Approaches to Understanding the Mechanism of Specific Protein–Transfer RNA Interactions

Paul R. Schimmel

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

Received February 17, 1977

Although the different tRNAs share much in common, there is a class of enzymes—the aminoacyl-tRNA synthetases—which accurately discriminate between various tRNA species.[1-4] For each amino acid there is at least one aminoacyl-tRNA synthetase which catalyzes the esterification of the amino acid to a ribose hydroxyl on the 3′-terminal adenosine of its cognate transfer RNA.[1-4] This reaction must be done with absolute fidelity; if a synthetase were to attach its amino acid to a noncognate tRNA, the amino acid could then be incorporated into an incorrect position of a growing polypeptide chain, producing the same end result as a mutation.[5]

A central question is the mechanism by which these enzymes specifically discriminate between various tRNA molecules. This question is particularly intriguing in view of the many common structural features shared by the various tRNA molecules.[6]

In recent years it has become clear that tRNA molecules may be designed as receptors for many different proteins, and also that they may serve functions other than those directly involving protein synthesis (see the introductory article of this issue[6]). Thus, the question of the mechanism by which proteins recognize specific tRNA molecules is of broad interest. It obviously presents a great challenge.

The pursuit of this question has been given great impetus by the availability of a three-dimensional structural model for tRNA.[7,8] With this in mind, let us consider in detail some experimental approaches to understanding the mechanism of specific protein–tRNA complex formation. Studies of this kind have largely been carried out with the well-characterized synthetase–tRNA systems.[2-4] The discussion which follows therefore concentrates on these systems. However, the approaches used and considerations involved are applicable to a broad range of protein–nucleic acid systems, many of which (such as chromatin) are more complex than the synthetase–tRNA systems.

Stability of Synthetase–tRNA Complexes

At the outset it is useful to consider thermodynamic features of synthetase–tRNA interactions. Although some studies have estimated the strengths of the interactions from steady-state kinetic data on the aminoacylation reaction,[9-11] there have been few studies employing direct methods for measuring complex formation. Some illustrative data obtained by direct measurements of 1:1 complex formation are given in Table I.[12] Association constants measured by fluorescence quenching under optimal binding conditions are summarized in this table. These data were obtained with yeast Val-tRNA synthetase and *E. coli* Ile-tRNA synthetase; measurements were done with cognate and noncognate tRNAs from both organisms. In the case of Ile-tRNA synthetase, it is clear that the strongest association occurs with the cognate tRNA from *E. coli*, with binding to other species from *E. coli* being much weaker. However, a strong association is found with noncognate tRNAs from yeast. Conversely, yeast valine tRNA synthetase binds strongly to certain *E. coli* noncognate tRNAs.

Table I
Association Constants for Cognate and Noncognate Enzyme-tRNA Complexes at pH 5.5, 17 °C

Enzyme	tRNA	K, M^{-1}
Yeast ValRS	Val (yeast)	$\simeq 10^8$
	Val (*E. coli*)	3.2×10^7
	Ile (*E. coli*)	9.1×10^6
	Phe (*E. coli*)	2.8×10^6
	Glu (*E. coli*)	$<10^4$
E. coli IleRS	Ile (*E. coli*)	$\simeq 10^8$
	Val (yeast)	2.0×10^7
	Phe (yeast, -Y)[a]	2.5×10^6
	Phe (*E. coli*)	1.1×10^5
	Tyr (*E. coli*)	$<10^4$
	Glu (*E. coli*)	$<10^4$

[a] Phe (yeast, -Y) has the Y base removed from $tRNA^{Phe}$. Data are from ref 12.

The data in Table I are incomplete in that not every possible tRNA species has been tested with each synthetase. Certainly the limited results given for *E. coli* Ile-tRNA synthetase indicate that among the *E. coli* tRNAs tested there is considerable preference to bind *E. coli* $tRNA^{Ile}$. However, some kinetic data imply that

Paul Schimmel is Professor of Biochemistry and Biophysics at Massachusetts Institute of Technology. He received his A.B. degree from Ohio Wesleyan University in 1962 and went on to do graduate studies at Tufts University School of Medicine, Massachusetts Institute of Technology, and Cornell University. He also carried out postdoctoral work at Stanford University. During the years of graduate and postdoctoral study he was associated with Gordon Hammes and Paul Flory. His research interests are on structure–function relationships of proteins, nucleic acids, and protein–nucleic acid complexes, with particular emphasis on transfer RNA and protein–transfer RNA interactions. He is the recipient of the 1978 American Chemical Society Pfizer Award in Enzyme Chemistry.

(1) G. D. Novelli, *Annu. Rev. Biochem.*, **36**, 449 (1967).
(2) L. L. Kisselev and O. O. Favorova, *Adv. Enzymol. Relat. Areas Mol. Biol.*, **40**, 141 (1974).
(3) D. Söll and P. R. Schimmel, *Enzymes, 3rd Ed.*, **10**, 489 (1974).
(4) J. Ofengand in "Molecular Mechanisms of Protein Biosynthesis", H. Weissbach and S. Peska, Ed., Academic Press, New York, N.Y. in press.
(5) F. Chapeville, F. Lipmann, G. von Ehrenstein, B. Weisblum, W. J. Ray, Jr., and S. Benzer, *Proc. Natl. Acad. Sci. U.S.A.*, **48**, 1086 (1962).
(6) A. Rich and P. R. Schimmel, *Acc. Chem. Res.*, first paper in this issue.
(7) S. H. Kim, F. L. Suddath, G. J. Quigley, A. McPherson, J. L. Sussman, A. H.-J. Wang, N. C. Seeman, and A. Rich, *Science*, **185**, 435 (1974).
(8) J. D. Robertus, J. E. Ladner, J. T. Finch, D. Rhodes, R. S. Brown, B. F. C. Clark, and A. Klug, *Nature (London)*, **250**, 546 (1974).
(9) M. Yarus, *Biochemistry*, 11, 2352 (1972).
(10) B. Roe, M. Sirover, and B. Dudock, *Biochemistry*, **12**, 4146 (1973).
(11) J. P. Ebel, R. Giegé, J. Bonnet, D. Kern, N. Befort, C. Bollack, F. Fasiolo, J. Gangloff, and G. Dirheimer, *Biochimie*, **55**, 547 (1973).
(12) S. S. M. Lam and P. R. Schimmel, *Biochemistry*, **14**, 2775 (1975).

even within the same organism the noncognate binding interaction in certain cases may be quite strong.[11] In these cases, specificity of aminoacylation is still achieved by an extremely low maximal velocity associated with the noncognate complexes.[11]

The temperature dependence of the stability of the Ile-tRNA synthetase–tRNA^{Ile} complex has been studied at pH 6.5. These data show that $\Delta H°$ of the reaction is about 0 kcal mol^{-1} while $\Delta S°$ is 34 cal deg^{-1} mol^{-1}.[12] Thus, in this case the association is entropically driven. The strong interaction between *lac* repressor and *lac* operator DNA also appears to be entropically driven.[13] The large positive entropy changes associated with formation of these protein–nucleic acid complexes could be due to the neutralization of negative charges on the nucleic acid with positive groups on the protein. In such cases solvation changes associated with charge neutralization can give rise to large positive entropy changes.[14]

While interesting and of obvious importance, these thermodynamic data do not give real insight into the structural features of synthetase–tRNA complexes. Also, they say nothing about the physical basis for the specificity of complex formation. This is a common limitation of thermodynamic studies of complex biological systems; the data obtained are simply too phenomenological in nature to yield detailed insights. For this reason it is necessary to resort to special approaches that give a more informative picture.

Approaches to Architectural Features of Synthetase–tRNA Complexes

Rationale. A prime goal is to determine the parts of the nucleic acid that are in actual contact with the protein. This information enables one to deduce how the protein positions itself on the known three-dimensional structure of tRNA. Moreover, once the enzyme–nucleic acid contact points have been identified, further experiments may be directed at these sites in order to determine which of them are responsible for the specificity and strength of complex formation.

In each of the approaches described below, the aim is to determine enzyme–tRNA contact points. Because the *E. coli* Ile-tRNA synthetase–tRNA^{Ile} complex has been studied with each of them, most of the discussion focuses on this particular complex. This enables one to compare with a single system the results obtained by several different methods. The less extensive data obtained with other complexes are also considered and are compared with those on the isoleucine system.

Oligonucleotide Hybridization. The isoleucine enzyme is a single polypeptide chain with a molecular weight of 112 000.[15] An early set of experiments with this enzyme was based on the observation that unshielded single-stranded regions of transfer RNA can hybridize to complementary tri- and tetranucleotides.[16] This suggested experiments in which the hybridization pattern of tRNA^{Ile} is compared in the presence and absence of bound Ile-tRNA synthetase. The idea is that sites on tRNA^{Ile} that are blocked by bound synthetase should not be accessible to hybridization. These experiments showed that the anticodon and the 3′ terminus on tRNA^{Ile} are simultaneously shielded by bound Ile-tRNA synthetase.[17] Since these two sections are widely separated (75 Å) in the transfer RNA three-dimensional structure (see below),[7,8] these data gave the first indication that the enzyme spans across a large part of the tRNA molecule.

Nuclease Digestion. Another approach is based on the idea that parts of a bound tRNA not covered by synthetase should be accessible to an added ribonuclease.[18,19] Ideally, the nuclease should digest away those parts of the nucleic acid that are not complexed to the protein and leave only a bound core. Unfortunately, it is generally not possible to obtain an extensively digested core, presumably because the complex falls apart when digestion proceeds too far. However, it is possible to obtain useful results on a lightly, or partially, digested complex.

The anticodon of the major species of tRNA^{Ile} is GAU.[20] Because the anticodon is well exposed to the solvent in the tRNA structure, it is possible to find conditions under which the enzyme T_1 ribonuclease (which cleaves specifically after G residues[21]) will cleave specifically after the anticodon G, without significant cleavage at other sites.[22] One should ask whether this cleavage site is masked when the tRNA is bound to the synthetase. Experiments of this kind showed that the bound synthetase protects the anticodon G against cleavage by T_1 ribonuclease.[22] Thus, these experiments confirmed the results obtained by oligonucleotide hybridization which showed that the anticodon is in close proximity to bound enzyme.

Photochemical Cross-Linking. The oligonucleotide hybridization and nuclease digestion approaches are subject to severe limitations. In the case of oligonucleotide hybridization, one can only probe those regions that are capable of binding complementary tri- and tetranucleotides.[16] However, most parts of the tRNA molecule are involved in cloverleaf helical sections or are shielded by the three-dimensional tertiary structure.[7,8] For this reason, very few parts of the molecule can be probed by oligonucleotide hybridization. In the case of nuclease digestion, one is limited by the base specificity of the nuclease so that only those sites containing the proper base can be probed. More seriously, after the first cleavage the structure of the tRNA may be significantly altered so that the relevance of data based on cleavages after the first one is open to question. For these reasons, it is clearly desirable to have alternate approaches not subject to such limitations.

An obvious possibility is to attempt to cross-link covalently regions of the protein and tRNA that are in close proximity. However, it is difficult to find chemical reagents that will satisfactorily accomplish this. There are at least two difficulties. First, there is currently no reagent available which will react with a sufficient variety of protein side chains and of functional groups

(13) A. D. Riggs, S. Bourgeois, and M. Cohn, *J. Mol. Biol.*, **53**, 401 (1970).
(14) W. Kauzmann, *Adv. Protein Chem.*, **14**, 1 (1959).
(15) A. N. Baldwin and P. Berg, *J. Biol. Chem.*, **241**, 831 (1966).
(16) O. C. Uhlenbeck, J. Baller, and P. Doty, *Nature (London)*, **225**, 508 (1970).
(17) P. R. Schimmel, O. C. Uhlenbeck, J. B. Lewis, L. A. Dickson, E. W. Eldred, and A. A. Schreier, *Biochemistry*, **11**, 642 (1972).
(18) W. Hörz and H. G. Zachau, *Eur. J. Biochem.*, **32**, 1 (1973).
(19) S. K. Dube, *Nature (London)*, *New Biol.*, **243**, 103 (1973).
(20) M. Yarus and B. G. Barrell, *Biochem. Biophys. Res. Commun.*, **43**, 729 (1971).
(21) K. Takahashi, *J. Biochem.*, **49**, 1 (1961).
(22) L. A. Dickson and P. R. Schimmel, *Arch. Biochem. Biophys.*, **67**, 638 (1975).

on nucleotide bases so as to be able to cross-link together the diversity of side-chain–base combinations that might be encountered in an actual complex. Second, reaction conditions required for some chemical reactions are too severe for delicate protein–nucleic acid complexes; thus, the problem of distortion or denaturation of the complexes may be severe.

As an alternative, highly reactive photoaffinity labels that permit protein–nucleotide cross-linking might be introduced into the enzyme or tRNA. However, this approach, which has been valuable for studying protein–small ligand complexes,[23-25] is seriously limited in the case of macromolecule–macromolecule complexes such as synthetase–tRNA complexes. A main difficulty is the problem of attaching a photoreactive group to specific sites on a macromolecule. At best, this will only be possible with a few selected sites. Another problem is that introduction of a photoreactive function may seriously perturb the conformation of the macromolecule so as to inhibit complex formation or to produce a distorted complex. There is no clear solution to these problems.

As another alternative, direct photochemical cross-linking by means of irradiation at 254 nm may be attempted. The rationale is that many amino acids or amino acid-like compounds may be photochemically cross-linked to purines and pyrimidines.[26,27] For example, direct irradiation of mixtures of amino acids and uracil suggests that ten or so of the amino acids may react with the pyrimidine.[28] With such a broad spectrum of potential cross-linking reactions, it was plausible to expect that direct irradiation of synthetase–tRNA complexes would facilitate cross-linking at many of the points where amino acid side chains are positioned near nucleotide bases.

It was indeed found that stable joining between a synthetase and tRNA is achieved by direct irradiation.[29-31] This was demonstrated for six different complexes. Although stable cross-links were introduced, a major question was whether such cross-links were formed in specific fashion or whether the proteins and tRNAs were joined together randomly. If the latter, the photochemical approach would be of no value for exploring the topology of synthetase–tRNA complexes. Another concern is whether irradiation introduces structural alterations in the tRNAs and enzymes, so that cross-linking occurs with the reacting species disturbed somewhat from their native conformations.

To explore these questions, some control experiments were carried out. Irradiation of $tRNA^{Ile}$ alone under the conditions used to establish cross-links was found to cause no detectable fragmentation of $tRNA^{Ile}$, and chromatographic data gave no evidence for intramolecular cross-links within the molecule.[30] Furthermore, the kinetics of aminoacylation of irradiated was similar to that of unirradiated $tRNA^{Ile}$. In the case of Ile-tRNA synthetase, irradiation gives a rapid inactivation of catalytic ability but does not greatly damage $tRNA^{Ile}$ binding ability.[30] These observations are reminiscent of those of Iaccarino and Berg who showed that modification of the enzyme's highly reactive sulfhydryl group gives a sharp decrease in catalytic ability with little effect on tRNA binding capacity.[32] The sulfhydryl group is believed to play a role in aminoacyl adenylate formation but is not critical for tRNA binding.[32] Although we cannot be certain of the photochemical lesion that reduces catalytic activity of the free enzyme, it is clear that the tRNA binding capacity is less sensitive to the irradiation. This suggests that the receptor sites involved were not greatly disturbed by the irradiation.

Table II
Photoinduced Cross-Linking of *Escherichia coli* tRNAs and Enzymes[a]

Enzyme	tRNA	% joining
Ile-tRNA synthetase	$tRNA^{Ile}$	40
Bovine albumin	$tRNA^{Ile}$	0
Ile-tRNA synthetase	$tRNA_1^{Tyr}$	0
Ile-tRNA synthetase	$tRNA^{Phe}$	0
Tyr-tRNA synthetase	$tRNA_1^{Tyr}$	50
Bovine albumin	$tRNA_1^{Tyr}$	0
Tyr-tRNA synthetase	$tRNA^{Ile}$	7

[a] All irradiations were done at pH 5.2 or 5.5 and involved an effective dosage of about 0.07–0.12 μeinstein/mm^2 at 254 nm. Data are from ref 30.

Another type of control is to test whether photocross-linking occurs in systems where synthetase–tRNA complex formation is known to be quite weak. This is true, for example, for many noncognate pairs from the same organism (cf. Table I). The occurrence of cross-linking in these systems would raise serious doubt that photojoining requires specific complex formation. To tackle this question, irradiations were carried out with a variety of protein–tRNA mixtures.[30] The results are summarized in Table II,[33] which shows the amount of nucleic acid that is joined to protein when the protein is in excess. In cases where the reaction mixture involved a cognate enzyme–tRNA pair, the concentration of each species was well above the dissociation constant of the complex. The table shows that, whereas Ile-tRNA synthetase cross-links to 40% of $tRNA^{Ile}$, no joining to the noncognate tRNAs—$tRNA^{Tyr}$ and $tRNA^{Phe}$—is detectable; conversely, the nonspecific protein bovine serum albumin does not cross-link to $tRNA^{Ile}$. Table II also shows that, while Ile-tRNA synthetase cannot cross-link to $tRNA^{Tyr}$, good joining occurs between Tyr-tRNA synthetase and $tRNA^{Tyr}$; this demonstrates that the inability of $tRNA^{Tyr}$ to react with Ile-tRNA synthetase is not due to some intrinsic inability of the $tRNA^{Tyr}$ to undergo photochemical cross-linking. Finally, a small amount of joining between Tyr-tRNA synthetase and $tRNA^{Ile}$ occurs. This is consistent with the fact that some complex formation between these species has been detected.[22]

(23) D. J. Brunswick and B. S. Cooperman, *Proc. Natl. Acad. Sci. U.S.A.*, **68**, 1801 (1971).
(24) J. Knowles, *Acc. Chem. Res.*, **5**, 155 (1972).
(25) B. E. Haley and J. F. Hoffman, *Proc. Natl. Acad. Sci. U.S.A.*, **71**, 3367 (1974).
(26) K. C. Smith in "Photochemistry and Photobiology of Nucleic Acids", S. Y. Wang, Ed., Academic Press, New York, N.Y., 1975.
(27) K. C. Smith, Ed., "Aging, Carcinogenesis, and Radiation Biology", Plenum Press, New York, N.Y., 1976.
(28) K. C. Smith, *Biochem. Biophys. Res. Commun.*, **34**, 354 (1969).
(29) H. J. P. Schoemaker and P. R. Schimmel, *J. Mol. Biol.*, **84**, 503 (1974).
(30) G. P. Budzik, S. S. M. Lam, H. J. P. Schoemaker, and P. R. Schimmel, *J. Biol. Chem.*, **250**, 4433 (1975).
(31) H. J. P. Schoemaker, G. P. Budzik, R. Giegé, and P. R. Schimmel, *J. Biol. Chem.*, **250**, 4440 (1975).
(32) M. Iaccarino and P. Berg, *J. Mol. Biol.*, **42**, 151 (1969).
(33) The data in Table II were obtained under optimal binding conditions, i.e., around pH 5.5. However, cross-linking can also be achieved at pH 7, and in one case studied in depth it was shown that regions cross-linked at pH 5.2 are identical with those joined at pH 7.[31]

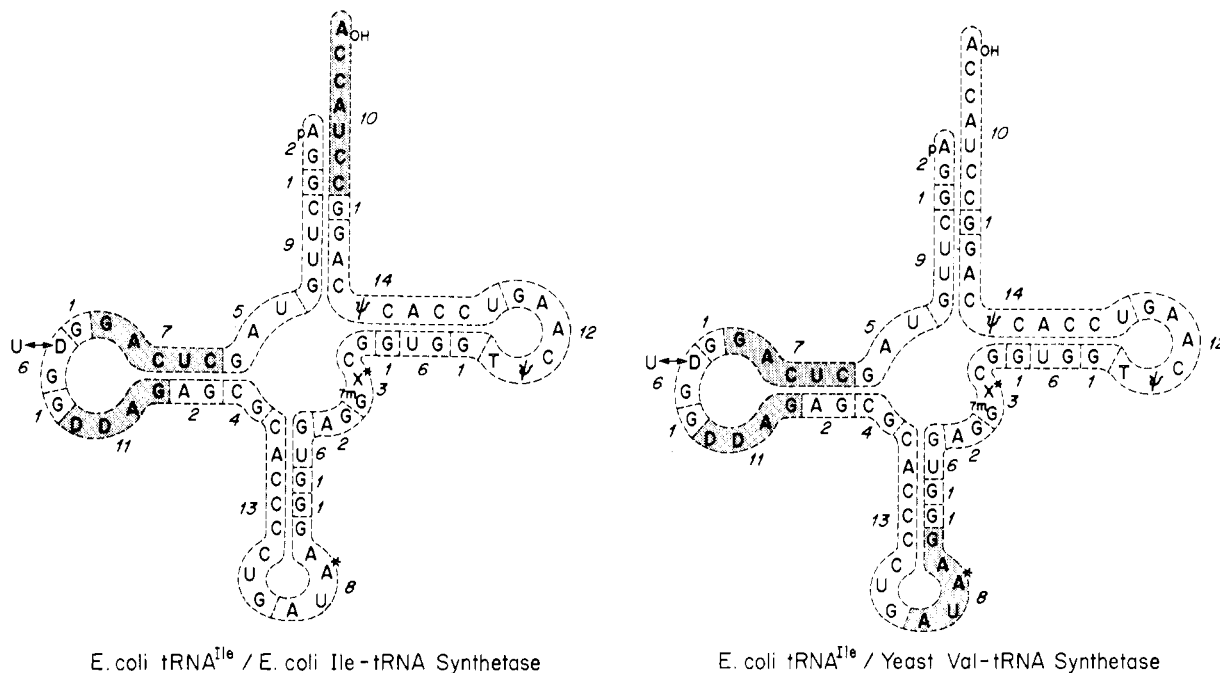

Figure 1. (a, left) Sequence and cloverleaf structure of *E. coli* $tRNA^{Ile}$.[20] T_1 fragments are enclosed by dashed lines and are numbered in accordance with their positions on a chromatogram. Shaded areas are those T_1 fragments found cross-linked to *E. coli* Ile-tRNA synthetase. Adapted from ref 30. (b, right) Sequence and cloverleaf structure of *E. coli* $tRNA^{Ile}$ with T_1 fragments found cross-linked to yeast Val-tRNA synthetase indicated by shading. Adapted from ref 30.

The results of Table II offer strong support for the notion that cross-linking requires specific complex formation. The data argue against the possibility that irradiation generates highly reactive species which slam together in random ways to generate nonspecific protein–nucleic acid cross-links. An additional control reinforced the idea that the nucleic acid must preexist in a bound form in order for cross-linking to be achieved. In these experiments, the Tyr-tRNA synthetase–$tRNA^{Tyr}$ complex was subjected to changes in solvent conditions which are known to weaken the stability of the synthetase–tRNA complex. It was found that, as the stability is weakened, there is a progressive decrease in the cross-linking yield.[29]

These several controls give reasonable confidence in the cross-linking procedure. The next major question is the location on the tRNA of the cross-linking sites. One might naively approach this problem by digesting a cross-linked complex with a nonspecific nuclease that releases all nucleotide units not joined to the protein. One could then hopefully analyze the remaining enzyme–nucleotide adducts. However, this approach would yield little useful information since one would not know the location in the tRNA structure of the covalently linked mononucleotide units. In order to identify the structural location of the cross-linked units, one must digest first with a specific nuclease that cuts at defined sites so as to subdivide the molecule into segments which originate from known parts of the tRNA cloverleaf. For example, T_1 ribonuclease[21] typically cleaves a transfer RNA into 14 or so segments, many of which are derived from a unique part of the structure. Therefore, if a particular T_1 fragment is identified as cross-linked to the synthetase, it is usually possible to locate in the tRNA structure the location of that fragment.

The results of this kind of analysis for the cross-linked Ile-tRNA synthetase–$tRNA^{Ile}$ complex are summarized in Figure 1a. In the cloverleaf structure of $tRNA^{Ile}$, T_1 fragments are enclosed by dotted outline and are numbered in accordance with their position on a chromatogram. The shaded segments show the T_1 fragments that are involved in cross-linking. The base within the fragment that is the actual cross-linking site has not yet been determined. Three regions are found to join to the enzyme. These are derived from the amino acid acceptor terminus and from in and around the dihydrouridine stem of the molecule. Since cross-links are not scattered throughout the tRNA structure, the data provide further support for the nonrandom nature of the cross-linking reaction. Moreover, the regions which do cross-link are ones that are believed to play some role in synthetase–tRNA interactions. For example, the 3′ terminus is the amino acid attachment site, and certainly part of the enzyme must come in close proximity to this section of the tRNA. The dihydrouridine stem region, which is enclosed in part by fragments 7 and 11, is also believed to play some role in synthetase interactions.[2,3] Therefore, the results of Figure 1a suggest that cross-linking has occurred at true enzyme–tRNA proximity or contact points.

We saw earlier that both nuclease digestion and oligonucleotide hybridization experiments suggest that the anticodon region of $tRNA^{Ile}$ is close to the synthetase in the complex. However, this part of the molecule is not found to cross-link. This illustrates an important limitation of the cross-linking approach. Although one has some confidence that the cross-linked regions are in close proximity to the protein, nothing can be said about those sections that do not cross-link. Regions in close proximity may not cross-link simply because of intrinsic photochemical constraints.

We have seen that in certain situations rather stable complexes are formed between a synthetase and a noncognate tRNA species. Thus, the stability constant

(Table I) for the yeast Val-tRNA synthetase complex with *E. coli* $tRNA^{Ile}$ is roughly within an order of magnitude of that for the cognate yeast Val-tRNA synthetase–yeast $tRNA^{Val}$ complex. The existence of such noncognate complexes offers an opportunity to explore an interesting question, namely, whether there is some universality or common features to the various synthetase–tRNA interactions. If, for example, yeast Val-tRNA synthetase is oriented on the noncognate *E. coli* $tRNA^{Ile}$ in much the same way as the cognate *E. coli* Ile-tRNA synthetase orients on this tRNA, one would suspect there are some common features to the molecular pattern of all synthetase–tRNA complexes. Cross-linking is suitable to attack this question.

When the yeast Val-tRNA synthetase–*E. coli* $tRNA^{Ile}$ complex is irradiated, good cross-linking is achieved.[30] Further analysis showed that three regions on $tRNA^{Ile}$ are cross-linked, and these are represented as shaded sections in Figure 1b. Two of the three cross-linking regions are identical with those found with the cognate complex. These two regions are fragments 7 and 11 from the dihydrouridine region. In the noncognate complex, the anticodon section is also cross-linked. As stated above, other data show that this section is close to the enzyme in the cognate complex. Moreover, although fragment 10 (which is cross-linked in the cognate complex) is not cross-linked in the noncognate complex, the enzyme must come close to this area, because it is possible for yeast Val-tRNA synthetase to attach valine to *E. coli* $tRNA^{Ile}$;[34] this necessitates that the 3′-terminus of $tRNA^{Ile}$ come close to the enzyme. In summary, the results shown in Figures 1a,b strongly suggest there is much similarity between the structural features of the cognate complex and a noncognate complex in which the reacting partners are derived from different organisms. Other photochemical cross-linking studies with yeast $tRNA^{Phe}$ also suggest that, for this tRNA, cognate and noncognate complexes have much in common.[31]

Tritium Labeling. The cross-linking approach has clearly yielded some novel structural information, but it has at least two limitations. First, as mentioned previously, no conclusions can be drawn about sites which do not cross-link. Second, there is always the fear that cross-linking itself introduces some perturbation into the structure of the complex, even though a variety of controls suggest that such perturbations, if any, are not large. For these reasons it is desirable to have an alternate approach not subject to these limitations.

An alternative approach is illustrated in Figure 2. The nucleic acid is envisioned as a continuous wavy line and the protein is depicted as a globular structure. The idea is to incubate the protein–nucleic acid complex in the presence of a reagent, X, which will react with all parts of the nucleic acid that are not in contact with the protein, and which will do so without disturbing the structure of the complex. After incubation with compound X, the nucleic acid may be removed and analyzed for sites that have incorporated X. These sites may then be viewed as parts of the molecule not in contact with the protein.

It is a challenge to find a suitable reagent X. Virtually any reactive chemical that will combine with nucleic acid components is also likely to disrupt the complex after reacting with only a few sites. Moreover, many reagents are so bulky that they are not apt to penetrate anything but the most exposed portions of a nucleic acid. For these reasons, we examined a simple isotope exchange reaction.

(34) D. Kern, R. Giegé, and J. P. Ebel, *Eur. J. Biochem.*, 31, 148 (1972).

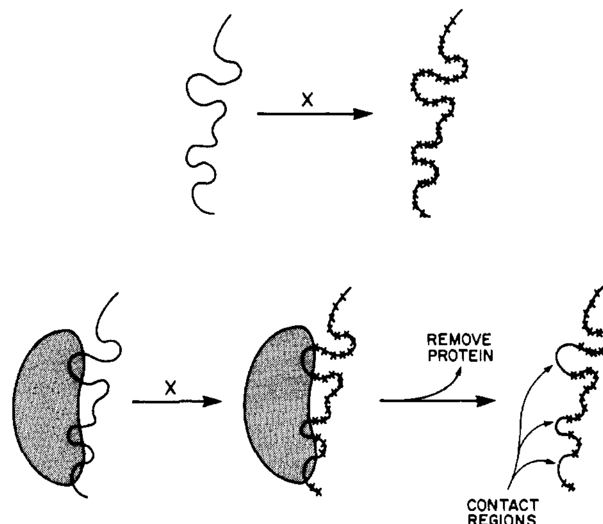

Figure 2. Schematic illustration of a nucleic acid (wavy line) reacting with a reagent (X) in the presence and absence of a bound protein (shaded globular structure). Contact regions in the protein–nucleic acid complex are envisioned as those sections protected from X by the bound protein.

It is well-known that the C-8 hydrogen of purines slowly exchanges.[35-38] At 37 °C the apparent first-order rate constant for exchange is about 10^{-3} h^{-1}.[39] Although this rate constant is low, tritiated water with high specific activity transfers easily measurable radioactivity into the purine units of tRNA in a period of several hours.[39,40] It is then possible to isolate the tRNA and to digest and analyze it in such a way as to determine the amount of tritium that has been incorporated into specific purine units. This analysis is done under conditions where exchange-out does not occur. Since ^3H is seemingly an innocous substitute for ^1H, no tRNA structural perturbations should occur upon incorporation of ^3H. Moreover, the amount of ^3H incorporation in a given interval of time can be measured with great accuracy.[39] Thus, tritium exchange into the C-8 positions of the purine units would appear to be an ideal candidate for reagent X of Figure 2.

We found that the amount of ^3H labeling of a purine is extremely sensitive to its microenvironment.[39-42] For example, when a purine nucleotide is bound to its receptor site on a protein, the rate of ^3H labeling is reduced severalfold.[42] Although certain chemical factors should affect the rate of tritium incorporation, our studies indicate steric factors to be the most important.

(35) M. Maeda, M. Saneyoshi, and Y. Kawazoe, *Chem. Pharm. Bull.*, 19, 1641 (1971).
(36) J. A. Elvidge, J. R. Jones, and C. O'Brien, *Chem. Commun.*, 394 (1971).
(37) M. Tomasz, J. Olson, and C. M. Mercado, *Biochemistry*, 11, 1235 (1972).
(38) J. Livramento and G. J. Thomas, Jr., *J. Am. Chem. Soc.*, 96, 6529 (1974).
(39) R. C. Gamble, H. J. P. Schoemaker, E. Jekowsky, and P. R. Schimmel, *Biochemistry*, 15, 2791 (1976).
(40) R. C. Gamble and P. R. Schimmel, *Proc. Natl. Acad. Sci. U.S.A.*, 71, 1356 (1974).
(41) H. J. P. Schoemaker, R. C. Gamble, G. P. Budzik, and P. R. Schimmel, *Biochemistry*, 15, 2800 (1976).
(42) H. J. P. Schoemaker and P. R. Schimmel, *J. Biol. Chem.*, 251, 6823 (1976).

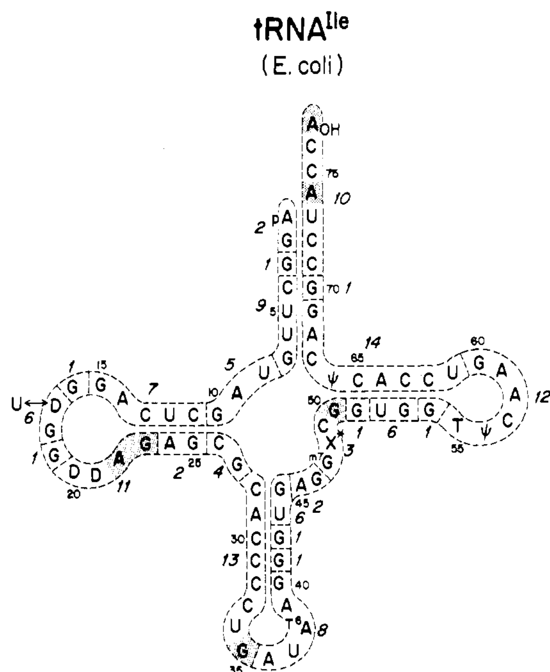

Figure 3. Sequence and cloverleaf structure of *E. coli* tRNA^{Ile}. Dotted outlines enclose T_1 fragments that are indexed with large numbers. The position from the 5′ end of every fifth base in the sequence is indicated by a small number. Shaded bases are those found perturbed in the tritium labeling experiment. Adapted from ref 42.

When a nucleotide unit is shielded, there is a reduction in tritium incorporation.

Tritium labeling of the purines in tRNA^{Ile} was carried out in the presence and absence of Ile-tRNA synthetase. The results are summarized in Figure 3, which indicates by shading the bases that have reduced tritium labeling rates in the presence of Ile-tRNA synthetase. Altogether six purines are perturbed. Comparison with Figure 1a shows that all four of the purines in two of the cross-linked fragments (fragments 10 and 11) are identified as close to the synthetase in the tritium-labeling experiment. The only purines in the cross-linked fragments that are not perturbed in the tritium experiment occur in fragment 7. However, it could well be that cross-linking in this fragment occurs to one of the pyrimidines, while the purines themselves are not in contact with the enzyme.

Figure 3 also shows that G-35 in the anticodon is one of the perturbed purines. As mentioned earlier, both oligonucleotide hybridization and nuclease digestion experiments indicated the anticodon is close to the synthetase in the complex. The remaining purine identified by tritium labeling is G-50. At present, no other data indicate a role for this base.

In summary, five of the six purines identified by tritium labeling occur at sites in the molecule that other techniques have also pinpointed as synthetase-tRNA proximity points. Thus, conclusions obtained by the various methods are well reinforced.

Discussion

Comments on Approaches Used. Several approaches described above have given useful information on the architectural features of synthetase–tRNA complexes. Of the four approaches, photochemical cross-linking and tritium labeling have proved most fruitful. Conclusions obtained from all of the approaches are generally in good agreement. Since the methods used are quite diverse, each appears to give information on the native complex and not on artifactual features that are uniquely sensed by a particular method.

In the case of photochemical cross-linking, at least two issues require further research and clarification. First, there is the problem of the effect of the irradiation on the catalytic activity of the enzyme. Secondly, there is a question of the number of and location of the cross-linking sites on the protein. It would be helpful to know if cross-linking at a particular fragment in the tRNA molecule involves only a single peptide or whether several peptides are cross-linked to a given fragment. The latter eventuality would suggest that there is some distortion or movement of the nucleic acid binding region during the cross-linking reaction. Although both these concerns are to a large extent mitigated by many other controls and by the good agreement between the cross-linking and other studies, further research on these questions is clearly desirable.

A more detailed analysis has been carried out with a photo-cross-linked complex involving Ile-tRNA synthetase and ATP.[43] This nucleotide is used as an energy source to drive the aminoacylation reaction;[1-4] consequently, each synthetase has at least one receptor site for ATP. When the Ile-tRNA synthetase–ATP complex is irradiated in the same fashion as done for the tRNA complexes, coupling between the synthetase and ATP is achieved.[43] Because this is a simpler system, a detailed analysis is more straightforward. In these experiments it was shown that, although the enzyme is catalytically inactivated by UV irradiation, ATP only cross-links to molecules that have retained their catalytic activity; no cross-linking occurs to species that have been inactivated by irradiation. Moreover, protease digestion of the cross-linked synthetase–ATP complex yields a single peptide–nucleotide adduct. This indicates that cross-linking occurs to a specific and not to multiple sites on the enzyme. The sequence of this peptide corresponds to a segment in the partially determined sequence of the enzyme that has been worked out by another investigator.[44] Thus, the more detailed investigations with ATP suggest that the direct irradiation procedure under proper conditions yields cross-linked complexes that are biologically and structurally significant.

Altogether six photo-cross-linked synthetase–tRNA complexes have been investigated in detail.[29-31] Viewed in total, these data show no evidence that cross-linking occurs to parts of the tRNA molecules that are intrinsically more photoreactive. That is, there appears to be no evidence for photoselection to favor cross-linking to regions containing modified bases or unusual structural features. The cross-links appear to represent a reasonable sampling of some of the enzyme–tRNA contact or proximity points in the various complexes.

Architectural Features of the Complex. It is of interest to view on the tRNA three-dimensional structure the regions identified as close to the synthetase in the complex. Figure 4a gives a schematic representation of the backbone of yeast tRNA^{Phe}. The molecule is approximately L-shaped, with one end of

(43) V. T. Yue and P. R. Schimmel, *Biochemistry*, in press.
(44) M. R. Kula, *FEBS Lett.*, **35**, 299 (1973).

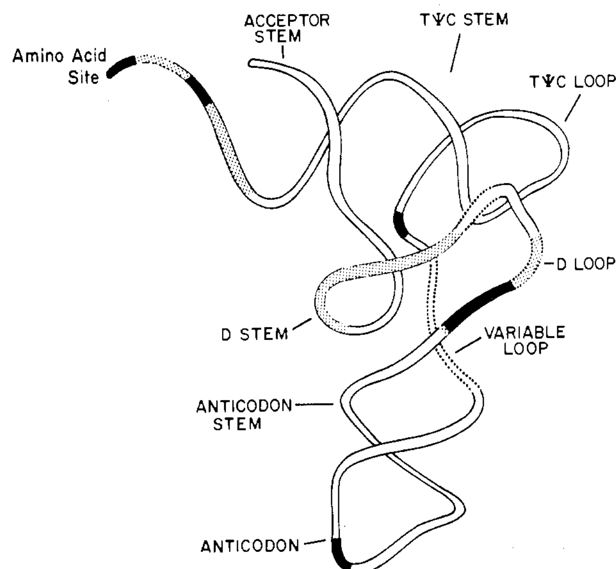

Figure 4. Schematic illustration of three-dimensional conformation of the tRNA backbone.[45] Regions that vary in length among various tRNA species are indicated by dotted outline. Major landmarks on the tRNA structure are identified. Gray areas are T_1 fragments in *E. coli* $tRNA^{Ile}$ found to photo-cross-link to *E. coli* Ile-tRNA synthetase (cf. Figure 1a). The actual base that is cross-linked within each gray fragment has not been identified, so that the gray areas probably overestimate enzyme–tRNA contact or close proximity regions. Dark areas are locations on the backbone of bases identified as close to bound synthetase in the tritium labeling experiment.

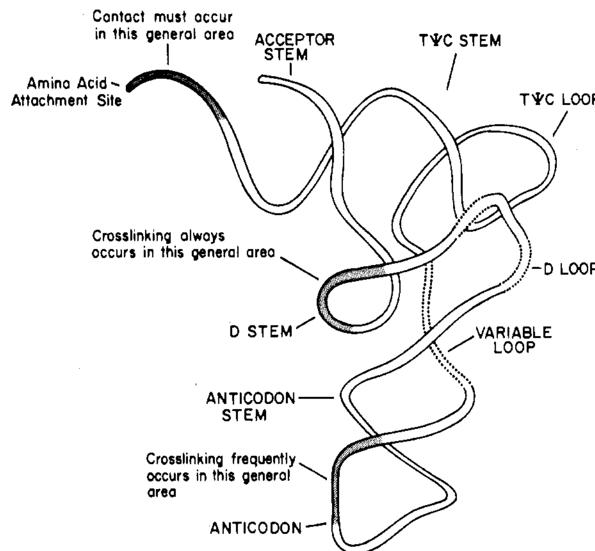

Figure 5. Summary of some of important areas (shaded) for synthetase interactions with tRNAs, based on studies with several systems.[46]

the L at the 3′-acceptor terminus and the other end at the anticodon.[7,8] These two parts of the molecule are separated by about 75 Å.[7,8] It appears that most or all tRNA species may be folded into this basic three-dimensional conformation, with length variations between tRNA species being accommodated in the dotted regions which appear as bulges from the basic structure.[45]

The lightly shaded areas in Figure 4a represent the T_1 fragments on $tRNA^{Ile}$ that are cross-linked to Ile-tRNA synthetase (cf. Figure 1a). The darkly shaded areas are those identified by the tritium-labeling approach as synthetase–tRNA contact, or close proximity, points (cf. Figure 3). It must be kept in mind that only the backbone is visualized in Figure 4, and that the bases (which are the moieties of interest in the tritium labeling and presumably in the cross-linking reaction) project off the backbone.

Clearly the contacts between synthetase and tRNA must span across a large part of the tRNA structure. Many of these areas are confined to the region along the inner part of the L-shaped structure. However, the enzyme also comes close to other parts that are on the side of the L. Thus, a model for the complex is that the enzyme binds along the inner part of the L-shaped structure but that some protein segments also make contact with other parts of the structure, such as, for example, along a side.

As mentioned earlier, other complexes have also been investigated, particularly by the photochemical cross-linking approach. Let us consider some general features of these results. Figure 5 gives a schematic illustration of the tRNA three-dimensional structure; attention is directed to three shaded areas which are believed to be close to or in contact with bound enzyme in many of the synthetase–tRNA complexes investigated.[46] One of these areas is the 3′ terminus, which must be close to the synthetase in all complexes involving cognate pairs since it is the amino acid acceptor site. A second is on the 5′ side of the dihydrouridine stem (e.g., fragment 7 in Figure 1a). This region is found to cross-link to bound synthetase in all complexes investigated.[29–31,47] It is located along the inside of the "L" where the two helical arms of the tRNA structure come together at an approximately right angle. As for the third area, the anticodon region frequently, although not always, cross-links to bound synthetase. These three sections (shaded in Figure 5) all fall along the inner part of the L-shaped structure. Thus, a common feature of most synthetase–tRNA complexes may be that the enzymes bind at least in part along the inner part of the L-shaped structure. In some systems, the synthetase binding may extend as far as the anticodon region, while in others it may fall considerably short of this region.

In addition to the common aspects, however, each enzyme may vary in its interactions with other parts of the tRNA molecule, and these other sections may not lie along the inside of the L-shaped structure. For example, Figure 4 indicates that Ile-tRNA synthetase makes contact with a side of the tRNA molecule, as well as interacting along the inside of the L. Thus, the

(45) S. H. Kim, J. L. Sussman, F. L. Suddath, G. J. Quigley, A. McPherson, A. H.-J. Wang, N. C. Seeman, and A. Rich, *Proc. Natl. Acad. Sci. U.S.A.*, 71, 4970 (1974).

(46) Figure 5 is based in part on photo-cross-linking studies of six complexes involving three different tRNA species with cognate and noncognate enzymes. There is some complication in deciding how much to shade for the areas that appear prominent in photo-cross-linking studies. This is due to the distinct T_1 fragmentation pattern of the different tRNA species. Even if exactly the same position in the tRNA sequences is cross-linked to synthetases in several different complexes, the T_1 fragments that encompass the cross-linked position will not be exactly identical for each tRNA species. In Figure 5 only the overlapping portions shared in common by cross-linked T_1 fragments from the same region are shaded. For example, in three different tRNAs a cross-linked T_1 fragment from the same general area might be derived from bases 32–40, 30–37, and 33–38, respectively; in this case bases 33–37 are common to each of these fragments and would be the shaded area in Figure 5.

(47) Due to an error in the characterization of a chromatogram,[39] the cross-linking of this area was not noted in a previous publication on complexes of yeast $tRNA^{Phe}$.[31] Subsequent work established this to be among the cross-linked segments, as well as the ones reported.

complexes may have an overall topological similarity, with individual structural variations superimposed on the general features.

Concluding Remarks

The results of these various analyses have given considerable insight into the architectural features of synthetase–tRNA complexes. However, these approaches have not revealed the molecular basis for specificity, although they do indicate regions on the tRNA on which to focus further attention. It should also be pointed out that many other approaches have also been employed to explore synthetase–tRNA interactions, such as studies of mutant tRNAs, chemically modified tRNAs, and fragments of tRNAs. References 2–4 review much of the literature in this area.

To consider the specificity and strength of the interaction, let us return to the thermodynamics. As shown in Table I, the association constant under optimal conditions for cognate synthetase–tRNA pairs is on the order of 10^8 M^{-1}. Interestingly enough, a stability constant of this magnitude can readily be achieved if the enzyme only binds to two or three nucleotide units in the entire chain. Thus, Michaelis constants for the interaction of small ligands with enzymes commonly fall in the range of 10^3–10^5 M^{-1}.[48] For example, the complex of Ile-tRNA synthetase with ATP has association constant of around 10^4 M^{-1} at pH 8, 25 °C.[49] Hence twice the free energy of association of ATP with synthetase gives rise to a stability constant on the order of 10^8 M^{-1}. These considerations simply show that subtle effects in the tRNA structure—for example, the presence or absence of a particular base at a particular point—can have large effects on the stability of the complex and also on the probability of forming a particular transition state in the aminoacylation reaction. These facts serve as a reminder that, although our understanding of the complexes has proceeded a great way, the remaining objectives present a vigorous challenge.

As mentioned earlier, tRNAs are believed to complex with a variety of other proteins. Data on these systems may be expected to be forthcoming in due course. It will be interesting to see, for example, if each general class of interactions involves a different area of the tRNA molecule. For instance, it may be that the extra loop, a region that varies in length among different tRNA species, serves as a particular type of receptor (see discussion in ref 6). These kinds of questions will obviously be prime issues in future investigations, and the approaches described above may go a long way toward yielding the answers.

This work was supported by Grant No. GM-15539 from the National Institutes of Health. The many contributions to this work of G. P. Budzik, L. A. Dickson, R. C. Gamble, E. Jekowsky, S. S. M. Lam, H. J. P. Schoemaker, and V. T. Yue are gratefully acknowledged.

(48) L. Stryer, "Biochemistry", W. H. Freeman, San Francisco, Calif., 1975, p 128.

(49) F. X. Cole and P. R. Schimmel, *Biochemistry*, **9**, 480 (1970).

Seven Terminal Steps in a Biosynthetic Pathway Leading from DNA to Transfer RNA

William H. McClain

Department of Bacteriology, University of Wisconsin, Madison, Wisconsin 53706

Received February 15, 1977

Much of the work of science takes the form of pathway elucidation. As the tools of exploration become sharper, they reveal in finer and finer detail the object under investigation. In an accompanying article, Alexander Rich traces the backbone of the polynucleotide chain in yeast phenylalanine transfer RNA ($tRNA^{Phe}$), revealing a fantastically detailed architecture.[1] Living organisms manufacture dozens of different tRNAs, each to precise specifications of nucleotide sequence and chain length. The nucleotide sequence is determined by the DNA, but the chain length is specified by a complex series of steps that form a biosynthetic pathway. In this Account we shall see how it is possible to elucidate such a pathway, with the revelation of detail right down to the molecular level.

The transfer of primary sequence information from DNA to RNA is referred to as transcription. In this process, ribonucleoside 5′-triphosphates pair with their Watson–Crick complements in the DNA. Incoming residues are added as nucleoside 5′-monophosphates by the formation of a covalent sugar–phosphate bond, with the release of pyrophosphate; this reaction is catalized by the enzyme RNA polymerase. But the complete process leading to tRNA is far more complex, since the initial product of transcription is a longer RNA molecule (precursor RNA) containing extra nucleotide residues not present in the tRNA. The presence of these extraneous nucleotide residues means that one or more enzymatic cleavages of the precursor RNA are

William H. McClain received his B.S. degree from Wesleyan College in Mt. Pleasant, Iowa, and in 1968, the Ph.D. from Purdue University. He was a Research Associate at Purdue in 1968, and a postdoctoral fellow from 1969 to 1971 at the Medical Research Council Laboratory of Molecular Biology in Cambridge, England. He then joined the faculty of the University of Wisconsin, where he is Associate Professor. Dr. McClain's principal research interests are in the areas of mutation, genetic recombination, and transfer RNA biosynthesis.

(1) A. Rich, *Acc. Chem. Res.*, **10**, second paper in this issue.