

# Assembly of a catalytic unit for RNA microhelix aminoacylation using nonspecific RNA binding domains

Joseph W. Chihade\* and Paul Schimmel†

Skaggs Institute for Chemical Biology, Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037

Contributed by Paul Schimmel, September 8, 1999

**An assembly of a catalytic unit for aminoacylation of an RNA microhelix is demonstrated here. This assembly may recapitulate a step in the historical development of tRNA synthetases. The class-defining domain of a tRNA synthetase is closely related to the primordial enzyme that catalyzed synthesis of aminoacyl adenylate. RNA binding elements are imagined to have been added so that early RNA substrates could be docked proximal to the activated amino acid. RNA microhelices that recapitulate the acceptor stem of modern tRNAs are potential examples of early substrates. In this work, we examined a fragment of *Escherichia coli* alanyl-tRNA synthetase, which catalyzes aminoacyl adenylate formation but is virtually inactive for catalysis of RNA microhelix aminoacylation. Fusion to the fragment of either of two unrelated nonspecific RNA binding domains activated microhelix aminoacylation. Although the fusion proteins lacked the RNA sequence specificity of the natural enzyme, their activity was within 1–2 kcal·mol<sup>-1</sup> of a truncated alanyl-tRNA synthetase that has aminoacylation activity sufficient to sustain cell growth. These results show that, starting with an activity for adenylate synthesis, barriers are relatively low for building catalytic units for aminoacylation of RNA helices.**

**W**e report here an example of a pathway by which one component of the genetic code could be assembled in evolution. Aminoacyl-tRNA synthetases establish the genetic code by attaching amino acids to transfer RNAs that bear the corresponding anticodons defined by the code (1–3). Because they establish the code, these enzymes are believed to be amongst the earliest proteins (4). The enzymes activate amino acids by coupling them to ATP to form an aminoacyl adenylates and subsequently transfer the amino acid to the 3' acceptor end of tRNA (5, 6). Extant aminoacyl-tRNA synthetases are divided into two classes based on the structures of their active sites. Class I enzymes are characterized by a Rossmann nucleotide binding fold having an 11-aa signature sequence that ends in the tetrapeptide HIGH and a KMSKS pentapeptide (7–9). Class II enzymes have a distinct fold made up of an antiparallel  $\beta$ -sheet and three  $\alpha$ -helices that encompass three highly degenerate sequence motifs known as motifs 1, 2, and 3 (5, 10). Members of the two classes of synthetases are thought to have evolved independently from two distinct primordial enzymes that catalyzed the formation of aminoacyl adenylate. Specificity for distinct transfer RNAs and amino acids evolved through changes within the ancestral core and through the incorporation of additional domains that are idiosyncratic to particular synthetases (11, 12).

Transfer RNAs also may have evolved in a piecemeal fashion. Modern tRNAs form an L-shaped structure made up of two domains. One domain is a 12-bp minihelix that is made up of the T $\Psi$ C and acceptor stems and that terminates at the 3'-end in the single-stranded tetranucleotide N<sup>73</sup>CCA with the amino acid acceptance site. (N<sup>73</sup> is any of the four nucleotides and is referred to as the “discriminator base.”) The second domain contains the anticodon loop at the end of a 10-bp helix comprised of the dihydrouridine and anticodon stems. The two domains segregate

distinct functions of the tRNA. Thus, the anticodon-containing domain serves as a reading head that interacts with messenger RNAs whereas the acceptor-T $\Psi$ C domain contains the site for amino acid attachment and interacts with the peptidyl transfer center in the ribosome.

Several lines of evidence suggested that the two domains of tRNAs arose independently. For example, each domain of tRNA interacts with a separate rRNA on the ribosome (13). In addition, the minihelix is an ancient tag for RNA genomes (14, 15) and is the part of contemporary tRNAs that is used as a primer for viral reverse transcription (16–18). Finally, at least 10 tRNA synthetases isolated across evolution from bacterial to human cells can aminoacylate helical RNA substrates corresponding to the acceptor-T $\Psi$ C domain or to the acceptor stem alone (termed mini- and microhelices, respectively) (19–27). Aminoacylation of these truncated substrates generally depends on the identity of the N73 discriminator base and on one or two base pairs within the last four pairs of the acceptor stem. The relationship between acceptor stem nucleotides/structures and specific amino acids constitutes an “operational RNA code” for amino acids. This code may have predated the addition of the anticodon branch to the tRNA structure (28).

In one evolutionary scheme for tRNA synthetases, ancient “core enzymes” activated amino acids to form aminoacyl adenylates. The active sites are imagined to be the progenitors of either of the two structural formats that determine the two classes of modern tRNA synthetases. The reaction to form (from an adenylate) an aminoacyl ester linkage with the 3'-hydroxyl of an RNA is thermodynamically favorable. Thus, the challenge in evolution may have been to dock potential RNA substrates proximal to the bound aminoacyl adenylates. These early substrates could have been RNA minihelix-like because such structures are thought to be prevalent in the RNA world (14, 15, 29). When these early RNA substrates incorporated an anticodon-containing branch, the primitive synthetase then could incorporate additional RNA binding elements to make contacts with the second domain of the tRNA.

The dissection by deletions of a modern tRNA synthetase can be viewed as a form of retroevolution. For example, deletion of the C-terminal half of *Escherichia coli* alanyl-tRNA synthetase (AlaRS) yields a 461-aa N-terminal fragment (461N) that aminoacylates microhelix substrates with the same efficiency as the full length enzyme (30). Because the efficiency of aminoacylation of full-length tRNA substrates by fragment 461N is reduced  $\approx$ 100-fold when compared with that of the full-length enzyme, this deletion is believed to remove parts of the enzyme that make

Abbreviations: AlaRS, alanyl-tRNA synthetase; 461N, 461-aa N-terminal fragment.

\*Present address: Department of Chemistry, Bronfman Science Center, Williams College, Williamstown, MA 01267.

†To whom reprint requests should be addressed. E-mail: schimmel@scripps.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.

contact with the second (anticodon-containing) domain of the L-shaped tRNA. (Contacts of AlaRS with the second domain of the tRNA do not include the anticodon triplet, which is unprotected from nuclease attack by the bound enzyme (31).) However, when expressed from a multicopy plasmid, the activity of fragment 461N is sufficient to maintain cell growth (32, 33). Thus, fragment 461N might be seen as an intermediate in the evolutionary development of AlaRS: that is, as a protein that efficiently aminoacylates the acceptor stem domain in isolation.

Further truncation of *E. coli* AlaRS to fragment 368N yields an enzyme that is fully competent for formation of alanyl adenylate but is inactive under standard conditions for aminoacylation of RNA substrates (32, 34). The fragment contains the three sequence motifs that are characteristic of the active sites of class II enzymes. Fragment 368N lacks RNA binding activity, with a  $K_d$  for tRNA at pH 5.2 that is reduced by more than three orders of magnitude compared with that of fragment 461N (35). Therefore, the deletion of the region from Thr369-Asp461 creates an enzyme that may resemble the functional progenitor in the evolutionary scheme described above: that is, one that catalyzes formation of aminoacyl adenylate but cannot effectively bind or aminoacylate RNA.

We imagined that a step in the evolution of aminoacylation activity might be the fusion of a nonspecific RNA binding domain to the catalytic unit for adenylate synthesis. The presence of nonspecific RNA binding domains in modern aminoacyl-tRNA synthetases such as *E. coli* isoleucyl-tRNA synthetase (36) and *Saccharomyces cerevisiae* glutamyl-tRNA synthetase (37) is consistent with these domains playing a role in the development of tRNA synthetases. To test the feasibility of conferring aminoacylation activity on fragment 368N of AlaRS, we fused two different nonspecific RNA binding domains to the core fragment for adenylate synthesis. It was not obvious that a fusion of this sort would result in aminoacylation activity without further refinements by protein engineering.

## Materials and Methods

**Construction and Purification of Truncated and Fusion Proteins.** All plasmid constructs are based on pQE-AlaS (38), which encodes the *E. coli* alanyl-tRNA synthetase with a C-terminal His<sub>6</sub> tag sequence under the control of an inducible T5 promoter. PCR mutagenesis (39) was used to produce silent mutations that introduce *Stu*I and *Sac*I restriction sites after the codons for Leu374 and Leu465, respectively, yielding the plasmid pAlaS-ESS. To construct a plasmid encoding fragment 465N with a C-terminal His<sub>6</sub> tag, an oligonucleotide cassette encoding six His residues and a stop codon was inserted between the *Sac*I and *Spe*I sites of pAlaS-ESS to give plasmid pESS-465. The N-terminal 227-residue appended domain of *S. cerevisiae* glutamyl-tRNA synthetase was amplified by PCR from plasmid pEFW110, which encodes the complete sequence of yeast glutamyl tRNA synthetase (40). For this purpose, the primers GCGCGCAGGCTGATGTCTTCTGTAGAAGAATTGACTCAG and GCGCGCAGCTCCAAATCACCTAGGAAACCTTCATTGAACATAGTCC (restriction sites in italics) introduced *Stu*I and *Sac*I sites, respectively, at the 5'- and 3'-ends of the gene for glutamyl tRNA synthetase. Subcloning into pESS-465 gave plasmid pESS-AQ, which encodes the first 374 residues of *E. coli* AlaRS fused to the appended domain of *S. cerevisiae* glutamyl tRNA synthetase, with the sequence ELHHHHHH at the C terminus of the fusion. The resulting protein was designated as fusion 374N-Ad<sup>ScQRS</sup>. Fusion of the C-terminal domain of Arc1p was achieved analogously by PCR from pCW119 [which encodes the complete Arc1p protein (37)] and ligation of the coding sequence of the C-terminal domain into pESS-465 to give plasmid pESS-AA. The protein expressed from this plasmid was designated as fusion 374N-C<sup>Arc1p</sup>.

Fusion proteins and His<sub>6</sub>-tagged fragment 465N were purified

by nickel affinity chromatography as described (38). Cells were grown in LB medium (supplemented with 50 mg/ml ampicillin) to OD<sub>600</sub> of 0.8 and then were inoculated with isopropyl β-D-thiogalactoside to a final concentration of 1.5 mM. The cells were grown an additional 4 hours and then were harvested and lysed in a French press. The His<sub>6</sub>-tagged enzyme was purified on a Ni-nitrilotriacetic acid agarose (Qiagen, Santa Clarita, CA) column in 50 mM Na<sub>2</sub>HPO<sub>4</sub> (pH 8.0), 300 mM NaCl, and 10% glycerol. Elution was with a gradient of 20–250 mM imidazole. Fragment 368N was prepared as described (41) by trypsin digestion and was purified by gel filtration chromatography on Superose-12.

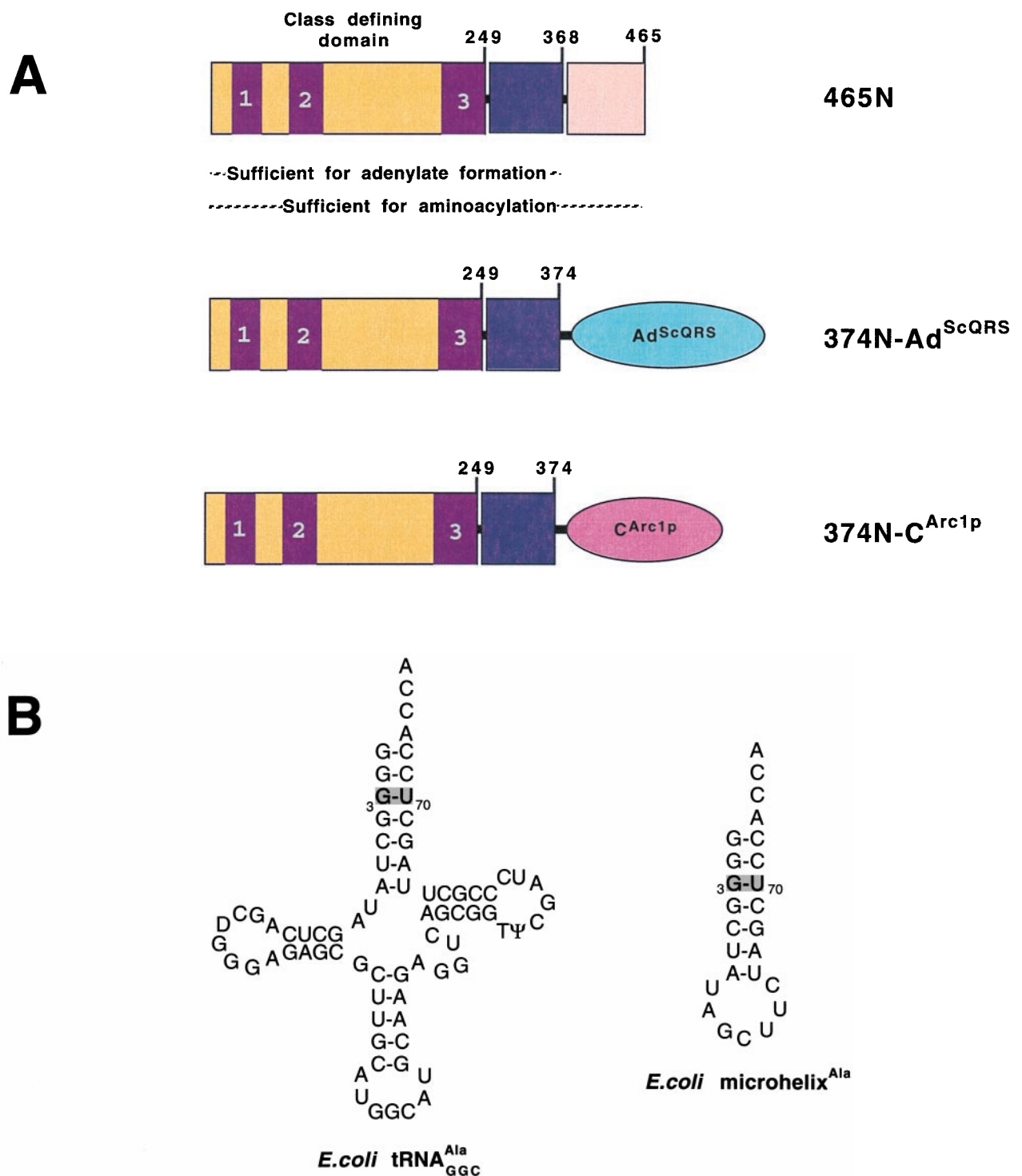
**RNA Substrates and Aminoacylation Assays.** Microhelix RNA substrates were chemically synthesized on a Pharmacia Gene Assembler synthesizer by using published procedures (42). Products of the syntheses were purified on denaturing 20% polyacrylamide gels.

Aminoacylation assays were carried out at ambient temperature (25°C) in 50 mM Hepes (pH 7.5), 20 mM alanine, 4 mM ATP, 20 mM KCl, 10 mM MgCl<sub>2</sub>, 20 mM β-mercaptoethanol, and 0.1 mg/ml BSA, as described (43). To insure proper base pairing, microhelices were heated to 80°C and were cooled in the absence of magnesium before addition to the aminoacylation reaction mixture. In assays in which different constructions were being directly compared, RNA was at a concentration of 2.5 or 10 μM whereas enzyme concentrations were 0.5 or 1 μM. [Enzyme concentrations were determined by the alanine-dependent ATP-pyrophosphate exchange activity (44).] Kinetic parameters were determined by direct fitting of data points to the Michaelis-Menten equation. Initial rates were determined at ambient temperature (25°C), with RNA concentrations from 10 to 400 μM (465N), 5 to 400 μM (374N-C<sup>Arc1p</sup>), or 2 to 20 μM (374N-Ad<sup>ScQRS</sup>). Enzyme concentrations of 0.5 or 1 μM were used.

## Results

**Nonspecific RNA Binding Domains.** Two well-studied nonspecific RNA binding domains were chosen as fusion partners for this work. The 228 amino acid N-terminal appended domain of *S. cerevisiae* glutamyl-tRNA synthetase binds RNA nonspecifically, with a  $K_D$  for tRNA, pseudoknots, and poly-A RNA of 0.6–1.8 μM (37). Arc1p specifically binds methionyl-tRNA synthetase and glutamyl-tRNA synthetase in the yeast *S. cerevisiae*. It is thought to deliver tRNA substrates to these enzymes by enhancing the local effective concentration of the substrates around the proteins (45, 46). On the basis of secondary structure predictions, the protein has been divided into three domains. The N-terminal 132 residues are thought to interact with the aminoacyl-tRNA synthetases whereas the middle (133–201) and C-terminal domains (202–376) are responsible for RNA binding. A truncation composed of the only the middle and C-terminal domains nonspecifically binds tRNA with the same affinity as does the full-length protein, with a  $K_D$  of 5–10 nM. The C-terminal domain alone also binds tRNA, but with a >50-fold reduction in affinity (46).

C-terminal His<sub>6</sub>-tagged fusions were constructed by attaching the appended domain of yeast GlnRS (Ad<sup>ScQRS</sup>) or the C-terminal domain of Arc1p (C<sup>Arc1p</sup>) to the C terminus of fragment 374N of *E. coli* AlaRS (Fig. 1). (We chose fragment 374N instead of 368N for reasons of technical convenience in constructing the plasmids that encode the fusion proteins. Previous work showed that fragments 368N and 385N have the full adenylate synthesis activity of the native enzyme (35). For that reason we assumed that a fusion protein with the first 374 amino acids of AlaRS would have full activity for adenylate synthesis (see below).) Attempts to express fusions to full length Arc1p and the middle and C-terminal domains of Arc1p yielded no soluble protein and



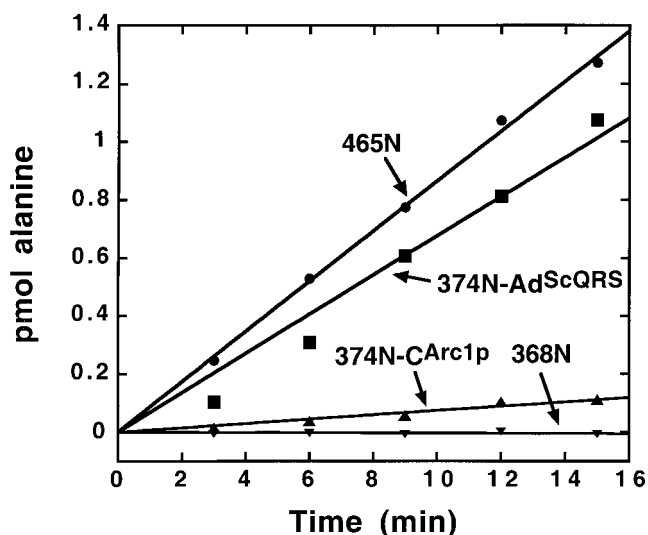
**Fig. 1.** (A) Schematic diagram of fragment 465N of *E. coli* alanyl-tRNA synthetase and the fusions to nonspecific RNA binding domains C<sup>Arc1p</sup> and Ad<sup>ScQRS</sup>. The class-defining active site region extends to His249 whereas a fragment produced by extensive proteolysis terminates at Arg368. The nonspecific RNA binding domains are schematically shown as ellipsoids. Conserved class II motifs in the class defining domain are labeled as 1, 2, and 3. (B) Sequence and cloverleaf structure of *E. coli* tRNA<sup>Ala</sup> and sequence and hairpin structure of microhelix<sup>Ala</sup>. The G3:U70 base pair is indicated by a shaded box.

were not pursued further. The 374N-Ad<sup>ScQRS</sup> and 374N-C<sup>Arc1p</sup> fusion proteins were isolated by nickel affinity chromatography to a purity of >90%.

Concentrations of fusion proteins 374N-Ad<sup>ScQRS</sup> and 374N-C<sup>Arc1p</sup> were determined by adenylate synthesis activity and by standard Bradford assays. Based on the agreement between

these assays, the fusions had full adenylate synthesis activity, indicating that alanine activation is not affected by the addition of the heterologous RNA binding domains or of the His<sub>6</sub>-tag.

**Aminoacylation of RNA Microhelices with Fusion Proteins.** For the aminoacylation assays, a hairpin microhelix comprised of the



**Fig. 2.** Aminoacylation of microhelix<sup>Ala</sup> by fusion constructs. Aminoacylation assays were at pH 7.5, 25°C, with 2.5 μM RNA and 500 nM enzyme.

seven base pairs of the acceptor stem of *E. coli* tRNA<sup>Ala</sup> was used as a substrate (Fig. 2). This substrate is efficiently aminoacylated by fragment 461N. Aminoacylation is highly specific with respect to the sequence of the microhelix. In particular, it depends on a G3:U70 base pair that marks the molecule for aminoacylation with alanine (47, 48). Our experiments were directed at determining whether the fusion constructions would charge the microhelix, and, if so, then we wanted to compare the specificity of charging with that of fragment 465N of AlaRS. The sensitivity of charging to changes in the essential G3:U70 base pair provided a simple way to test for specificity.

In these experiments, fragment 465N of AlaRS with a 6-His tag appended to the C terminus was used as a standard for sequence-specific aminoacylation of the microhelix<sup>Ala</sup>. Fragment 368N was used as a negative control. [Although fragment 374N could have been used in principle, the purification of this protein could have easily resulted in fragment 368N as a product of proteolysis by cellular proteases (34, 49).] Fragment 465N charged microhelix<sup>Ala</sup> (Fig. 2). As expected, no charging under these conditions was seen with Fragment 368N.

In contrast, both the 374N-Ad<sup>ScQRS</sup> and 374N-C<sup>Arc1p</sup> fusion proteins charged microhelix<sup>Ala</sup> at rates clearly above the background of Fragment 368N (Fig. 2). Remarkably, the activity seen with fusion 374N-Ad<sup>ScQRS</sup> was comparable to that of fragment 465N. As for the 374N-C<sup>Arc1p</sup> fusion construction, the efficiency of aminoacylation was ≈15-fold lower than that with fragment 465N.

Kinetic parameters were measured for each of the proteins over a range of at least 10-fold in the concentration of microhelix<sup>Ala</sup>. This analysis showed that both fusion proteins have  $k_{cat}$  values ≈20-fold lower than that for fragment 465N. [The value of  $k_{cat}$  for the His-tagged fragment 465N is ≈9-fold lower than that previously measured for aminoacylation of microhelix<sup>Ala</sup> by full length AlaRS (50) at 37°C. The value of  $K_m$  (100 μM) for fragment 465N is similar to that previously reported for full length AlaRS. Because fragment 461N aminoacylates truncated RNA substrates with the same efficiency as the full length enzyme (51), this lowering of  $k_{cat}$  could be attributable to the difference in temperature (25°C versus 37°C) and, additionally or alternatively, to the presence of the His-tag appended to fragment 465N. It is possible that the tag also attenuates  $k_{cat}$  for the fusion proteins, but this possibility has not been proven.] However, fusion 374N-Ad<sup>ScQRS</sup> has a  $K_m$  of 7 μM for microhe-

lix<sup>Ala</sup>. This  $K_m$  is ≈15-fold lower than that of fragment 465N. Consequently, the combined effect of the reduced  $k_{cat}$  and  $K_m$  for fusion 374N-Ad<sup>ScQRS</sup> results in  $k_{cat}/K_m$  being comparable to that of fragment 465N.

In the case of fusion 374N-C<sup>Arc1p</sup>, the entire difference between it and fusion 374N-Ad<sup>ScQRS</sup> is explained by the difference in  $K_m$  for microhelix<sup>Ala</sup>. This difference is consistent with the previously measured apparent 10-fold higher tRNA binding affinity of Ad<sup>ScQRS</sup> when compared with a large fragment of Arc1p that includes the C-terminal domain (37).

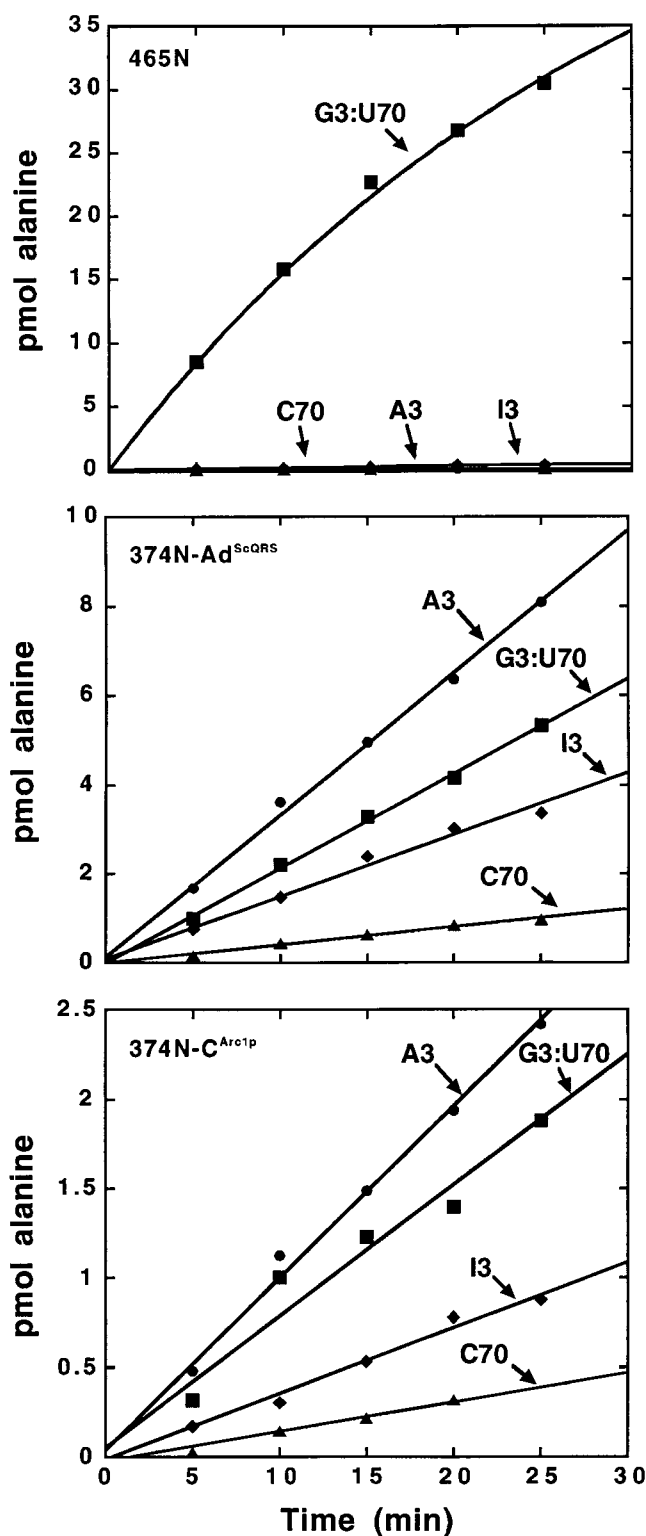
**Specificity of Microhelix<sup>Ala</sup> Aminoacylation by Fusion Proteins.** As stated above, alanyl-tRNA synthetases require a G3:U70 base pair for aminoacylation of tRNA<sup>Ala</sup> and of RNA substrates based on the acceptor stem of tRNA<sup>Ala</sup>. The unpaired exocyclic 2-amino group of G3 in the minor groove is a critical atomic determinant, as demonstrated by replacement of G3 by inosine (I) in duplex substrates that recapitulate the acceptor stem of tRNA<sup>Ala</sup>. The only difference between the G:U and I:U pair is the presence or absence of the 2-amino group. *E. coli* AlaRS fails to charge substrates with an I3:U70 pair (52). Moreover, inserting a base pair comprised of 2-amino adenosine (2-AA) and isocytidine (isoC) at the 3:70 position can restore aminoacylation (53, 54). With the 2-AA3:isoC70 pair, major and minor groove atomic determinants are different from that of the G:U pair, but the presence of the unpaired exocyclic amino group is maintained. Microhelices incorporating the 2-AA3:U70 base pair are efficiently charged with alanine, further demonstrating the importance of the unpaired 2-amino group at the 3-position of the helix.

Thus, to test the RNA sequence specificity of the fusion proteins, the comparison of a microhelix containing the G3:U70 pair with one containing I3:U70 was considered essential. We confirmed that fragment 465N charged microhelix<sup>Ala</sup> containing the G3:U70 base pair and that substitution with I3:U70 eliminated aminoacylation (Fig. 3 *Top*). In contrast, both fusion proteins charged microhelices containing the I3:U70 base pair, albeit with a relatively small decrease in efficiency (Fig. 3 *Middle* and *Bottom*). Thus, the fusion proteins do not recapitulate the critical specificity for the unpaired 2-amino group at the 3-position of the RNA helix. This result is consistent with previous work that demonstrated the importance of the region from Thr369 to Asp461 for recognition of G3:U70 (41).

We also checked substrates containing an A3:U70 and a G3:C70 base pair. As expected, neither of these substrates was charged by fragment 465N. However, both variants were aminoacylated by the fusion constructions. Of the four substrates (G3:U70, I3:U70, A3:U70, and G3:C70 microhelix<sup>Ala</sup>), the A3:U70 microhelix<sup>Ala</sup> was the most efficiently and the G3:C70 microhelix<sup>Ala</sup> the least efficiently charged by both fusion proteins (Fig. 3). These results demonstrate that, although the fusion constructions are active for aminoacylation and show some sensitivity to the base pair at the 3:70 position, they lack the high selectivity of AlaRS for G3:U70.

## Discussion

The portion of AlaRS that extends to His249 contains the structural unit that defines the class II tRNA synthetases. This unit is thought to be the original, historic tRNA synthetase that contained the structural components needed to catalyze synthesis of the aminoacyl adenylate. RNA binding elements are imagined to have been added to this structural unit to give rise to the 10 distinct class II enzymes. These RNA binding elements are more or less specific to the synthetase; that is, enzymes in the same class often do not share common RNA binding elements. Thus, these elements may have been added in an idiosyncratic way. Any peptide or protein that docked the end of an RNA close to the bound aminoacyl adenylate should facilitate synthesis of



**Fig. 3.** Specificity of aminoacylation for a G3:U70 base pair. Aminoacylation of microhelix<sup>Ala</sup> containing the indicated nucleotides at the 3:70 position with fragment 465N (Top), fusion 374N-Ad<sup>ScQRS</sup> (Middle), and fusion 374N-C<sup>Arc1p</sup> (Bottom). Aminoacylation assays were at pH 7.5, 25°C, with 10  $\mu$ M RNA and 1  $\mu$ M enzyme.

the aminoacyl RNA. From this perspective, the key to success lies in the orientation in which the RNA substrate is docked onto the moiety that holds the aminoacyl adenylate.

**Table 1.** Kinetic parameters for aminoacylation of microhelix<sup>Ala</sup> at pH 7.5, 25°C

Enzyme	$k_{cat}$ , min <sup>-1</sup>	$K_M$ , $\mu$ M	$k_{cat}/K_M$ (relative)
465N	1.9	100	(1.0)
374N-Ad <sup>ScQRS</sup>	0.09	7	0.68
374N-C <sup>Arc1p</sup>	0.11	110	0.05

Kinetic parameters were determined over a microhelix<sup>Ala</sup> range of 10–400  $\mu$ M (465N), 5–400  $\mu$ M (374N-C<sup>Arc1p</sup>), or 2–20  $\mu$ M (374N-Ad<sup>ScQRS</sup>).

In the present study, the success of the experiment was undoubtedly influenced by using a fragment of AlaRS that extended beyond His249 (to Leu374). This extension was necessary because C-terminal truncations of AlaRS significantly beyond Arg368 result in a loss of the capacity to synthesize alanyl adenylate. For example, fragment 257N is inactive for adenylate synthesis (32). By extending the fragment of AlaRS to Leu374, we incorporated elements that are important to recognize the discriminator base A73 and the first two base pairs of microhelix<sup>Ala</sup> (41). In particular, Asp285 and Arg314 are especially important for binding microhelix<sup>Ala</sup> to AlaRS (55). These highly conserved residues are not involved in adenylate synthesis but are part of a predicted two-helix pair that establishes a platform for binding RNA helices. However, as demonstrated by the lack of aminoacylation activity for fragment 368N (Fig. 2), this two-helix pair does not provide sufficient binding interactions to achieve productive docking of microhelix<sup>Ala</sup>. Thus, the nonspecific RNA binding domains added here provided reinforcement for weak interactions—including those involving the platform around Asp285 and Arg314—that place the microhelix substrate in the proper orientation for aminoacylation.

The lack of specificity of aminoacylation with the fusion constructs provides further support for the earlier conclusion that determinants for recognition of G3:U70 lie between Arg368 and Asp461 (41). These fusion constructions have  $k_{cat}$  values that are  $\approx$ 20-fold below that of fragment 465N. This difference corresponds to  $<2$  kcal·mol<sup>-1</sup> in transition state stabilization free energy. Possibly the bulk of this difference is an entropic component that arises from unproductive orientations of the 3'-end of the substrate when it is docked with the assistance of the nonspecific RNA binding domains from *S. cerevisiae* glutamyl-tRNA synthetase and Arc1p. With specificity for G3:U70 (i.e., specific recognition of the 2-amino group of G3), the orientation of the 3'-end of the bound substrate undoubtedly is made precise. Thus, the energetic benefit of specific contacts with G3:U70 may be used to offset the entropic cost of fixing precisely the orientation of the 3'-end of the microhelix substrate.

In this connection, it is of interest that the  $K_M$  parameters for the fusion proteins are equal to or less than that for fragment 465N (Table 1). This circumstance suggests that evolving (within the context of either nonspecific RNA binding domain) an energetically favorable specific contact with G3:U70 could offset the entropic cost of raising  $k_{cat}$  without having to make a sacrifice in  $K_M$ .

Aminoacyl-tRNA synthetases must perform several distinct functions, which include amino acid activation, RNA binding, RNA discrimination and, often, editing of incorrectly aminoacylated tRNAs (32, 56, 57). In several cases, these distinct functions are carried out by specific domains, which can function independently of the larger enzyme. “Mixing and matching” domains from distinct proteins affords the possibility of creating “artificial” synthetases, which could be engineered to have different RNA or amino acid specificities than those of the wild-type enzymes. By fusing the catalytic site of a synthetase to

unrelated RNA binding domains, we have taken a first step in this direction.

In our experiments, we made a simple head-to-tail linear fusion of N terminus of C<sup>Arclp</sup> or Ad<sup>ScQRS</sup> to the C terminus of fragment 465N of AlaRS. We suspect that placing the nonspecific RNA binding domain on the C-terminal side of fragment 465N may have been important for the success of the experiment. In an earlier study, we showed that fragment 461N of AlaRS could be split after His249—the end of the class-defining catalytic domain—to give fragments 249N and 212C. These fragments spontaneously assembled *in vivo* to give a protein that executed G3:U70-specific charging of microhelix<sup>Ala</sup> (58). These experiments supported the concept that, in principle, early

tRNA synthetases could have developed as noncovalent assemblies of fragments that synthesize aminoacyl adenylates with those that bind RNA. These and earlier experiments (summarized in refs. 32 and 59) also demonstrated explicitly that the domain for adenylate synthesis and RNA binding elements were arranged in a linear “head-to-tail” way along the sequence for AlaRS. Thus, our fusion proteins duplicated the strategy used by the natural enzyme.

We thank Drs. Lluís Ribas de Pouplana, Shana Kelley, Brian Steer, and Tamara Hendrickson for many helpful discussions. This work was supported by National Institutes of Health Grant GM23562. J.W.C. was a National Institutes of Health postdoctoral fellow (1997–1999).

1. Schimmel, P. (1987) *Annu. Rev. Biochem.* **56**, 125–158.
2. Meinnel, T., Mechulam, Y. & Blanquet, S. (1995) in *tRNA: Structure, Biosynthesis, and Function*, eds. Söll, D. & RajBhandary, U. L. (Am. Soc. Microbiol., Washington, DC), pp. 251–292.
3. Arnez, J. G. & Moras, D. (1997) *Trends Biochem. Sci.* **22**, 211–216.
4. Nagel, G. M. & Doolittle, R. F. (1995) *J. Mol. Evol.* **40**, 487–498.
5. Eriani, G., Delarue, M., Poch, O., Gangloff, J. & Moras, D. (1990) *Nature (London)* **347**, 203–206.
6. Moras, D. (1992) *Trends Biochem. Sci.* **17**, 159–164.
7. Webster, T. A., Tsai, H., Kula, M., Mackie, G. A. & Schimmel, P. (1984) *Science* **226**, 1315–1317.
8. Hountondji, C., Dessen, P. & Blanquet, S. (1986) *Biochimie* **68**, 1071–1078.
9. Ludmerer, S. W. & Schimmel, P. (1987) *J. Biol. Chem.* **262**, 10801–10806.
10. Cusack, S., Härtlein, M. & Leberman, R. (1991) *Nucleic Acids Res.* **19**, 3489–3498.
11. Delarue, M. & Moras, D. (1993) *BioEssays* **15**, 675–687.
12. Schimmel, P. & Ribas de Pouplana, L. (1995) *Cell* **81**, 983–986.
13. Noller, H. F. (1993) in *The RNA World*, eds. Gesteland, R. F. & Atkins, J. F. (Cold Spring Harbor Lab. Press, Plainview, NY), pp. 137–156.
14. Weiner, A. M. & Maizels, N. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 7383–7387.
15. Maizels, N. & Weiner, A. M. (1993) in *The RNA World*, eds. Gesteland, R. F. & Atkins, J. F. (Cold Spring Harbor Lab. Press, Plainview, NY), pp. 577–602.
16. Kikuchi, Y., Ando, Y. & Shiba, T. (1986) *Nature (London)* **323**, 824–826.
17. Chapman, K. B., Bystrom, A. S. & Boeke, J. D. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 3236–3240.
18. Isel, C., Ehresmann, C., Keith, G., Ehresmann, B. & Marquet, R. (1995) *J. Mol. Biol.* **247**, 236–250.
19. Francklyn, C. & Schimmel, P. (1989) *Nature (London)* **337**, 478–481.
20. Francklyn, C., Musier-Forsyth, K. & Schimmel, P. (1992) *Eur. J. Biochem.* **206**, 315–321.
21. Martinis, S. A. & Schimmel, P. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 65–69.
22. Frugier, M., Florentz, C. & Giegé, R. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 3990–3994.
23. Nureki, O., Niimi, T., Muto, Y., Kanno, H., Kohno, T., Muramatsu, T., Kawai, G., Miyazawa, T., Giegé, R., Florentz, C. & Yokoyama, S. (1993) in *The Translation Apparatus*, eds. Nierhaus, K. H., Franceschi, F., Subramanian, A. R., Erdmann, V. A. & Wittmann-Liebold, B. (Plenum, New York), pp. 59–66.
24. Frugier, M., Florentz, C. & Giegé, R. (1994) *EMBO J.* **13**, 2218–2226.
25. Hamann, C. S. & Hou, Y.-M. (1995) *Biochemistry* **34**, 6527–6532.
26. Quinn, C. L., Tao, N. & Schimmel, P. (1995) *Biochemistry* **34**, 12489–12495.
27. Saks, M. E. & Sampson, J. R. (1996) *EMBO J.* **15**, 2843–2849.
28. Schimmel, P., Giegé, R., Moras, D. & Yokoyama, S. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 8763–8768.
29. Wang, H. & Lambowitz, A. M. (1993) *Cell* **75**, 1071–1081.
30. Buechter, D. D. & Schimmel, P. (1993) *Biochemistry* **32**, 5267–5272.
31. Park, S. J. & Schimmel, P. (1988) *J. Biol. Chem.* **263**, 16527–16530.
32. Jasin, M., Regan, L. & Schimmel, P. (1983) *Nature (London)* **306**, 441–447.
33. Ho, C., Jasin, M. & Schimmel, P. (1985) *Science* **229**, 389–393.
34. Putney, S. D., Sauer, R. T. & Schimmel, P. R. (1981) *J. Biol. Chem.* **256**, 198–204.
35. Regan, L., Bowie, J. & Schimmel, P. (1987) *Science* **235**, 1651–1653.
36. Glasfeld, E. & Schimmel, P. (1997) *Biochemistry* **36**, 6739–6744.
37. Wang, C. C. & Schimmel, P. (1999) *J. Biol. Chem.* **274**, 16508–16512.
38. Ribas de Pouplana, L. & Schimmel, P. (1997) *Biochemistry* **36**, 15041–15048.
39. Higuchi, R., Krummel, B. & Saiki, R. K. (1988) *Nucleic Acids Res.* **16**, 7351–7367.
40. Whelihan, E. F. & Schimmel, P. (1997) *EMBO J.* **16**, 2968–2974.
41. Buechter, D. D. & Schimmel, P. (1995) *Biochemistry* **34**, 6014–6019.
42. Francklyn, C., Musier-Forsyth, K. & Schimmel, P. (1992) *Eur. J. Biochem.* **206**, 315–321.
43. Calendar, R. & Berg, P. (1966) *Biochemistry* **5**, 1690–1695.
44. Calendar, R. & Berg, P. (1966) in *Procedures in Nucleic Acids Research*, eds. Cantoni, R. & Davies, D. R. (Harper & Row, New York), pp. 384–399.
45. Simos, G., Segref, A., Fasiolo, F., Hellmuth, K., Shevchenko, A., Mann, M. & Hurt, E. C. (1996) *EMBO J.* **15**, 5437–5448.
46. Simos, G., Sauer, A., Fasiolo, F. & Hurt, E. C. (1998) *Mol. Cell* **1**, 235–242.
47. Hou, Y. M. & Schimmel, P. (1988) *Nature (London)* **133**, 140–145.
48. McClain, W. H. & Foss, K. (1988) *Science* **240**, 793–796.
49. Putney, S. D., Royal, N. J., Neuman de Vegvar, H., Herlihy, W. C., Biemann, K. & Schimmel, P. (1981) *Science* **213**, 1497–1500.
50. Francklyn, C. & Schimmel, P. (1989) *Nature (London)* **337**, 478–481.
51. Buechter, D. D. & Schimmel, P. (1993) *Biochemistry* **32**, 5267–5272.
52. Musier-Forsyth, K., Usman, N., Scaringe, S., Doudna, J., Green, R. & Schimmel, P. (1992) *Science* **253**, 784–786.
53. Musier-Forsyth, K., Shi, J.-P., Henderson, B., Bald, R., Fürste, J. P., Erdmann, V. A. & Schimmel, P. (1995) *J. Am. Chem. Soc.* **117**, 7253–7254.
54. Beuning, P. J., Yang, F., Schimmel, P. & Musier-Forsyth, K. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 10150–10154.
55. Ribas de Pouplana, L., Buechter, D., Sardesai, N. Y. & Schimmel, P. (1998) *EMBO J.* **17**, 5449–5457.
56. Wayne, M. M. Y., Winter, G., Wilkinson, A. J. & Fersht, A. R. (1983) *EMBO J.* **2**, 1827–1829.
57. Lin, L., Hale, S. P. & Schimmel, P. (1996) *Nature (London)* **384**, 33–34.
58. Sardesai, N. & Schimmel, P. (1998) *J. Am. Chem. Soc.* **120**, 3269–3270.
59. Schimmel, P. & Ripmaster, T. (1995) *Trends Biochem. Sci.* **20**, 333–334.