Subtle Functional Interactions in the RNA Minor Groove at a Nonessential Base Pair

Barry S. Henderson,[†,||] Penny J. Beuning,[‡] J.-P. Shi,[†,⊥] Rolf Bald,[§] Jens Peter Fürste,[§] Volker A. Erdmann,[§] Karin Musier-Forsyth,[†,▽] and Paul Schimmel*[,†,#]

Department of Biology, Massachusetts Institute of Technology Cambridge, Massachusetts 02139 Institute für Biochemie der FU Berlin Thielallee 63, 14195 Berlin, Germany Department of Chemistry, University of Minnesota Minneapolis, Minnesota 55455

Received March 18, 1998

We report here a remarkable example of functional interactions with a nonessential base pair in an RNA helix. The results suggest that interactions at a "nonconserved" pair can contribute significantly to the specificity of RNA−protein interactions.

Aminoacyl-tRNA synthetases specific for alanine, aspartic acid, cysteine, glutamine, glycine, histidine, isoleucine, methionine, serine, tyrosine, and valine catalyze the sequence-specific aminoacylation of short helical RNAs that mimic the acceptor stem of their cognate tRNAs.[1−3] In these cases, the helical RNA substrates are composed of as few as four base pairs affixed to the common single-stranded $NCCA_{OH}$ present at the 3′-end of all tRNAs (Figure 1). The specificity of these reactions remains high even though the efficiency of aminoacylation is reduced relative to that observed with full tRNA. The sequences/structures embedded within tRNA acceptor stems constitute an "operational RNA code" for amino acids.[4] This code may have predated the genetic code.

A G3:U70 base pair within the acceptor stem of $tRNA^{Ala}$ is a major determinant for aminoacylation with alanine (Figure 1).[5,6] The unpaired, exocyclic 2-amino group of G3 marks RNA substrates for alanine acceptance and contributes more than 3 kcal/mol to transition-state stabilization for aminoacylation.[7] (This contribution is significantly greater than that which might be attributed to a helix distortion.)[8,9] The 2′-OHs at positions 4, 70, and 71 are also important for RNA recognition by alanyl-tRNA synthetase (AlaRS), with each contributing between 1 and 2 kcal/mol to transition-state stabilization.[10] These elements of RNA recognition form a cluster of atoms that is centered around the essential 2-amino group of G3.

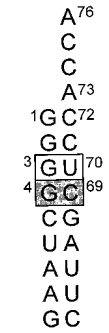

Figure 1. RNA duplex substrate used for aminoacylation experiments. The G3:U70 base pair that is essential for aminoacylation with alanine is boxed. The adjacent 4:69 base pair that was studied here is boxed and shaded. Nucleotides are numbered on the basis of their position in *E. coli* $tRNA^{Ala}$ (GGC isoacceptor).

We speculated that, because functional interactions with the 2′-OH occurred at position 4, base-specific substitutions that changed the character of the minor groove at 4:69 might reveal additional functional contacts in this region. Paradoxically, in extending our previous work, we were particularly surprised to see that many substitutions at 4:69 (e.g., C:G, I:C, G:U, A:U, U:A)[11] had a minimal effect on aminoacylation efficiency (Figure 2A). These observations showed that the enzyme does not recognize a specific base pair at this position. We surmised that, if there are minor groove functional effects at 4:69, then these effects must be subtle. The question was whether substitutions at the 4:69 position could be found that severely affected aminoacylation. For this purpose, the placement of an unpaired exocyclic amino group at the 69-position was of special interest (such as in U4:G69). While this pair places the 2-amino group in the same location on the dyad axis as that obtained with the active G4:U69-containing substrate,[12] the angle at which the amino group protrudes onto the dyad is different.

Chemical synthesis[13] was used to prepare RNA duplex substrates[14] containing either G4:C69 (wild-type (a control)), U4:G69, or U4:I69 base pairs. The remainder of the duplex sequence was based on the first nine base pairs of the *Escherichia coli* $tRNA^{Ala}$ acceptor-TΨC helix plus the single-stranded $ACCA_{OH}$-3′ end (Figure 1). (For this alanine system, results of substitutions in duplex substrates have been in accord with those seen in full tRNA substrates.)[9] The initial rate of aminoacylation for the U4:G69 duplex was about 5% of that of the wild-type duplex (Figure 2B).[15,16] In contrast, the U4:I69 duplex variant was efficiently aminoacylated. Thus, the defect in aminoacylation efficiency observed with the U4:G69 duplex appears to be specifically related to the presence of the unpaired amino group in the minor groove at position 69.

* Corresponding author.
† Massachusetts Institute of Technology.
‡ University of Minnesota.
§ Institute für Biochemie der FU Berlin.
|| Present address: Duke University Medical Center, Department of Microbiology, Box 3020, Durham, NC 27710-3020.
⊥ Present address: Department of Brain and Cognitive Science, Massachusetts Institute of Technology, Cambridge, MA 02139.
▽ Present address: Department of Chemistry, University of Minnesota, Minneapolis, MN 55455.
Present address: The Skaggs Institute for Chemical Biology, The Scripps Research Institute, Beckman Center, 10550 North Torrey Pines Road, La Jolla, CA 92037.

(1) Frugier, M.; Florentz, C.; Giegé, R. *EMBO J.* **1994**, *13*, 2219−2226.
(2) Martinis, S. A.; Schimmel, P. In *Escherichia coli and Salmonella*, 2nd ed.; Neidhardt, F., Ed.; ASM Press: Washington, D. C., 1996; pp 887−901.
(3) Hamann, C. S.; Hou, Y. M. *Biochemistry* **1997**, *36*, 7967−7972.
(4) Schimmel, P.; Giegé, R.; Moras, D.; Yokoyama, S. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 8763−8768.
(5) Hou, Y. M.; Schimmel, P. *Nature* **1988**, *333*, 140−145.
(6) McClain, W. H.; Foss, K. *Science* **1988**, *240*, 793−796.
(7) Musier-Forsyth, K.; Usman, N.; Scaringe, S.; Doudna, J.; Green, R.; Schimmel, P. *Science* **1991**, *253*, 784−786.
(8) Gabriel, K.; Schneider, J.; McClain, W. H. *Science* **1996**, *271*, 195−197.
(9) Beuning, P. J.; Yang, F.; Schimmel, P.; Musier-Forsyth, K. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 10150−10154.
(10) Musier-Forsyth, K.; Schimmel, P. *Nature* **1992**, *357*, 513−515.
(11) Duplexes containing the substitutions C:G, G:U, A:U, and U:A, at 4:69 were analyzed in this work, as described in ref 15. Activity was previously reported for a duplex containing I4:C69 (ref 7).
(12) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1984.
(13) Oligonucleotides were synthesized on either a Pharmacia gene assembler plus or an Applied Biosystems 394 automated synthesizer as described by Musier-Forsyth et al. (refs 14 and 17).
(14) Musier-Forsyth, K.; Scaringe, S.; Usman, N.; Schimmel, P. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 209−213.
(15) Assays were carried out at 20 °C essentially as described by Musier-Forsyth et al. (ref 14) using 4.5 μM preannealed RNA duplexes and 45 nM *E. coli* AlaRS (active site concentration was determined by BioRad protein assay and the adenylate burst assay (Fersht, A. R.; Ashford, J. S.; Bruton, C. J.; Jakes, R.; Koch, G. L.; Hartley, B. S. *Biochemistry* **1975**, *14*, 1−4)).
(16) The perturbations in aminoacylation activity observed were not related to helix stability. Thermal denaturation studies revealed that all duplexes reported here formed stable helices under the assay conditions.

S0002-7863(98)00915-9 CCC: $15.00
Published on Web 08/25/1998

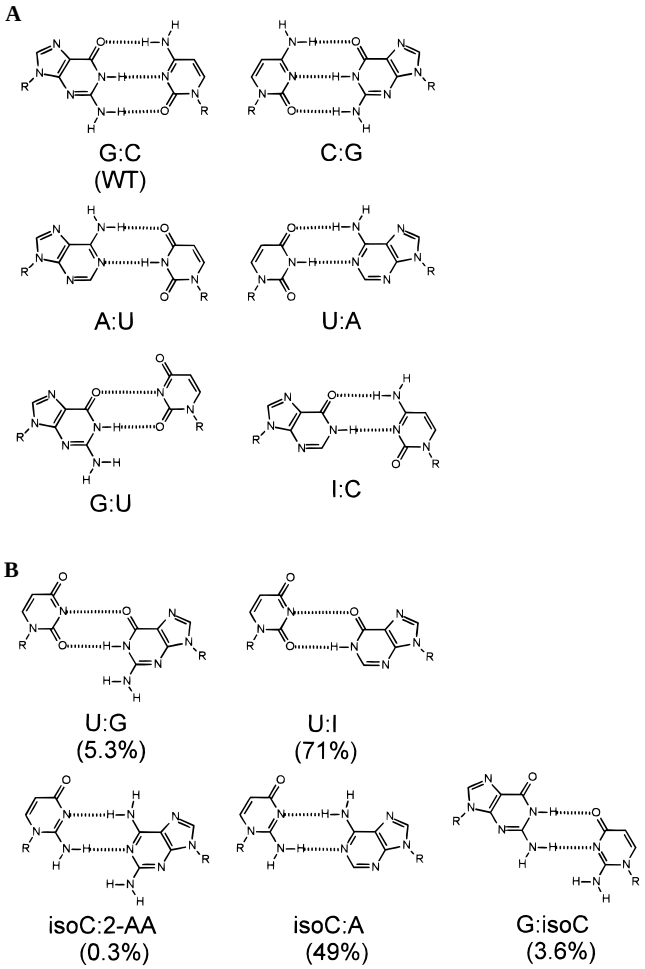

Figure 2. (A) Structure of base-pair substitutions evaluated at position 4:69 that were charged at a rate within a factor of 3 of that of the wild-type duplex. (B) Illustration of base pairs incorporated at position 4:69 of duplex substrates to assess the effect of an unpaired amino group presented in the minor groove at position 69. The efficiency of aminoacylation (as a percentage of apparent k_{cat}/K_m for the wild-type (G4:C69) substrate) for each pair is shown in parentheses.

In contrast to the U4:G69 duplex, the G4:U69 substrate was efficiently aminoacylated with an activity similar to that of the wild-type duplex (Figure 2A). Because the unpaired 2-NH_2 of G is on the dyad axis,[12] it is located in the same position in G:U and U:G. Thus, the deleterious effect of the U4:G69 substitution is unlikely to be due to direct steric blocking of the enzyme by the 2-NH_2.

We also synthesized a duplex containing an isoC:2-AA (isocytidine:2-amino adenosine) pair at position 4:69.[17] The isoC:2-AA pair also places an unpaired 2-amino group in the minor groove at position 69. The activity of this duplex was reduced 300-fold relative to the wild-type duplex (Figure 2B). In contrast, ablation of the unpaired 2-amino group with an isoC4:A69 substitution restored the activity (Figure 2B). Here again, with a completely different base pair, an unpaired 2-amino group at position 69 severely reduced aminoacylation efficiency.

A G4:isoC69 duplex was also evaluated. This base pair is similar to the U:G and isoC:2-AA base pairs in that it positions an unpaired 2-amino group in the minor groove at position 69. In this case, the pendant amino group is attached to the isoC pyrimidine ring. The G4:isoC69 construct was aminoacylated at about 3% of the rate of the wild-type substrate (Figure 2B), further supporting the notion that an unpaired amino group presented from position 69 is detrimental to RNA recognition by AlaRS.[18]

The results demonstrate that alanyl-tRNA synthetase is sensitive to base substitutions at the 4:69 position of duplex RNA substrates. The remarkable subtlety of the interaction at this position undoubtedly enhances the specificity of the overall interaction with the duplex substrate. Possibly, the effect of presenting an unpaired amino group in the minor groove at position 69 perturbs interaction of AlaRS with the minor groove 2′-OH of position 4. This 2′-OH has previously been shown to be thermodynamically significant for RNA recognition.[10] One mechanism by which this perturbation could occur is through a water-mediated hydrogen bond between an unpaired NH_2 at position 69 and the 2′-OH at position 4. (Crystallographic analysis of several G:U-containing RNA helices, including the acceptor stem of *E. coli* $tRNA^{Ala}$, elucidated a network of water-mediated interactions that connect,[19−24] for example, the free 2-amino group of the G with adjacent atomic groups (including 2′-OH) in the minor groove.) Thus, the water-mediated interactions seen in the structures of protein−nucleic acid complexes[25−35] might manifest themselves in functional effects at nonconserved base pairs. However, regardless of the detailed interpretation, these subtle functional effects in the RNA minor groove demonstrate the higher order sophistication of the determinants of specificity of protein−RNA complexes and the constraints on sequences of even nonconserved pairs.

Acknowledgment. This work was supported by Grants GM15539 (to P.S.) and GM49928 (to K.M.-F.) from the National Institutes of Health and Grants to V.A.E. from the Deutsche Forschungsgemeinschaft, the Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie (BMBF), the Deutsche Akademische Austauschdienst (DAAD), and the Fonds der Chemischen Industrie e. V. B.S.H was a postdoctoral fellow of the National Institutes of Health. P.J.B. is a Louise T. Dosdall Fellow.

JA9809152

(17) Musier-Forsyth, K.; Shi, J.-P.; Henderson, B.; Bald, R.; Fürste, J. P.; Erdmann, V. A.; Schimmel, P. *J. Am. Chem. Soc.* **1995**, *117*, 7253−7254.

(18) Attempts were made to synthesize oligonucleotides containing 4-oxo-pyrimidine in order to directly assess the contribution to transition state destabilization of the unpaired amino group presented in the minor groove by G4:isoC69. However, this base analogue proved insufficiently stable to numerous protection/deprotection schemes rendering this approach impractical.

(19) Westhof, E.; Dumas, P.; Moras, D. *Biochimie* **1988**, *70*, 145−165.
(20) Holbrook, S. R.; Cheong, C.; Tinoco, I., Jr.; Kim, S. H. *Nature* **1991**, *353*, 579−581.
(21) Betzel, C.; Lorenz, S.; Fürste, J. P.; Bald, R.; Zhang, M.; Schneider, T. R.; Wilson, K. S.; Erdmann, V. A. *FEBS Lett.* **1994**, *351*, 159−164.
(22) Cruse, W. B.; Saludjian, P.; Biala, E.; Strazewski, P.; Prange, T.; Kennard, O. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 4160−4164.
(23) Biswas, R.; Wahl, M. C.; Ban, C.; Sundaralingam, M. *J. Mol. Biol.* **1997**, *267*, 1149−1156.
(24) Müller, U.; Schübel, H.; Sprinzl, M.; Heinemann, U., personal communication.
(25) Otwinowski, Z.; Schevitz, R. W.; Zhang, R. G.; Lawson, C. L.; Joachimiak, A.; Marmorstein, R. Q.; Luisi, B. F.; Sigler, P. B. *Nature* **1988**, *335*, 321−329.
(26) Cheng, X.; Balendiran, K.; Schildkraut, I.; Anderson, J. E. *EMBO J.* **1994**, *13*, 3927−3935.
(27) Oubridge, C.; Ito, N.; Evans, P. R.; Teo, C. H.; Nagai, K. *Nature* **1994**, *372*, 432−438.
(28) Shakked, Z.; Guzikevich-Guerstein, G.; Frolow, F.; Rabinovich, D.; Joachimiak, A.; Sigler, P. B. *Nature* **1994**, *368*, 469−473.
(29) Hirsch, J. A.; Aggarwal, A. K. *EMBO J.* **1995**, *14*, 6280−6291.
(30) Keller, W.; Konig, P.; Richmond, T. J. *J. Mol. Biol.* **1995**, *254*, 657−667.
(31) Elrod-Erickson, M.; Rould, M. A.; Nekludova, L.; Pabo, C. O. *Structure* **1996**, *4*, 1171−1180.
(32) Passner, J. M.; Steitz, T. A. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 2843−2847.
(33) Swaminathan, K.; Flynn, P.; Reece, R. J.; Marmorstein, R. *Nat. Struct. Biol.* **1997**, *4*, 751−759.
(34) Tucker-Kellogg, L.; Rould, M. A.; Chambers, K. A.; Ades, S. E.; Sauer, R. T.; Pabo, C. O. *Structure* **1997**, *5*, 1047−1054.
(35) Valegard, K.; Murray, J. B.; Stonehouse, N. J.; van den Worm, S.; Stockley, P. G.; Liljas, L. *J. Mol. Biol.* **1997**, *270*, 724−738.