

Current Research
By
William H. McClain,
Halvorson Professor of Bacteriology and Molecular Biology,
Emeritus
Department of Bacteriology
University of Wisconsin-Madison

Statement of purpose: The following two experimental descriptions are not a comprehensive report of my research activities. Rather, they represent part of my total activity and are presented to inform the reader of my ongoing general areas of active research, the novelty of my observations and analyses, and their theoretical content.

Statement of contributions: All novel and original ideas expressed below are entirely mine, not those of collaborators. I am also responsible for all the writing of this document, with the exception of editing by Gabriele Varani of project number 2. I expect that continued future work, experiments, and publication will be shared with the indicated collaborator and possibly their respective student(s).

1. The Nucleotide Degeneracy in RNA-Protein Recognition Hypothesis

My basic research is beyond the forefront of current scientific thought. In all organisms, particular RNAs are recognized in a highly specific process by their specific, cognate proteins. This recognition process plays critical roles in thousands of essential cellular processes.

Transfer RNA (tRNA) recognition by aminoacyl-tRNA synthetase (aaRS) enzymes is a paradigm for the study of the fundamentals of RNA-protein recognition, a process that results in aaRS attachment of the correct amino acid to their cognate tRNA accepting type during the translation of the genetic code in gene expression. Because there are 20 amino acids and 20 tRNA acceptor types, a matrix of 400 potential interactions exist, but specificity demands that only 20 (5%) are the correct ones.

Scrutinizing RNA-protein recognition via that latter attack mode includes the requirement for accurate translation of the genetic code into protein sequences for cell survival; the observed error rates in amino acid attachment to tRNA are less than one in ten thousand. In addition, powerful genetic, biochemical, and computational mathematics allow dissection of this process in utmost detail. Finally, a rich repertoire of 3-D crystal structures is available of tRNA and aaRS complexes. While we know how different aaRS proteins and their respective tRNAs fit together structurally, it is important

to emphasize how little we understand the molecular basis of recognition specificity that leads to complex formation. For example, are nucleotides in tRNA recognized individually and independently of one another or are they recognized as groups of nucleotides that function as a unit in the interaction with their partner aaRS? In addition, what is the functional relevance of the many tRNA backbone-mediated interactions with partner aaRSs observed in crystal structures? Do backbone-mediated interactions with tRNA cause local shifts in the geometry of the pendent bases, which themselves also frequently interact with the proper protein? These theoretical considerations and questions are the center of attention one of my active research explorations.

In experiments involving aspartic acid tRNA and its AspRS enzyme carried out in living *Escherichia coli* cells and subsequently demonstrated biochemically in test tube experiments, I have observed that the sequence of a tRNA “acceptor stem” can be extensively changed (5 to 8 new nucleotides of 15 total) in a series of mutants to new, radically different nucleotide sequences while retaining the associated aspartic acid tRNA aminoacylation specificity and efficiency; in fact, the alternative sequence mutants support aspartic acid tRNA knock out cell growth. Only particular nucleotide combinations are active, and, in fact, most (>99.99%) alternative combinations are inactive. Other workers in the RNA-protein recognition field do not question the validity of these results, but nevertheless consider them eccentric if not bizarre.

In retrospect, the predominance of tRNA backbone-mediated interactions observed between AspRS and the acceptor stem of aspartic acid tRNA in the 3-D complex suggests that many of the corresponding wild-type bases should be replaceable by alternative sequences so long as they preserve the initial backbone structure of the tRNA. A corollary of structure-based recognition is that *the smallest unit of function is a group of nucleotides rather than a single nucleotide or base pair*. An aminoacylation specificity that depends on molecular structure rather than base sequence can explain why aspartic acid tRNA with divergent acceptor stem sequences can be functionally active.

It would be incorrect to link the functional activity of any mutant tRNA to an individual base pair or imperfect, non-Watson-Crick pair that it contains since the sequence context in which that element exists undoubtedly contributes to the molecule's functional activity. Clearly, the amino acid acceptor stem as a whole is being recognized. The twist angles and overall pitch determine acceptor stem backbone geometry, which is critical for the terminal amino acid accepting residue to fit into the enzyme active site. I have experimentally dissected other tRNA-aaRS pairs to learn the extent to which RNA-protein recognition systems in general use structure-base, rather than individual nucleotide focused recognition.

In further experiments performed in the alanine tRNA system, active mutants were selected in alanine tRNA knockout *E. coli* cells that drastically mutated all of the 7 critical recognition nucleotide residues that are located in the tRNA's acceptor stem and define this its acceptor type. They are critical to alanine acceptor function because substituting any one of the 7 wild type residues inactivates the acceptor activity of the

modified tRNA. Most notable in the selected mutant was the obliteration of a G-U 'wobble pair' in the acceptor stem that is the major tRNA acceptor recognition tag, not only in bacterial alanine tRNAs, but also in those of all other organisms. When we, and subsequently others, discovered the critical function of this G-U pair in 1988, the popular press (e.g., NYT front page) dubbed it a 'Second Genetic Code.' To the level of the precision of our experimental determinations, the selected novel mutant retains the full repertoire of cellular and biochemical functions relative to those of wild type alanine tRNA molecule.

In view of the structural recognition hypothesis derived above in the aspartic acid system, and the massive changes in the alanine tRNA mutant, I reasoned that most or all of the 7 nucleotide substitutions would be necessary for retention of the acceptor activity of the mutant tRNA. When each substituted nucleotide in the mutant tRNA was individually changed back to the corresponding wild type nucleotide at that position, the resulting mutant was inactive. Astonishingly, substituting individual nucleotides at the 7 critical sites yielded the identical outcomes for both the selected mutant and wild type alanine tRNA.

The functional properties of selected alanine tRNA provides a second example that recognition cannot be reduced to isolated nucleotide structural elements or functional groups. Clearly, the amino acid acceptor stem as a whole is being recognized. The twist angles, overall pitch and the deformability between base pair adjacent to the location of the G-U pair in the wild type tRNA are key determinants in order for that part of the tRNA to fit into the enzyme active site. Thus, one can understand how a different combination of bases in the wild type and the selected mutant tRNAs can lead to a similar overall topology and function. Our genetic selection system has allowed us to uncover and penetrate this remarkable structure-function complexity in a way that single-mutation analysis does not. The results emphasize the interplay of sequence and structure and how difficult it is to disentangle them in RNA-protein recognition. Our understanding of recognition must be reconfigured to emphasize the role of RNA structure as opposed to RNA sequence or its pendent atomic groups.

I intend to provide a more precise description of the structural recognition hypothesis at the atomic level. What are the roles of water molecules that bridge interactions between the tRNA and aaRS? Are the active multiple mutants and wild type tRNAs able to explore alternative macromolecular conformations on a more frequent time scale relative to other, inactive multiple mutant tRNAs? I intend to determine in utmost detail the molecular and atomic factors that underpin the type of highly novel structural recognition presented above. In this endeavor, I am collaborating with Eric Westhof, a 3-D RNA-protein biochemist with expertise in the Brownian dynamics, time-sampled structural space in RNA-protein complexes, building and modeling RNA 3-D structures. Although he works in France, we readily correspond by e-mail, and confer at conferences. His title and address information is given below.

Professor Eric Westhof
Architecture et Réactivité de l'ARN

Université de Strasbourg
Institut de biologie moléculaire et cellulaire du CNRS,
15 rue René Descartes
F-67084 Strasbourg France
Tel + 33 (0) 3 88 41 70 46
Secrétariat (D. Werling) + 33 (0) 3 88 41 70 56
FAX + 33 (0) 3 88 60 22 18

E-mail: E.Westhof@ibmc.u-strasbg.fr
<http://www-ibmc.u-strasbg.fr/arn/Westhof/>

2. A new mechanism for stimulation of aminoacylation by the tRNA anticodon

A longstanding and perplexing question in protein synthesis is how binding of the tRNA anticodon nucleotides, removed from the 3' acceptor end of the molecule by about 40 Å, profoundly enhances the catalytic rate of amino acid attachment. This key step in protein synthesis contributes to the precision of amino acid attachment by aaRS enzymes, so that only the correct amino acids attach to their cognate 3' acceptor ends of tRNAs. Aminoacylated tRNAs then insert their amino acids at their associated codons in the mRNA during translation; a mistake in attachment results in miss-incorporation. Because it is so crucial to cell viability, multiple mechanisms have evolved to scrutinize the correctness of nascent aminoacyl-tRNAs.

A wealth of biochemical and genetic results, as well as crystal structures of tRNA-aaRS complexes, demonstrates that productive intermolecular interaction between a tRNA and its aaRS involve the anticodon region. Surprisingly, nucleotide substitutions in the anticodon region reduce not so much the affinity of the RNA-protein interaction, but rather the steady-state k_{cat} rate of aminoacylation by two- to four-orders of magnitude, with little impact on the K_m of tRNA binding. Thus, aaRS binding to the correct tRNA anticodon has a very large effect on catalysis occurring at the 3' end of the tRNA, despite the large distance separating the two sites.

A first mechanism to explain the role of the anticodon in aminoacylation efficiency relies on signaling with the enzyme active site residues by conformational communication. More precisely, binding to the anticodon causes a conformational change in the protein (and tRNA, see below) that brings the enzyme's active-site residues into proper alignment to form the catalytically proficient structure. In support of this mechanism, the crystal structure of an *E. coli* glutamine tRNA-GlnRS complex shows a beta-barrel domain straddling these sites that could provide the putative communication route. Consistent with this suggestion, binding to incorrect tRNAs leads to different crystal structures of the enzymatic domain, lending support to the presence of conformational communication between the two sites.

Two independent genetic studies identified a second motif in the glutamine system implicated in rate enhancement. The part of the protein relevant to this second communication pathway is a 26 amino acid motif composed of a strand-alpha-helix-strand that binds to the central nucleotide in the anticodon, the inside corner of the tRNA's L-shape, and abuts the enzyme's catalytic site. Genetic data were separately collected in *E. coli* for glutamine tRNA and GlnRS, but we merge the two sets of data here for clarity. Starting with a glutamine tRNA gene that contained an anticodon or other substitutions that curtailed aminoacylation, mutations that compensated the null phenotype were observed in the genes corresponding to either the 26 amino acid motif in GlnRS or in the sub-domains of glutamine tRNA that interact with segments of this motif; mutations were not observed in the beta-barrel domain. The compensating mutations substantially restored efficient and specific aminoacylation. Thus, *in vivo* genetic data from two independent sources, the enzyme or the tRNA, provides robust support for the involvement in aminoacylation of this motif and its complementary binding patch on the tRNA. It would be incorrect to think that a compensating mutation uncoupled from its initial inactivating mutation would by itself boost aminoacylation activity of the wild type system because the initial mutation plus the compensating mutation work synergistically as a pair. In other words, both mutations are needed to simultaneously for viability, and are therefore highly unlikely to arise as a double mutation without some sort of genetic pressure.

If conformational communication via the 26 amino acid motif were the dominant mechanism by which the information encoded in the anticodon sequence is transmitted to the enzyme active site, one would expect that binding to the anticodon would lead to structural rearrangements in this motif. However, the crystal structure of the free, unbound GlnRS failed to reveal significant structural changes in the 26 amino acid motif relative to that of the glutamine tRNA-GlnRS complex. Thus, either the particular crystallographic snapshots of the complexed GlnRS or that of the free enzyme did not capture the presumed altered structures in the GlnRS motif, or the interaction of the tRNA with this motif primarily serves another purpose. Thus, we are left with a dilemma: while structural studies do not support a role for the 26 amino acid motif or the beta-barrel domain in the enhancement of catalytic activity by the tRNA anticodon, independent genetic analyses of GlnRS as well as tRNA^{Gln} strongly argue for the importance of the 26 amino acid motif in rate enhancement.

In summary, the general model, which posits that anticodon binding is coupled to a conformational reconfiguration of its active site residues into the catalytically proficient pocket, fails to fully explain the very large stimulation of the rate of aminoacylation when the aaRS binds the anticodon and nearby nucleotides in the tRNA substrate. We consider here a new mechanism, as an alternative for, or supplement to, the mechanism of conformational communication for tRNA anticodon stimulation of aminoacylation efficiency.

We propose that enzyme binding to the anticodon and nearby nucleotides anchors the tRNA to the enzyme to ensure that the substrate does not diffuse away during the

myriad of conformational changes that ensue during the catalytic cycle. The two mechanisms are distinct: unlike the first mechanism, this second one ascribes binding to the anticodon and other nucleotides to ensure the substrate remains bound to the enzyme during the conformational changes so aminoacylation can occur.

The conformational changes in the tRNA anticodon region that accompany aaRS binding are substantial. They include the unstacking and splaying out of anticodon nucleotides in a process that is energetically driven by the formation of new direct and indirect non-covalent bonds in partner macromolecules. This process encapsulates and tightly anchors the anticodon segment in a quasistable pocket. This structurally unique anticodon conformation does not exist (or is highly disfavored) in the absence of bound enzyme. Conformational distortion of the 3'-acceptor end of the complexed tRNA can also occur and, as mentioned above, anticodon binding by the aaRS is associated with a conformational reconfiguration of the enzyme to form its catalytically active site pocket. In the *E. coli* glutamine system, the first base pair of the tRNA's acceptor stem is broken by the insertion of a bulky amino acid side chain. Moreover, five nucleotides of the 3' end rotate back toward the anticodon, forming a hairpin loop rather than continuing the A-form helical path observed in free, unbound tRNA.

Clearly, these examples of tRNA conformational change entail large movements accompanied by making and breaking many bonds within and between the partner macromolecules. This process could foster the detachment of the bound tRNA, were it not anchored to the enzyme through kinetically stable interactions between the anticodon and the anticodon-binding domain of the synthetase.

Further arguing for the need for kinetic anchoring is the observation that the conformation of the acceptor end of many tRNAs undergoes pronounced structural changes subsequent to anticodon binding by the aaRS. In some aminoacylation systems, the acylated acceptor end of the tRNA rotates by about 30 Å into an editing site before the aminoacyl-tRNA is released from the enzyme. The purpose of the editing site is to verify the accuracy of the match of the amino acid and the tRNA's acceptor identity. The overall frequency of cytoplasmic mismatched aminoacyl-tRNA is about 10^{-4} . This low frequency calls for a proofreading mechanism subsequent to amino acid attachment to ensure that chemically similar but smaller amino acids are not miss-aminoacylated. Only after the editing site confirms the correctness of the match, the enzyme releases aminoacyl-tRNA to participate in protein synthesis. Conversely, if the pair does not match, the transiently esterified amino acid is hydrolyzed in the editing site and the 3' acceptor end of the deacyl tRNA then relocates back to the initial catalytic site where the aminoacylation and proofreading cycle repeat. If the tRNA were not firmly anchored to the enzyme, the dramatic conformational changes could allow the bound tRNA to diffuse away from the synthetase, resulting in an energetically costly, non-productive cycle.

In summary, we suggest that the mechanism widely believed to be primarily or even exclusively responsible for the communication between the anticodon-binding domain and the synthetase's active site does not fully explain the remarkable accuracy of aminoacylation and how well the identity of the anticodon can control aminoacylation

efficiency. Rather, we suggest that a second major contributor to specificity is the ability of the interaction between anticodon and surrounding nucleotides and the specificity domain of the synthetase to hold on to the cognate tRNA long enough for catalysis to occur. First, these interactions may be required to prevent the tRNA from dissociating during aminoacylation, and to overcome the energetically costly disruption of many intermolecular interactions during the large conformational changes that occur as the tRNA-synthetase complex forms. Second, substrate release may be much more efficient when a non-cognate anticodon is mistakenly recognized.

This work is collaboration with Gabriele Varani whose research interests encompass how proteins and nucleic acids interact with each other. The aim of the group is to understand at the physical chemical level how proteins bind nucleic acids and exploit this knowledge to rationally design new proteins and small molecule drugs that control human regulatory networks or repress viral replication. He uses spectroscopic (NMR), crystallographic, computational and biochemical techniques. His title and address information is given below

Professor Gabriele Varani
University of Washington
Department of Chemistry
Bagley Hall
Campus Box 351700
Seattle, WA 98195, USA
Tel: 206 543 7113
E-mail: varani@chem.washington.edu
<http://depts.washington.edu/varani2/>