LETTER TO THE EDITOR

Recognition of functional groups in an RNA helix by a class I tRNA synthetase

YA-MING HOU,1 MALLIKARJUN SUNDARAM,2 XIAOLIN ZHANG,3 JASON A. HOLLAND,2 and DARRELL R. DAVIS2

1Department of Biochemistry and Molecular Pharmacology, Thomas Jefferson University, Philadelphia, Pennsylvania 19107, USA
2Department of Medicinal Chemistry, University of Utah, Salt Lake City, Utah 84112, USA
3Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA

Keywords: aminoacylation; cysteine; major groove; RNA microhelices; tRNA acceptor stem

RNA helices that recapitulate sequences of the tRNA acceptor stem, including the 3′ NCCA nucleotides, can be substrates for aminoacyl–tRNA synthetases (Frugier et al., 1994; Hamann & Hou, 1995; Martinis & Schimmel, 1995; Quinn et al., 1995). Although the catalytic efficiency of aminoacylation of RNA helices is reduced from that of the full-length parent tRNA, the specificity is maintained. The specific aminoacylation lies in the ability of aminoacyl–tRNA synthetases to recognize functional groups within the RNA helices. Analysis of tRNA–synthetase structures has suggested a general principle (Rould et al., 1989; Ruff et al., 1991; Arnez & Moras, 1997). The class I synthetases, which attach an amino acid initially to the 2′-OH of the terminal ribose, approach the acceptor and NCCA end from the minor groove side. The class II synthetases, which attach an amino acid to the terminal 3′-OH, approach from the major groove side (Arnez & Moras, 1997). The class-specific approach leads to tRNA–synthetase complexes that are near mirror images of each other and provides a structural rationale for the stereochemistries of aminoacylation. We report here the identification of a functional group in the acceptor end of Escherichia coli tRNA^Cys that is important for the class I cysteine–tRNA synthetase. This functional group makes one of the largest energetic contributions to aminoacylation. However, it is located on the major groove side of the acceptor stem. Kinetic analysis of the contribution of this functional group to aminoacylation suggests new features that are not anticipated from the class-specific approach of synthetases.

The acceptor stem of E. coli tRNA^Cys (Fig. 1A) is a substrate for aminoacylation (Hamann & Hou, 1995). For example, an RNA microhelix that contains the acceptor stem, the UCCA end, and the UUCG tetra-loop (Fig. 1B) is specifically aminoacylated with cysteine. An RNA minihelix (not shown) that extends the microhelix by including the TΨC stem is also specifically aminoacylated with cysteine, with a catalytic efficiency (kcat/Km) of aminoacylation similar to that of the microhelix. In these RNA helices, the determinant for aminoacylation is U73 (Hamann & Hou, 1995). Substitution of U73 with A73, C73, or G73 completely eliminates aminoacylation. Further, transfer of U73 to RNA helices of a different specificity confers on the latter the ability to accept cysteine. For example, introduction of U73 to the RNA helix of tRNA^Ala enables aminoacylation with cysteine, even though the major determinant for aminoacylation with alanine (G3:U73) is still present in the acceptor stem. In the transplanted helix, the kcat/Km value of aminoacylation with cysteine is virtually identical to that of the helix of tRNA^Cys, whereas kcat/Km of aminoacylation with alanine is reduced by 30-fold (Hamann & Hou, 1995). The dominant role of U73 in aminoacylation with cysteine is also observed in the full-length tRNA^Cys. In E. coli tRNA^Cys, U73 is the most important nucleotide for aminoacylation, and it accounts for 7.1 kcal/mol of the free energy change of activation (Komatsoulis & Abelson, 1993; Hou, 1997). The significance of U73 in E. coli tRNA^Cys is followed by the GCA anticodon (4.3 kcal/mol) and a G15:G48 tertiary base pair (2.8 kcal/mol) (Komatsoulis & Abelson, 1993; Lipman & Hou, 1998).

U73 is conserved in all cysteine-specific tRNAs and its role in aminoacylation appears conserved in evolution. Studies of the human, yeast, and several microbial cysteine tRNAs have confirmed the significance of U73 (Lipman & Hou, 1998). The prominence of U73...
prompted us to investigate the functional groups in U73 that are important for recognition by cysteine–tRNA synthetase. We used the microhelix of E. coli tRNA^Cys as a framework for the design of substrates (Fig. 1B). Although the k_cat/K_m of aminoacylation of the microhelix is reduced from that of the full-length tRNA by five to six orders of magnitude (Table 1), the dependence on U73 for aminoacylation is the same. Extension of the helix to include the TΨC stem does not improve the k_cat/K_m value; neither does the addition of a second helix that consists of the D-anticodon domain (Hamann & Hou, 1995). This argues that the acceptor stem is an independent domain. Further, kinetic studies show that substitutions in the helix (not at U73) have the same effect as those in the full-length tRNA. For example, alteration of the second and third base pairs of E. coli microhelix decreases k_cat/K_m of aminoacylation by two-fold. The same decrease is observed when alteration is made at the second and third base pairs of E. coli tRNA^Cys (Hou et al., 1995). This indicates a close parallel between the microhelix and the acceptor stem of the tRNA and suggests that the microhelix can be a model system to examine the interaction between the acceptor stem and cysteine–tRNA synthetase.

Kinetic parameters of the wild-type RNA microhelix were determined under the steady-state kinetic condi-
Figure 2. The Lineweaver–Burk plot of aminoacylation with cysteine of the wild-type RNA helix (A) and the m\textsuperscript{5}isoC73 variant (B) as a function of substrate concentration. The linear regression of $y = ax + b$, together with the $R^2$ value (the correlation coefficient), is shown for each plot.

Table 1. Kinetic parameters of aminoacylation with cysteine.

<table>
<thead>
<tr>
<th>RNA Substrate</th>
<th>$k_{cat}$ (s(^{-1}))</th>
<th>$K_m$ ((\mu)M)</th>
<th>$k_{cat}/K_m$ ((\text{M}^{-1}\text{s}^{-1}))</th>
<th>Percentage of charging(^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild-type tRNA\textsuperscript{Cys} transcript</td>
<td>1.3 ± 0.11</td>
<td>1.3 ± 0.12</td>
<td>1.0 × 10(^6)</td>
<td>33</td>
</tr>
<tr>
<td>C73 transcript</td>
<td>(0.48 ± 0.05) × 10(^{-4})</td>
<td>9.6 ± 1.1</td>
<td>0.5 × 10(^1)</td>
<td>30</td>
</tr>
<tr>
<td>Wild-type (U73) helix</td>
<td>(1.8 ± 0.22) × 10(^{-4})</td>
<td>160 ± 6.3</td>
<td>1.12</td>
<td>62</td>
</tr>
<tr>
<td>C73 helix</td>
<td>n.d.</td>
<td>n.d.</td>
<td>&lt;10(^{-5})</td>
<td>n.d.</td>
</tr>
<tr>
<td>m\textsuperscript{5}isoC73 helix</td>
<td>(0.40 ± 0.05) × 10(^{-4})</td>
<td>69 ± 6.4</td>
<td>0.58</td>
<td>65</td>
</tr>
<tr>
<td>dU73 helix</td>
<td>n.d.</td>
<td>n.d.</td>
<td>2.10</td>
<td>60</td>
</tr>
<tr>
<td>dT73 helix</td>
<td>n.d.</td>
<td>n.d.</td>
<td>0.96</td>
<td>60</td>
</tr>
<tr>
<td>s\textsuperscript{4}U73 helix</td>
<td>(2.0 ± 0.48) × 10(^{-4})</td>
<td>1,300 ± 110</td>
<td>0.15</td>
<td>80</td>
</tr>
<tr>
<td>s\textsuperscript{2}U73 helix</td>
<td>n.d.</td>
<td>n.d.</td>
<td>1.00</td>
<td>80</td>
</tr>
</tbody>
</table>

\(^a\)The $k_{cat}/K_m$ values were obtained by one of two methods. First, it was derived from the Lineweaver–Burk equation, based on individual $k_{cat}$ and $K_m$ values (as shown in the table). Values obtained by this method include those for the wild-type transcript, the C73 transcript, the wild-type helix, the m\textsuperscript{5}isoC73 helix, and the s\textsuperscript{4}U73 helix. Alternatively, the $k_{cat}/K_m$ values were estimated from initial rates of aminoacylation at substrate concentrations significantly below the $K_m$ of the respective wild-type RNA. With this method, the individual $k_{cat}$ and $K_m$ values for a substrate were not determined (n.d.). Each value is an average of at least two determinations.

\(^b\)The percentage of charging indicated the percentage of each RNA substrate that was functional for aminoacylation (see Materials and methods). Based on the percentage of charging, the concentration of a substrate was then adjusted accordingly.

The Lineweaver–Burk plot of aminoacylation with cysteine of the full-length transcript of E. coli tRNA\textsuperscript{Cys} (1.27 s\(^{-1}\)) by about 10\(^4\)-fold. The $K_m$ was 160 \(\mu\)M, which is higher than the $K_m$ of the full-length tRNA (1.28 \(\mu\)M) by about 10\(^2\)-fold (Table 1). Both the $k_{cat}$ and $K_m$ values were similar to those of our previous report (Hamann & Hou, 1995). Together, the $k_{cat}/K_m$ of aminoacylation of the microhelix is reduced from that of the full-length tRNA by six orders of magnitude.

We synthesized variants of the RNA microhelix that contained natural and unnatural pyrimidines to replace U73. These variants were to test the atomic groups of the uridine base. The closely related natural pyrimidine, C73, was investigated first. The U73C substitution decreased $k_{cat}/K_m$ of aminoacylation by approximately 10\(^5\)-fold. This effect was estimated by comparing the initial rate of the U73C mutant to that of the wild-type microhelix. The condition of aminoacylation was such that the concentrations of the RNA substrate were at least fivefold below the $K_m$ of the wild type. Under this condition, the initial rate was approximately proportional to $k_{cat}/K_m$.

The 10\(^5\)-fold effect of the U73C substitution was reproduced in the full-length tRNA. The $k_{cat}$ and $K_m$ values for the U73C variant of the tRNA were individually obtained. By using elevated levels of the enzyme, we determined the $k_{cat}$ of the U73C variant as 4.8 × 10\(^{-5}\) s\(^{-1}\) and the $K_m$ as 9.59 \(\mu\)M. These values were valid, as the concentration of the enzyme was maintained at 10-fold below the concentration range of the RNA substrate. Compared to values of the wild-type tRNA, the decrease of $k_{cat}$ by almost 10\(^5\)-fold and the increase of $K_m$ by 10-fold led to a decrease of $k_{cat}/K_m$ by five to six orders of magnitude. Thus, the effect of the U73C substitution in the full-length tRNA is the same as that in the microhelix. This provided additional support that the minihelix domain had quantitatively recapitulated features of the acceptor stem of the full-length tRNA.
U73 differs from C73 at the 3- and 4-positions of the pyrimidine ring. U73 contains the 4-carbonyl group instead of the 4-amino group and has the N3 as a hydrogen donor instead of an acceptor (Fig. 1C). The large effect of the U73C substitution on aminoacylation suggests that the functional groups at the 3- and 4-positions may be the basis for the effect. Several RNA helices were chemically synthesized to test these functional groups (Fig. 1C). The methods for chemical synthesis have all been previously described (Strobel et al., 1994; Kumar & Davis, 1995, 1997). One helix has 5-methyl-isoC (m5isoC) at position 73, which maintains the 4-carbonyl group but converts the N3 into a hydrogen bond acceptor. The m5isoC substitution also replaces the 2-carbonyl group with an amino group and introduces a methyl group to the 5 position of the pyrimidine ring. The kcat/Km of aminoacylation of this helix is dramatically improved over that of the U73 variant. The improved kcat/Km is within twofold of that of the wild type (Table 1; Fig. 2B). Comparison of m5isoC with C (Fig. 1C) suggests that this improvement was not due to the 3-position of the pyrimidine ring, which is common to both nucleosides. Rather, the improvement might come from the 4-carbonyl group, which is also present in U. To determine if the 5-methyl group was responsible for the improved kcat/Km, helices containing dT (deoxy thymidine) and dU at position 73 were synthesized. The helix with dU73 has only a twofold effect (Table 1), therefore any change seen for the dT73 helix was expected to be due to the methyl group. The kcat/Km of the dT73-containing RNA is similar to that of the wild type, suggesting that the 5-methyl group has little effect on aminoacylation.

To determine the importance of the 4-carbonyl group of U73, we investigated an RNA helix containing 4-thiouridine (s4U) at position 73. In s4U, the 4-carbonyl group of uridine is replaced by a thio group but the N3 imino is maintained. The kcat/Km of the s4U-containing RNA is reduced from that of the wild type by 10-fold (Table 1). This modest effect suggests that the 4-carbonyl group is important, but that the interactions important for aminoacylation are tolerant of the greater size and reduced hydrogen-bonding potential of the 4-thio group. To further confirm the significance of the 4-position, but not the 2-position, we tested a helix containing 2-thiouridine (s2U). In s2U, the 3-position is again maintained while the 2-position has the chemically similar thio group. The kcat/Km of the s2U-containing helix is virtually identical to that of the wild type. This supported the data of the m5isoC73 and C73 substitutions and further illustrated that aminoacylation is tolerant to substitutions at the 2-position but not at the 4-position.

The contribution of the 4-carbonyl group of U73 to aminoacylation is largely through a kcat effect. The U73C variant of the full-length tRNA lacks this functional group and suffers from a loss of kcat by almost five orders of magnitude (Table 1). The kcat of the U73C variant of the RNA helix was not measured, due to limitations in the amount of the enzyme needed to offset the severe kinetic defect. However, because of the close parallel between the microhelix domain and the acceptor stem of the full-length tRNA, the kcat of the U73C variant was expected to be impaired. The loss of kcat in the U73C variant and the recovery of kcat in the m5isoC and s4U mutants suggest that the 4-carbonyl group is important in the transition state of aminoacylation.

In standard RNA helices, the 4-carbonyl group of uridine is located in the major groove. Although U73 is a single-stranded nucleotide protruded from the acceptor helix, two solution NMR studies show that the structural environment of U73 is analogous to the major groove of a double helix. One NMR study focused on a variant of E. coli tRNA Met that contained U73 and the first 4 bp identical to those of E. coli tRNA Cys (Puglisi et al., 1994). The second focused on the acceptor stem of Mycoplasma pneumoniae tRNA Cys, which contains U73 and the same first 2 bp as those of E. coli tRNA Cys (Holland et al., 1999). The key finding of both studies is that U73 is well stacked between two nucleotides, the C72 below it and the C74 above it. The continuous stacking of C72, U73, and C74 creates an RNA “A-like” conformation that places U73 on the major groove side of the acceptor stem. Further, the NMR data indicate that C75 and A76 of the CCA end bend over towards G1 such that the major groove position of the 4-carbonyl group of U73 is even more exposed. These major features of U73 are believed to be present in E. coli tRNA Cys, based on the common UCCA end and the first 2–4 bp at the acceptor end. Also, recent kinetic analysis of microhelices of E. coli tRNA Cys and of M. pneumoniae tRNA Cys has demonstrated strong similarities between the two (Y.M. Hou, X.L. Zhang, J.A. Holland, & D.R. Davis, in prep.).

The importance of the 4-carbonyl group of U73 in the major groove is unexpected, as cysteine–tRNA synthetase is a class I enzyme (Eriani et al., 1991; Hou et al., 1991). Studies of functional groups in the acceptor ends, particularly at the terminal 73 position and the 1:72 bp, have largely supported the class-specific principle. For example, crystal structures of the class I glutamine–tRNA synthetase and of the class II aspartate–tRNA synthetase show recognition of the 73 base on the minor and major groove sides, respectively (Rould et al., 1989; Ruff et al., 1991). The class II alanine– and proline–tRNA synthetases both recognize the 1:72 bp in the major groove (Liu et al., 1996; Fischer et al., 1999). Although the alanine enzyme recognizes the determinant G3:U70 in the minor groove (Musier-Forsyth et al., 1991, 1995; Beuning et al., 1997), this determinant is further away from the acceptor end and thus may not comply with the class-specific rule. The only known exception to the class-specific rule is recognition by the class II serine–tRNA synthetase of the 2- amino group of G73, which is located on the
minor groove side of the acceptor end (Breitschopf & Gross, 1996). Additional studies of other synthetases are necessary to gain more insights into the class-specific rule at the acceptor end.

The energetic contribution of the 4-carbonyl group of U73 to aminoacylation is large. It is estimated as 7 kcal/mol based on the change of five orders of magnitude in $k_{cat}/K_m$ from the C73 variant to the $m^5$isoC73 variant (Table 1). Among known groups that are important for aminoacylation (Musier-Forsyth & Schimmel, 1999), the value of 7 kcal/mol is the largest for a single atomic group. However, the actual contribution of the 4-carbonyl group may be smaller, because the 4-amino group of C73 may serve as a negative determinant to prevent aminoacylation by other synthetases. The delineation of how the 4-carbonyl group of U73 contributes to aminoacylation will be important for understanding the catalysis of tRNA aminoacylation by the class I cysteine enzyme. The 4-carbonyl group of U73 is located at the helical end where the major groove is widely accessible to RNA–protein interactions. One possibility is that it could directly contact cysteine–tRNA synthetase to form a transition-state complex. Alternatively, it could form a structure with another part of the RNA during the transition-state of aminoacylation. This RNA–RNA interaction would then be required for the cysteine enzyme to turn the CCA end to the minor groove side of the acceptor helix, where the class I enzyme is expected to approach.

**MATERIALS AND METHODS**

Kinetics of aminoacylation of cysteine was determined under previously established steady-state conditions (Hamann & Hou, 1995). The wild-type transcript was assayed in the range of 0.3–10 μM and with 0.8 nM of E. coli cysteine–tRNA synthetase. The C73 transcript was assayed in the range of 5–50 μM and with 0.5 μM of the cysteine enzyme. The wild-type helix and the $m^5$isoC73 variant were assayed in the range of 20–300 μM with 3 μM of the cysteine enzyme. For all of these RNA substrates, the reported $K_m$ values were in the same range of substrate concentrations that were tested. This gave credence to the accuracy of the $K_m$ values. However, the $s^4$U73 variant was assayed in the range of 50–400 μM with 4.2 μM of the cysteine enzyme. As the reported $K_m$ (1,300 μM) was outside the range of substrate concentrations, it was an estimation. All other microhelix variants were assayed at 30 μM with 3 μM of the cysteine enzyme and the initial rates of their aminoacylation reactions were used to estimate $k_{cat}/K_m$.

The percentage of charging for each substrate (in Table 1) was determined by aminoacylation with substrate levels of cysteine–tRNA synthetase over a 30-min time course. The plateau of aminoacylation indicated the amount of the substrate that was functional. This value was then compared to the standard of 1,600 pmol/OD (based on a 100% functional substrate) to determine the percentage of charging. In this study, all of the RNA substrates had been gel purified and ethanol precipitated. The percentage of charging for the trans-


Rould MA, Perona JJ, Söll D, Steitz TA. 1989. Structure of E. coli glutaminyl-tRNA synthetase complexed with tRNA(Gln) and ATP at 2.8 Å resolution [see comments]. Science 246:1135–1142.
