Polypeptide Sequences Essential for RNA Recognition by an Enzyme

LYNNE REGAN,* JAMES BOWIE, PAUL SCHIMMEL

Many RNAs are complex, globular molecules formed from elements of secondary and tertiary structure analogous to those found in proteins. Little is known about recognition of RNAs by proteins. In the case of transfer RNAs (tRNAs), considerable evidence suggests that elements dispersed in both the one- and three-dimensional structure are important for recognition by aminoacyl tRNA synthetases. Fragments of alanine tRNA synthetase were created by in vitro manipulations of the cloned *alaS* gene and examined for their interaction with alanine-specific tRNA. Sequences essential for recognition were located near the middle of the polypeptide, juxtaposed to the carboxyl-terminal side of the domain for aminoacyl adenylate synthesis. The most essential part of the tRNA interaction strength and specificity was dependent on a sequence of fewer than 100 amino acids. Within this sequence, and in the context of the proper conformation, a segment of no more than 17 amino acids was responsible for 25% or more of the total synthetase-tRNA free energy of association. The results raise the possibility that an important part of specific RNA recognition by an aminoacyl tRNA synthetase involves a polypeptide segment that is short relative to the total size of the protein.

AMINOACYL TRANSFER RNA SYNthetases catalyze the attachment of amino acids to their cognate transfer RNAs (tRNAs) (which have specific triplet anticodons) and in this way establish the rules of the genetic code. The sequences of many tRNAs have been determined (*1*), and all can be folded into the common cloverleaf structure (*2*). The cloverleaf structure is in turn folded into an L-shaped tertiary structure, with different tRNAs sharing common, stabilizing tertiary interactions (*3*).

Because of the folded, globular structure, there are more levels of complexity in the interaction of aminoacyl tRNA synthetases with tRNAs than, for instance, occur in the interactions of repressor proteins with DNA (*4*). Repressors interact with a discrete segment of linear double helix. In contrast, chemical modification, ultraviolet and chemical cross-linking, and nuclease protection experiments show that regions of a tRNA molecule that are contacted by an aminoacyl tRNA synthetase are not contiguous. In particular, contacts have typically been observed in the acceptor stem, the D-stem, and, in some cases, the anticodon stem and loop (*5*).

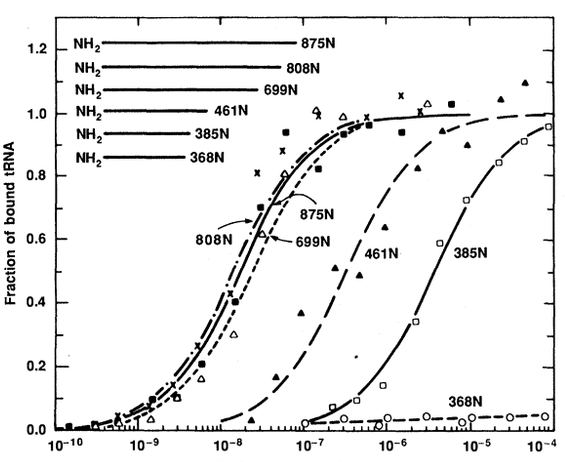

Fig. 1. Binding of ^{32}P-labeled tRNA^{Ala} to the proteins described in Table 1. The curves show fraction of the maximum tRNA^{Ala} that is bound versus the concentration of protein monomer. Protein monomer concentrations were standardized to those determined by active site titrations (*33*). The tRNA^{Ala} (Subriden RNA) was labeled with ^{32}P at the 5′ phosphate and then purified as described (*23, 34, 35*). Nitrocellulose filter-binding experiments were based on the method described by Yarus and Berg (*12*). Conditions (optimized for the Ala-tRNA synthetase interaction with tRNA) were 10 m*M* $MgCl_2$, 60 m*M* potassium phosphate (at *p*H 5), 100 μg of BSA per milliliter, and about 10^{-13} to 10^{-12}*M* [^{32}P]tRNA^{Ala}. For fragments 875N to 461N, the maximum concentration of protein, over most of the binding curve, was less than the concentration of BSA (100 μg/ml) that was introduced to counter nonspecific binding. For fragment 368N, no binding was observed up to 100 μg/ml. To determine the curve for this fragment, we measured nonspecific binding of tRNA^{Ala} to varying concentrations of BSA, up to and beyond 100 μg/ml; this was compared with the binding of fragment 368N over the same concentration range. No binding for fragment 368N above the nonspecific binding to BSA was detected. For fragment 385N, the binding was well above the nonspecific binding to BSA. To test the possibility that an inhibitory contaminant is present in the fragment 368N preparation, we measured the binding curve for wild-type enzyme (875N) in the presence of 4.8 μ*M* 368N protein. The binding of wild-type enzyme to tRNA^{Ala} was unaltered.

In comparison to the relative uniformity of tRNA structures, aminoacyl tRNA synthetases show a vast array of subunit molecular weights and quaternary structures, with limited primary sequence homologies (*6, 7*). Although there is some three-dimensional structural information on adenylate binding sites in these enzymes (*8, 9*), we know of no example of a high-resolution crystal structure of an entire tRNA-binding region of an aminoacyl tRNA synthetase.

Escherichia coli Ala-tRNA synthetase is a tetramer of four α chains of 875 amino acids encoded by *alaS* (*10*). To delineate the parts of the protein required for recognition of tRNA^{Ala}, we have used *alaS* gene deletions and investigated the resulting protein fragments. The native protein and proteins that have from 20 to 500 amino acids deleted from the COOH terminus were investigated (Table 1). These deletions affect the oligomeric state, as determined in vitro (*10*). Fragment 699N (synthesized from an *alaS* gene that encodes only the NH_2-terminal 699 codons) and the smaller NH_2-terminal fragments are monomers; fragments 808N, 852N, and 875N (wild-type enzyme) are tetramers.

Each fragment had alanyl adenylate synthesis activity indistinguishable from that of the wild-type enzyme. The aminoacylation activity was determined in vivo and in vitro. Except for 368N and 385N, each fragment complemented an *alaS* null strain (*11*), when it was expressed from a multicopy plasmid (column 5, Table 1). Only fragment 852N, however, had activity in vitro comparable to that of the wild-type enzyme (column 4). Fragments 461N to 808N had a greatly reduced catalytic constant (k_{cat}) for aminoacylation, which evidently was unrelated to the oligomeric state. The reason for the reduction in k_{cat} is not yet known.

To measure the ability of these proteins to bind to tRNA^{Ala} separately from their catalytic properties, we used a nitrocellulose filter-binding assay (*12*). The binding curves for fragments 875N, 808N, and 699N (Fig. 1) were essentially identical and show that the three proteins bind to tRNA^{Ala} with dissociation constants of 17 n*M*, 13 n*M*, and 26 n*M*, respectively. Because fragments 875N and 808N are tetramers, while fragment 699N is a monomer, the binding curves show that binding occurred to a single chain and that the strength of association was not altered by tetramer formation.

Deletion of an additional 238 amino acids

Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139.

*Present address: E. I. du Pont de Nemours and Company, Wilmington, DE 19898.

Table 1. Properties of fragments of alanine tRNA synthetase.*

Protein	Oligomeric state	Aminoacyl adenylate formation	Aminoacylation	
			In vitro	In vivo
875N	α_4	Yes	Yes	Yes
852N	α_4	Yes	Yes	Yes
808N	α_4	Yes	Reduced	Yes
699N	α	Yes	Reduced	Yes
461N	α	Yes	Reduced	Yes
385N	α	Yes	No	No
368N	α	Yes	No	No

*The wild-type protein is designated as 875N for the monomer chain length, which is 875 amino acids. Fragments 852N, 808N, 699N, 461N, and 385N were synthesized from recombinant plasmids that harbor *alaS* genes that encode NH_2-terminal peptides that extend to the respective codons. Each protein had the expected molecular weight within the resolution of SDS gel electrophoresis (*10*). The wild-type and *alaS* deletion mutant genes were cloned behind the Tac promoter (*22, 23*) to elevate production of the proteins. Fragments 875N to 461N each complement a chromosomal deletion of *alaS* (*11*) and were purified from the deletion background (*24, 25*). Fragment 385N, which does not complement the deletion strain, was purified from a wild-type background. The amount of the low level of wild-type contamination in the preparation of fragment 385N was determined quantitatively by use of Western blot analysis (*26, 27*). It was found to be at least 50 times less than what would be required to generate the binding curves observed through wild-type protein contamination. Fragment 368N was generated by trypsin digestion of fragment 461N, as described for its preparation from wild-type protein (*28*). The NH_2-terminal sequence was confirmed to be the same as that of the wild-type protein by Edman degradation. The COOH-terminal amino acid was established by a mass spectrometric analysis of digests of the protein (*29*). All protein preparations were free of interfering nuclease contamination. The oligomeric form of each protein was determined by gel-filtration chromatography (*10*). Aminoacyl adenylate synthesis activity was determined by the ATP-pyrophosphate exchange assay (*30*). In vivo complementation of a chromosomal deletion of *alaS* was as described (*10, 24*). In vitro aminoacylation assays were according to a standard procedure (*31, 32*). "Reduced" indicates that k_{cat} for this reaction was diminished about 1000-fold relative to the value for the wild-type enzyme.

from fragment 699N gives fragment 461N. This protein was able to bind tRNA with a dissociation constant 20 times that of the wild-type protein. This corresponds to a contribution by the deleted sequences of −1.8 kcal/mol to the total interaction free energy of −10.7 kcal/mol (assuming a standard state of 1 mol/liter).

Removal of 76 additional amino acids from the COOH terminus gives fragment 385N; the dissociation constant for binding to $tRNA^{Ala}$ was 3.7 μM. This was 200 times higher than the dissociation constant for the wild-type enzyme-$tRNA^{Ala}$ complex. This fragment had the full adenylate synthesis activity, but it did not aminoacylate tRNA (Table 1). The data in Fig. 1 suggest that the defect in aminoacylation activity was due not to tRNA binding but rather to a defect in the transfer of amino acid from the adenylate to the bound tRNA. This conclusion was strengthened by our failure to detect aminoacylation by in vitro assays conducted at $tRNA^{Ala}$ concentrations as high as 100 μM.

We detected no association of fragment 368N with $tRNA^{Ala}$, even though we estimate we could detect binding that is reduced 100-fold from that of fragment 385N (corresponding to a dissociation constant of 370 μM). Therefore, the dissociation constant for fragment 368N was $\geq$370 μM. The result suggests a crucial role in tRNA binding for some of the amino acids between positions 368 and 385.

The curves in Fig. 1 are calculated on the assumption that there is no cooperativity in the binding of protein fragments to $tRNA^{Ala}$ (*13*). Cooperativity would be manifest if the monomeric fragments formed oligomers upon binding to $tRNA^{Ala}$. That the protein binds tRNA as a monomer was confirmed for fragment 461N. Gel filtration chromatography was performed, under the conditions of the filter-binding assay, on fragment 461N alone and in the presence of saturating amounts of $tRNA^{Ala}$. Fragment 461N migrated on the column as a monomer, and the fragment-tRNA complex migrated as the sum of protein and tRNA molecular weights for a complex with a 1:1 stoichiometry. This result is in agreement with the hyperbolic tRNA binding curves observed for all of the proteins studied.

Although the fragments used in these studies were judged to be at least 95% pure, contamination with another synthetase could lead to nonspecific association of that contaminant with the $tRNA^{Ala}$ used in the assay (*14, 15*). We examined this possibility with gel filtration experiments on fragment 385N, which has the weakest association with $tRNA^{Ala}$. The elution profile of the column was monitored by assays of alanine-dependent adenosine triphosphate (ATP)–pyrophosphate exchange activity (adenylate synthesis) and $tRNA^{Ala}$ binding. The two activities coincided. As positive controls, we used fragments 699N and 808N, each of which binds strongly to $tRNA^{Ala}$. The gel filtration profiles obtained with these fragments also showed coincident peaks of alanine-dependent ATP-pyrophosphate exchange activity and of $tRNA^{Ala}$ binding at the expected locations in the profile. These results collectively establish that the fragments themselves are responsible for the observed binding to $tRNA^{Ala}$.

A second issue is the specificity of the association of the fragments with $tRNA^{Ala}$. All experiments were performed with 100 μg/ml of bovine serum albumin (BSA) in the binding mixture. In this way, we detected only interactions that were above the background of nonspecific association with this protein. For proteins 875N, 808N, 699N, and 461N, the concentration of BSA was in excess of the fragments over most of the concentration range in which the fragments bound to $tRNA^{Ala}$. For fragments 385N and 368N, concentrations higher than 100 μg/ml were used for some of the data points. For that reason, we investigated the weak nonspecific association of BSA with tRNA at higher protein concentrations. The binding of BSA to $tRNA^{Ala}$ is inconsequential compared with that of protein 385N (Fig. 1). In contrast, protein 368N shows no binding beyond the weak interaction observed with BSA. Because protein 368N was used at concentrations considerably higher than those of the other fragments, it also serves as an additional control against nonspecific association of tRNA with protein.

Experiments were also performed to compare the relative affinities of the fragments for $tRNA^{Ala}$ with those for a population of noncognate tRNAs. For this purpose, an amount of each protein was used such that, in the absence of unlabeled competitor tRNA, all of the ^{32}P-labeled $tRNA^{Ala}$ was bound. A constant amount of ^{32}P-labeled $tRNA^{Ala}$ was mixed with increasing amounts of either unlabeled $tRNA^{Ala}$ or of total tRNA and analyzed in the standard binding assay (Fig. 1). By comparing the amount of unlabeled total tRNA relative to unlabeled $tRNA^{Ala}$ that was able to compete away half of the bound ^{32}P-labeled $tRNA^{Ala}$, it is possible to estimate the affinity of $tRNA^{Ala}$ relative to that of unfractionated tRNA.

For fragment 875N (the wild-type enzyme) there was a tenfold preference in the apparent binding affinity for $tRNA^{Ala}$ versus the unfractionated tRNA. This is because approximately 5% of the unfractionated tRNA mixture is $tRNA^{Ala}$ (as determined by independent measurements). Thus, the preference for $tRNA^{Ala}$ was within a factor of 2 of what would be expected if there were no binding to the noncognate tRNAs in the unfractionated mixture. This confirms that, for the wild-type protein, the binding assay detects the specific synthetase-tRNA interaction.

With fragments 461N and 385N, the difference in binding affinities was reduced

Table 2. Summary of binding of $tRNA^{Ala}$ to alanine tRNA synthetase fragments at 25°C. Abbreviation: $\Delta G°$, Gibbs free energy.

Protein	Dissociation constant (n*M*)	$\Delta G°$ (kcal/mol)*
875N	17	10.7
808N	13	10.8
699N	26	10.4
461N	340	8.9
385N	3,700	7.6
368N	≥370,000	≤4.7

*The standard state is 1*M*.

only threefold relative to that observed with the wild-type enzyme. Thus, even fragments that have lost much of their affinity for $tRNA^{Ala}$ nonetheless retain significant specificity. The results support the interpretation that the smaller proteins are binding to $tRNA^{Ala}$ in essentially the same mode as the wild-type enzyme, but with fewer interactions.

Additional experiments have further addressed the question of specificity. Protein 461N, when expressed from a multicopy plasmid, can serve as the sole source of Ala-tRNA synthetase activity to sustain cell growth in vivo. This implies that misacylations with this enzyme fragment are minimal or nonexistent in vivo (*16*). We also aminoacylated unfractionated tRNA in vitro with alanine, by using the wild-type protein 875N. After completion of the aminoacylation, we added protein 461N to determine whether additional aminoacylation (for example, of species other than $tRNA^{Ala}$) could occur. No incremental aminoacylation of unfractionated tRNA by protein 461N was detected.

These experiments indicate that the fragments retain significant specificity for $tRNA^{Ala}$ even though the absolute affinity for the nucleic acid drops as certain critical sequences are removed from the protein.

Winter and co-workers have recently proposed a model for binding of $tRNA^{Tyr}$ to tyrosine-tRNA synthetase (*17, 18*). This enzyme is a dimer of subunits of 419 amino acids. On the basis of kinetic studies of mixed dimers of mutants, they propose that $tRNA^{Tyr}$ binds across the two subunits of the dimer. In Ala-tRNA synthetase, all of the sequences for tRNA binding are contained within an isolated chain, and it is therefore more straightforward to define and study the regions essential for RNA recognition.

Table 2 summarizes the affinities of the various fragments for $tRNA^{Ala}$ and tabulates the amount by which the free energy of association is reduced by removal of specific polypeptide sequences. The simplest interpretation of the results is that the change in free energy of interaction with $tRNA^{Ala}$ is attributable to removal of contacts made by groups in the deleted sequences, rather than to conformational changes induced at a distance. The following evidence supports this interpretation.

1) All of the proteins were at least as resistant as the wild-type enzyme to proteolytic degradation in crude extracts. The result suggests that each is stably folded. This is also consistent with our ability to purify large quantities of each of the proteins, without experiencing losses any greater than those normally incurred in preparations of the wild-type enzyme.

2) None of the fragments were altered in aminoacyl adenylate synthesis. This means that amino acid and ATP binding sites (and other residues important for catalysis) were not distorted.

3) We created a point mutation at position 409 in the wild-type enzyme, which is within the region implicated by the deletion analysis to be involved in tRNA binding. This mutation had an effect specifically on tRNA binding and not on the binding of ATP or alanine (*16*). This is independent evidence that the region interacts with $tRNA^{Ala}$.

4) The lack of binding by fragment 368N is probably not due to local denaturation of the COOH-terminal end of this protein. This fragment is generated by trypsin digestion of fragment 461N for several hours under native conditions (see footnote to Table 1). The COOH terminus of this protein was established as Arg368. Because fragment 368N is a limit digestion product, it is unlikely that the COOH-terminal end is untethered from the rest of the protein.

Fragment 461N has a free energy of association with $tRNA^{Ala}$ of −8.9 kcal/mol. More than 40%, and perhaps all, of its free energy of interaction with $tRNA^{Ala}$ depends on sequences between amino acids 368 and 461. Most of this interaction energy depends in turn on the surprisingly short segment between amino acids 368 and 385. This segment contributes −2.8 kcal/mol or more to the total synthetase-tRNA interaction free energy. This represents, at a minimum, 25% of the total free energy of interaction of wild-type enzyme with $tRNA^{Ala}$.

These critical sequences are juxtaposed to the COOH-terminal side of the domain for adenylate synthesis. This location is plausible because the 3′ end of the bound tRNA reacts with the adenylate in the aminoacylation reaction. This assignment of the location within the structure for a critical interaction site is supported by work on other systems. Several other synthetases, for example, have their adenylate synthesis domains located within their NH_2-terminal halves (*8, 10, 19*). Two cross-linking studies provide evidence that contacts with the 3′ end of the tRNA are made on the COOH-terminal side of the ATP binding sites (*20, 21*). Future experiments will determine whether, in alanine tRNA synthetase, the region between amino acids 368 and 385 is in close contact with the 3′ end of $tRNA^{Ala}$.

REFERENCES AND NOTES

1. M. Sprinzl, J. Moll, F. Meissner, T. Hartmann, *Nucleic Acids Res.* **13**, r1 (1985).
2. R. W. Holley *et al.*, *Science* **147**, 1462 (1965).
3. S. H. Kim *et al.*, *ibid.* **185**, 435 (1974).
4. C. O. Pabo and R. T. Sauer, *Annu. Rev. Biochem.* **53**, 293 (1984).
5. P. R. Schimmel, *Adv. Enzymol. Relat. Areas Mol. Biol.* **49**, 187 (1979).
6. J. Ofengand, in *Molecular Mechanisms of Protein Biosynthesis*, H. Weissbach and S. Pestka, Eds. (Academic Press, New York, 1977), p. 7.
7. P. R. Schimmel and D. A. Soll, *Annu. Rev. Biochem.* **48**, 601 (1979).
8. D. M. Blow *et al.*, *J. Mol. Biol.* **171**, 571 (1983).
9. A. R. Fersht *et al.*, *Nature (London)* **314**, 235 (1985).
10. M. Jasin, L. Regan, P. Schimmel, *ibid.* **306**, 441 (1983).
11. M. Jasin and P. Schimmel, *J. Bacteriol.* **159**, 783 (1984).
12. M. Yarus and P. Berg, *J. Mol. Biol.* **28**, 479 (1967).
13. The data were least-mean squares fit to the equilibrium equation for noncooperative binding of protein to tRNA. The fitting parameters were K_d, total concentration of $tRNA^{Ala}$, and the ratio of ^{32}P-labeled $tRNA^{Ala}$ to total concentration of $tRNA^{Ala}$. The data and model were then normalized to be unity at large total protein concentrations.
14. B. Roe, M. Sirover, B. Dudock, *Biochemistry* **12**, 4146 (1973).
15. S. S. M. Lam and P. R. Schimmel, *ibid.* **14**, 2775 (1975).
16. C. Ho, M. Jasin, P. Schimmel, *Science* **229**, 389 (1985); L. Regan *et al.*, in preparation.
17. H. Bedouelle and G. Winter, *Nature (London)* **320**, 371 (1986).
18. P. Carter, H. Bedouelle, G. Winter, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 1189 (1986).
19. T. Webster, H. Tsai, M. Kula, G. A. Mackie, P. Schimmel, *Science* **226**, 1315 (1984).
20. C. Houtondji, S. Blanquet, F. Lederer, *Biochemistry* **24**, 1175 (1985).
21. C. Houtondji, F. Lederer, P. Dessen, S. Blanquet, *ibid.* **25**, 16 (1986).
22. H. A. De Boer, C. J. Comstock, M. Vasser, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 21 (1983).
23. L. Regan, thesis, Massachusetts Institute of Technology (1986).
24. Z.-Q. Chen, J.-J. P. Kim, C.-S. Lai, A. H. Mehler, *Arch. Biochem. Biophys.* **233**, 611 (1984).
25. L. Regan, J. D. Dignam, P. Schimmel, *J. Biol. Chem.* **261**, 5241 (1986).
26. M. Jasin, L. Regan, P. Schimmel, *Cell* **36**, 1089 (1984).
27. W. N. Burnette, *Anal. Biochem.* **112**, 195 (1981).
28. W. C. Herlihy, N. J. Royal, K. Biemann, S. D. Putney, P. R. Schimmel, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 6531 (1980).
29. H. A. Scoble and K. Bierman, in preparation.
30. R. Calender and P. Berg, in *Procedures in Nucleic Acids Research*, G. L. Cantoni and D. R. Davies, Eds. (Harper & Row, New York, 1966), p. 384.
31. A. A. Schreier and P. R. Schimmel, *Biochemistry* **11**, 1582 (1972).
32. S. D. Putney, R. T. Sauer, P. R. Schimmel, *J. Biol. Chem.* **256**, 198 (1981).
33. A. R. Fersht *et al.*, *Biochemistry* **14**, 1 (1975).
34. H. Donis-Keller, A. M. Maxam, W. Gilbert, *Nucleic Acids Res.* **4**, 2527 (1977).
35. T. Maniatis and A. Efstratiadis, *Methods Enzymol.* **65**, 299 (1980).

36. Supported by NIH grant GM23562. L.R. was a Fulbright Scholar. We thank S. G. J. Mochrie for the fitting program and for help and advice in its application.

14 August 1986; accepted 30 December 1986