NUCLIN and NUCLSQ (Westhof et al., 1985) in order to ensure correct geometry and stereochemistry. Coordinates are available upon request through e-mail (westhof@lynch.u-strasbg.fr).

ACKNOWLEDGMENTS

This work was supported by the Deutsche Forschungsgemeinschaft and the Fonds der Deutschen Chemie. We

thank M. Kromayer for reading the manuscript and A. Krol and R. Walczak for sharing results on the SECIS element in eukaryotes prior to publication.

Received January 26, 1996; returned for revision February 26, 1996; revised manuscript received March 1, 1996

REFERENCES

Axley MJ, Böck A, Stadtman TC. 1991. Catalytic properties of an *Escherichia coli* formate dehydrogenase mutant in which sulfur replaces selenium. *Proc Natl Acad Sci USA 88*:8450–8454.

Baron C, Heider J, Böck A. 1993. Interaction of translation factor SELB with the formate dehydrogenase H selenopeptide mRNA. *Proc Natl Acad Sci USA 90*:4181–4185.

Berg BL, Baron C, Stewart V. 1991a. 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.

Berg BL, Li J, Heider J, Stewart V. 1991b. Nitrate-inducible formate dehydrogenase in *Escherichia coli* K-12. I. Nucleotide sequence of the *fdn*GHI operon and evidence that opal (UGA) encodes selenocysteine. *J Biol Chem 266*:22380–22385.

Berry ML, Banu L, Chen Y, Mandel SJ, Kieffer JD, Harney JW, Larsen PR. 1991. Recognition of UGA as a selenocysteine codon in the type I deiodinase requires sequences in the 3′ untranslated region. *Nature (Lond) 353*:273–276.

Böck A, Forchhammer K, Heider J, Leinfelder W, Baron C. 1991. Selenoprotein synthesis: An expansion of the genetic code. *Trends Biochem Sci 16*:463–467.

Chambers I, Goldfarb P, Hampton J, Affara N, Mc Bain W, Harrison PR. 1986. The structure of the mouse glutathione peroxidase gene: The selenocysteine in the active site is encoded by the "termination" codon TGA. *EMBO J 5*:1121–1127.

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

Heider J, Baron C, Böck A. 1992. Coding from a distance: Dissection of the mRNA determinants required for the incorporation of selenocysteine into protein. *EMBO J 11*:3759–3766.

Hüttenhofer A, Noller HF. 1992. Hydroxyl radical cleavage of tRNA in the ribosomal P site. *Proc Natl Acad Sci USA 89*:7851–7855.

Konnert JH, Hendrickson WA. 1980. A restrained-parameter thermal factor refinement procedure. *Acta Crystallogr A 36*:344–350.

Krzyzosiak WJ, Marciniec T, Wiewiorowski M, Romby P, Ebel JP, Giege R. 1988. Characterization of the lead(II)-induced cleavages in tRNAs in solution and effect of the Y-base removal in yeast $tRNA^{Phe}$. *Biochemistry 27*:5771–5777.

Latham J, Cech TR. 1989. Defining the inside and outside of a catalytic RNA molecule. *Science 245*:276–282.

Leinfelder W, Zehelein E, Mandrand-Berthelot MA, Böck A. 1988. Gene of a novel tRNA species that cotranslationally inserts selenocysteine. *Nature (Lond) 331*:723–725.

Maniatis T, Fritsch EF, Sambrock J. 1989. *Molecular cloning: A laboratory manual*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Massire C, Gaspin C, Westhof E. 1994. DRAWNA: A program for drawing schematic views of nucleic acids. *J Mol Graphics 12*:201–206.

Milligan JF, Groebe DR, Witherell GW, Uhlenbeck OC. 1987. Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Res 15*:8783–8798.

Mullis KB, Faloona FA. 1987. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods Enzymol 155*:335–350.

Ott G, Faulhammer HG, Sprinzl M. 1989. Interaction of elongation factor Tu from *Escherichia coli* with aminoacyl-tRNA carrying a fluorescent reporter group on the 3′ terminus. *Eur J Biochem 184*: 345–352.

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.

Puglisi JD, Chen L, Blanchard S, Frankel AD. 1995. Solution structure of a bovine immunodeficiency virus tat-TAR peptide–RNA complex. *Science 270*:1200–1203.

Schimmel P, Söll D. 1979. Aminoacyl-tRNA synthetases: General features and recognition of transfer RNAs. *Annu Rev Biochem 48*: 601–648.

Stern S, Moazed D, Noller HF. 1988. Analysis of RNA structure using chemical or enzymatic probing monitored by primer extension. *Methods Enzymol 164*:481–489.

Westhof E. 1993. Modelling the three-dimensional structure of ribonucleic acids. *J Mol Struct (Theochem) 286*:203–210.

Westhof E, Dumas P, Moras D. 1985. Crystallographic refinement of yeast aspartic acid transfer RNA. *J Mol Biol 184*:119–145.

Zinoni F, Birkman A, Leinfelder W, Böck A. 1987. Cotranslational insertion of selenocysteine into formate dehydrogenase from *Escherichia coli* directed by a UGA codon. *Proc Natl Acad Sci USA 84*:3156–3160.

Zinoni F, Birkman A, Stadtman TC, Böck A. 1986. Nucleotide sequence and expression of the selenocysteine-containing polypeptide of formate dehydrogenase (formate-hydrogen-lyase-linked) from *Escherichia coli*. *Proc Natl Acad Sci USA 83*:4650–4654.

Zinoni F, Heider J, Böck A. 1990. Features of the formate dehydrogenase mRNA necessary for decoding of the UGA codon as selenocysteine. *Proc Natl Acad Sci USA 87*:4660–4664.