Published in final edited form as:

J Mol Biol. 2008 April 11; 377(5): 1382–1405. doi:10.1016/j.jmb.2008.01.073.

Dynamics of Recognition between tRNA and Elongation Factor Tu

John Eargle^{*}, Alexis A. Black[†], Anurag Sethi[†], Leonardo G. Trabuco^{*}, and Zaida Luthey-Schulten^{*,†,§}

^{*}Center for Biophysics and Computational Biology, Urbana, IL

[†]Department of Chemistry University of Illinois, Urbana, IL

Abstract

Elongation factor Tu (EF-Tu) binds to all twenty standard aminoacyl-transfer RNAs (aa-tRNAs) and transports them to the ribosome while protecting the ester linkage between the tRNA and its cognate amino acid. We use molecular dynamics simulations to investigate the dynamics of the EF-Tu-GTP-aa-tRNA^{Cys} complex and the roles played by Mg²⁺ ions and modified nucleosides on the free energy of protein-RNA binding. Individual modified nucleosides have pronounced effects on the structural dynamics of tRNA and the EF-Tu-Cys-tRNA^{Cys} interface. Combined energetic and evolutionary analyses identify the coevolution of residues in EF-Tu and aa-tRNAs at the binding interface. Highly conserved EF-Tu residues are responsible for both attracting aa-tRNAs as well as providing nearby nonbonded repulsive energies which help fine-tune molecular attraction at the binding interface. In addition to the 3' CCA end, highly conserved tRNA nucleotides G1, G52, G53, and U54 contribute significantly to EF-Tu binding energies. Modification of U54 to thymine affects the structure of the tRNA common loop resulting in a change in binding interface contacts. In addition, other nucleotides, conserved within certain tRNA specificities, may be responsible for tuning aa-tRNA binding to EF-Tu. The trend in EF-Tu-Cys-tRNA^{Cys} binding energies observed as the result of mutating the tRNA agrees with experimental observation. We also predict variations in binding free energies upon misacylation of tRNA^{Cys} with D-cysteine or O-phosphoserine and upon changing the protonation state of L-cysteine. Principal components analysis in each case reveals changes in the communication network across the protein-tRNA interface and is the basis for the entropy calculations.

Introduction

With the number of copies in *Escherichia coli* estimated to be ten times the number of ribosomes, elongation factor Tu (EF-Tu in *Bacteria*, EF1A in *Archaea* and *Eucarya*) is one of the most abundant proteins in the cell¹. This GTP-dependent translation enzyme is responsible for protecting the ester linkage between a tRNA and its cognate amino acid *en route* to the ribosome. EF-Tu must bind and transport the aminoacyl-tRNAs (aa-tRNAs), including all isoacceptors, corresponding to the twenty standard amino acids as well as pyrrolysine. A protein homologous to EF-Tu transports the charged tRNA species specific to

© 2010 Published by Elsevier Ltd.

[§]Correspondence to: A544 CLSL, Box 41-6, 600 S. Mathews Ave., Department of Chemistry, University of Illinois at Urbana-Champaign, Urbana, IL 61801, Phone: (217) 333-3518, Fax: (217) 244-3186, schulten@scs.uiuc.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

selenocysteine². The dynamics and energetics involved in this universal process of RNA recognition are the central themes of this study.

Here, we use molecular dynamics (MD) simulations and evolutionary analysis to investigate EF-Tu-aa-tRNA^{Cys} binding under a variety of different conditions: with varying Mg²⁺ and K⁺ ion concentrations, with both modified and unmodified nucleosides, and with different amino acids charged onto tRNA^{Cys}. Our exhaustive study shows that through coevolution, certain features of both EF-Tu and tRNA were developed and maintained for their mutual association. Conservation of key interactions is important to establish a base commonality across all EF-Tu-aa-tRNA pairings. In addition, the long timescales of these MD simulations (20ns) allow us to study the stability of the tRNA and dynamics of the ternary complex under varying conditions. Using the methods of MM-PBSA and PCA, we analyze the binding free energies for the various modifications and mutations suggested by recent experiments.

Key Features of EF-Tu and tRNA

The first crystal structure of an EF-Tu-GTP-aa-tRNA ternary complex to be solved contained Phe-tRNA^{Phe} from yeast, EF-Tu from *Thermus aquaticus*, and GDPNP, a nonhydrolyzable GTP analogue³. In addition to other structures of EF-Tu-GTP-Phe-tRNA^{Phe}, the structure of an all-bacterial ternary complex with *T. aquaticus* EF-Tu, GDPNP, and *E. coli* Cys-tRNA^{Cys} is available (PDB ID 1B23)⁴. EF-Tu is a three-domain protein composed of an N-terminal GTPase domain and two OB-folds (Fig. 1). In this crystal structure, the contact area between any two of the three domains is large (995 Å² between domains 1 and 2, 612 Å² between domains 1 and 3, and 556 Å² between domains 2 and 3). The amino acid binding pocket is between domains 1 and 2. The effector loop connecting the GTP binding pocket to the tRNA acceptor stem is a helix in the EF-Tu GTP form but changes into two beta strands when EF-Tu moves to its GDP form after GTP hydrolysis.

Most of the conserved contacts between tRNA and EF-Tu are formed close to the domain-domain contact regions of EF-Tu (Fig. 2). With regard to tRNA structure terminology: since modified bases dihydrouridine (D) and thymine (T) are absent from most known archaeal tRNAs⁵, we adopt the terms “GG arm” and “common arm” instead of the “D arm” and “T arm” conventionally used in describing bacterial and eukaryotic systems. The tRNA acceptor stem and common arm bind along the interface between domain 1 and the other two domains of EF-Tu and form many contacts with domain 3. The tRNA helix composed of the GG and anticodon arms does not make direct contact with EF-Tu.

Because of its central role in translation, EF-Tu is extremely well conserved across all species as shown in Figure 2. The high degree of conservation in this universally distributed enzyme allowed for its use along with the paralogous EF-G (EF2 in *Archaea* and *Eucarya*) as an outgroup to identify the root of the universal phylogenetic tree which was then applied to the 16S rRNA tree⁶. It is notable that eukaryal and archaeal EF1As have diverged significantly from each other while EF-Tus remain remarkably similar within *Bacteria* (50% sequence identity). Previously, a lack of sequence and structure information had prevented an extensive study of EF-Tu/tRNA coevolution. Here we provide evolutionary analyses of both molecules in order to address questions of the interface design.

The existing ternary complex crystal structures show EF-Tu interacting with tRNA primarily through backbone contacts. If tRNA sequence differences play a role in binding, they do so through sequence-dependent changes to the RNA backbone⁷. Working with transcribed tRNAs, Uhlenbeck and co-workers have shown that the binding affinity of aa-tRNA to EF-Tu is determined by contributions from both the charged amino acid and the tRNA

sequence. Their findings lead them to hypothesize that the tRNA sequences for each specificity have been tuned so that the overall binding affinities for various aa-tRNAs are similar⁸. In their proposed compensation mechanism, tight-binding amino acids are paired with loose-binding tRNAs and *vice versa*. In more recent experiments they showed that binding affinity varies when different tRNA sequences are aminoacylated with the same amino acid^{9,10}. Kinetics studies reveal global effects of changing the tRNA sequence, but the atomic-level mechanisms can be seen only with MD simulations. The binding interface can be better understood in terms of the dynamic atomistic details of the complex. While we use evolutionary analyses to focus primarily on the most conserved regions of the EF-Tu-aa-tRNA^{Cys} binding interface, we also examine a tRNA^{Cys} mutant used to support the tRNA sequence tuning hypothesis. In addition, although structures of the complex are only known for *Bacteria*, our MD studies are complemented by the evolutionary analysis to estimate the universality of our findings. In particular, certain tRNA nucleotides at the binding interface are found to be conserved among some tRNA specificities. They may be conserved for tuning EF-Tu binding for their respective aa-tRNAs.

Role of Modified Bases

Of known RNA molecules, tRNA has the highest concentration of modified nucleosides, and it is well known that these bases play important roles in the decoding at the ribosome and recognition by the aa-tRNA synthetase (AARS). Their role, if any, in tRNA binding to EF-Tu has not been previously considered, and such an investigation is a major component of our computational study of the dynamics of tRNA^{Cys}. Aside from the backbone phosphate of T54, none of the modified nucleotides sit at the interface to EF-Tu.

An excellent review of the role of modified bases in RNA folding is given by Helm¹¹. Cys-tRNA^{Cys} from *E. coli* has eight modified bases of five different types: 4-thiouridine (4su), dihydrouridine, pseudouridine (Ψ), 2-methylthio-N⁶-isopentenyladenosine (Mia), and thymine. 4su at position 8, a bacterial modification, has been shown to delay growth under near-ultraviolet illumination in *E. coli*¹². Due to lack of aromaticity, D bases at positions 20 and 21, typically found in the tRNA GG loop of *Bacteria* and *Eucarya*, have greater conformational flexibility ensuring that these residues are not part of the GG stem and exhibit less structural stability when in contact with Mg²⁺ ions¹³. In previous studies, Ψ has been shown to stabilize the local structure of RNA through interactions with a bridging water between the base and RNA backbone phosphates^{14,15}. T54 and Ψ55 in the common loop provide increased structural stability through tighter base stacking and tertiary contacts to the GG loop^{16,15}. By far the largest base in this system, hypermodified Mia37 has been shown through NMR studies to be important for opening the anticodon loop since it disrupts the base pairing that would occur between unmodified A37 and U33¹⁷.

Recognition of Nonstandard aa-tRNAs

The traditional role of EF-Tu involves the transport of all twenty standard aa-tRNAs to the ribosome, but redundancy in the genetic code can be coopted for the addition of new amino acids. The genetic code has been expanded *in vivo* using the translational system, including EF-Tu, along with a strategy similar to that used by a few organisms in the *Methanosarcinales* genus to encode the amber codon as pyrrolysine, the twenty-first amino acid¹⁸. While EF-Tu can bind tRNAs aminoacylated with non-natural amino acids, it remains an open question whether EF-Tu can distinguish between the cognate aa-tRNA pairing and tRNAs that are misacylated with cellular metabolites or amino acid precursors. It has been shown that Asp-tRNA^{Asn} and Glu-tRNA^{Gln} do not successfully deliver their amino acids to a growing polypeptide at the ribosome^{19,20}, suggesting that discrimination is possible, but this could be related to the extra negative charge on these particular misacylations. Investigations into the nucleotides and amino acids responsible for the

discrimination between Asp-tRNA^{Asn} and Asn-tRNA^{Asn} in bacteria have recently been performed²¹. An example of misacylation with a D-amino acid exists in *Bacteria* and *Eucarya*. Tyrosyl tRNA-synthetase misacylates tRNA^{Tyr} with D-Tyr and has evolved a freestanding editing domain to hydrolyze the ester bond between D-Tyr and tRNA^{Tyr}²². Here we test whether EF-Tu is able to distinguish between L- and D-amino acids on the tRNA by comparing the dynamics of EF-Tu bound to D-Cys-tRNA^{Cys} with those of EF-Tu bound to L-Cys-tRNA^{Cys}.

In addition to misacylation of tRNAs with metabolites, in some organisms tRNAs are deliberately aminoacylated with amino acid precursors and then converted to the cognate aa-tRNA pairing in a tRNA-dependent fashion. For example, in all methanogens, tRNA^{Cys} can be charged through an indirect mechanism in which O-phosphoserine (Sep) is first loaded onto tRNA^{Cys}, and then Sep-tRNA^{Cys} is converted into Cys-tRNA^{Cys} by a second enzyme^{23,24}. A phylogenetic study using a combination of sequence and structure data has shown that both the direct and indirect mechanisms of cysteine aminoacylation are ancient and were present at the root of the universal phylogenetic tree²⁴. The question now arises whether Sep-tRNA^{Cys} is capable of loading Sep onto a growing peptide chain at the ribosome, either in natural or engineered systems. Since Sep could be a possible target for future protein engineering efforts, we look at the binding energies of Sep-tRNA^{Cys} to EF-Tu in bacteria.

Molecular Dynamics Simulations

While MD methods have been used to study GTPase proteins and tRNA separately^{25,26,27}, EF-Tu has not previously been simulated. The significant improvements in the application of MD to RNA have recently been reviewed^{28,29,30}. MD simulations of tRNA have been used to study the local dynamics of tRNA^{Asp} base pairs and base triples³¹ and patterns of hydration in a tRNA^{Asp} anticodon hairpin with the modified bases Ψ and 1-methylguanine included¹⁴. These simulations have been relatively short (<1ns) and did not include Mg²⁺ ions. The presence and concentration of monovalent and divalent cations have also been shown to play a role in the stability of tRNA. By shielding the negatively charged RNA backbone phosphates from each other, cations allow the nucleic acid backbone to form double helices and other more complicated, compact structures. Mg²⁺ ions in particular have been shown to stabilize overall tRNA tertiary structure³² although sometimes they can cause local destabilization¹³. Identifying Mg²⁺ binding sites is nontrivial, but recent crystal structures have shown several sites that are in agreement with Mg²⁺ placement procedures used in this work and others³³. In this study, we examine the effects of different Mg²⁺ concentrations on the dynamics of tRNA in the presence and absence of nucleoside modifications.

MM-PBSA (Molecular Mechanics-Poisson Boltzmann Surface Area) is a method developed to calculate trends in binding free energies from MD trajectories^{34,35,36,37}. It has previously been used to calculate free energies of binding for protein-RNA systems³⁸ as well as systems containing Mg²⁺³⁹. MM-PBSA together with entropy estimates from principal components analysis is used here to determine binding free energies for different EF-Tu-aa-tRNA^{Cys} systems.

Results and Discussion

Evolutionary Analyses of EF-Tu and tRNA

The evolutionary analyses of the sequences and structures of both EF-Tu and tRNA were performed with tools in the MultiSeq plugin⁴⁰ to VMD⁴¹ (see Methods). Evolutionary profiles (EPs) containing representatives from all three domains of life were generated using

the QR algorithm⁴² to obtain the most linearly independent, statistically well balanced sets of sequences that span the evolutionary phase spaces for EF-Tu and tRNA⁴³. In addition, a separate bacterial EP was created in order to determine the *Bacteria*-specific elements involved in tRNA recognition.

EF-Tu Interface—EF-Tu follows the canonical phylogenetic pattern with the root of the tree occurring inbetween the bacterial group and the distinct groups of *Archaea* and *Eucarya*. The protein also contains insertions specific to each domain of life. Despite these differences, the GTP-binding pocket and surrounding area, which contact the ribosomal sarcin/ricin loop, and most of the tRNA binding interface are well conserved with large areas of the protein surface conserved across the three domains of life as shown in Figure 2.

On EF-Tu, the tRNA binding interface can be divided into 6 regions: Regions I to IV (residues 50 to 57 and 85 to 92 from domain 1, and residues 224 to 240 and 271 to 302 from domain 2) interact with the acceptor arm or the amino acid charged onto the tRNA, while regions V and VI (residues 325 to 350 and residues 375 to 392 from domain 3) interact with the tRNA common arm. These regions are comprised of a number of highly conserved, acidic or basic residues which interact with the negative tRNA backbone phosphates. A76 and the cysteine attached to its ribose 3'-hydroxyl interact with a highly conserved pocket formed by residues Asp227, Glu271, and Asn285. The conserved residues Lys90, Asn91, and Arg300 are part of a small groove that specifically interacts with the base of G1. Asn91 also contacts the backbone of G2. The common arm forms contacts with the conserved residue Tyr338 (G52) on domain 3. The Lys376 sidechain stretches out from the bottom of domain 3 to make contact with the C67 phosphate, and Arg389, while not directly contacting tRNA, is close enough to electrostatically attract the nucleic acid backbone. There are conserved acidic residues such as Glu56, Asp87, and Asp227 present at the interface to ensure that the binding affinity of aa-tRNA to the protein is repulsively balanced so that undocking of aa-tRNA from EF-Tu-GDP at the ribosome occurs at a reasonable rate. The rate of undocking of the aa-tRNA on the ribosome is 100 times per second⁴⁴. Glu390, which hydrogen bonds with the G63 exocyclic amine, is conserved across *Bacteria*, but since it is replaced by similar amino acid Asp in *Archaea* and *Eucarya*, we consider it to be universally conserved. In *T. aquaticus*, Lys52, which remains hydrophilic in almost all organisms, and Lys376, which is conserved in *Eucarya* and *Archaea* but not in most bacterial sequences, interact with C74 and C67, respectively.

In addition to the universally conserved recognition elements on EF-Tu, there are bacterial domain specific recognition elements in contact with tRNA. The conserved bacterial insert corresponding to residues 327 to 331 in *T. aquaticus* interacts with tRNA^{Cys}. In *Bacteria*, conserved EF-Tu interface residues Glu55 (G1), Arg330 (G52), Arg339 (C51), Asp348 (A66), and Glu390 (G63) play an important role in the energetic recognition of the charged tRNA species (see Local Nonbonded Interaction Energies). Residues Tyr88 (C3), Arg274 (C76), His331 (G53), Gln341 (A65), Thr350 (A65), and Gly391 (G64) are also highly conserved but participate in weaker, hydrogen bonding or hydrophobic interactions. Lys325 is also conserved in *Bacteria* and may play a role in interacting with T54 even though they are about 10 Å from each other in the crystal structure.

The GTP-binding pocket of EF-Tu is composed of residues 18 to 26, 60 to 62, 81 to 85, 136 to 139, and 173 to 177 which are also conserved in all three phylogenetic domains. Aside from GTP binding, a number of these residues are conserved because they interact with the sarcin/ricin loop on the large subunit of the ribosome during the accommodation of the new aa-tRNA into the ribosomal A-site. The interdomain contacts, including some residues from region III, also remain highly conserved. These residues and conserved residues on the effector loop could be involved in allosteric communication between the GTP-binding

pocket and the tRNA-binding interface due to their close proximity to the tRNA binding regions.

tRNA Interface—There are several highly conserved nucleotides on the acceptor and common arms of tRNAs that participate in binding and interact with highly conserved residues on EF-Tu. The CCA hairpin is universally conserved in all tRNA molecules and interacts with a highly conserved Asp227, Glu271, and bacterially conserved Arg274 on the EF-Tu. G1 has at least 75% sequence identity among all isoacceptors within 17 specificities in *Archaea*, 15 in *Eucarya*, and 13 in *Bacteria*. In most cases where a mutation occurs at position 1, a correlated mutation is also observed at position 72, maintaining the Watson Crick base pair. The 5' phosphate of G1 in tRNA^{Cys} interacts with a conserved pocket in EF-Tu. The ribose and phosphate of G1 interact with the universally conserved Lys90 and Asn91, and the highly conserved Arg300. Lys90 and Asn91 are part of a helix in the EF-Tu GTPase domain which forms at the interface of all three protein domains. In addition, effector loop residue Glu55, which is conserved in *Bacteria* and *Archaea*, forms a hydrogen bond with the amide functional group on G1 and is the only interaction that is base specific in this region.

On the common arm, G52, G53, and the modified U54 (T54 in *Bacteria* and *Eucarya* and 1-methyl-Ψ54 in *Archaea*) have strong conservation across all tRNA specificities in all three domains of life, however, G52 has only 65% sequence identity across *Bacteria*^{45,46}. Conservation of nucleoside modifications is less clear than conservation of underlying tRNA gene sequence due to scarcity of available modified tRNA sequences, especially in *Archaea*. Conserved residues in EF-Tu interact with nucleotides 52–54, and most of these interactions are hydrogen bonds to the backbone of the common arm. A number of bases in the common arm (including the universally conserved bases 55, 56, and 58) are also highly conserved in order to maintain the three-dimensional structure of tRNA by forming tertiary interactions with bases on the loop of the GG arm. In addition, the conserved base 54 pairs with 58 to form a conserved base pair in the core of all tRNA molecules. The unconserved base 57 stacks with universal G18 and G19. This region is highly conserved in structure in all tRNA molecules, and EF-Tu interacts with nearby nucleotides. The modifications in this region of the tRNA differ between different domains of life (and sometimes specificities), but the overall effect of these modifications is to stabilize the three-dimensional arrangement of the tRNA. When the transcribed tRNA molecule is present, this region is floppier causing a slight shift in its contacts with EF-Tu (see Local Nonbonded Interaction Energies).

The remainder of the interface on the tRNA side (bases 63–67) is variable, even across bacterial tRNA^{Cys}. This is not surprising given the broad range of molecules to which EF-Tu has evolved to bind⁹. The features common to all tRNA specificities exist at both extremes of the tRNA binding surface: G1 and the CCA end at one side and the common arm nucleotides 52–54 at the other (Fig. 1, 2). The only base in the common arm that engages in significant interactions with EF-Tu is the unconserved G63, which hydrogen bonds with the conserved Glu390/Asp390.

The interface region between EF-Tu and tRNA^{Cys} is comprised of a number of conserved residues. Of the strongly interacting interface residues, more than 85% are conserved in *Bacteria* and 50% over all three domains of life. Given the coevolutionary nature of this interface, it is not surprising that these conserved or coevolving residues dominate the interaction energetics of binding as seen in Figure 5 and Figure 6.

Is the binding affinity of tRNA tuned?—It has been proposed that tRNA sequences are tuned so that the binding affinity of the aa-tRNA remains the same irrespective of the charged amino acid⁸. Since the EF-Tu side of the binding interface is so well conserved, if

such tuning were present, the nucleotides at some of the positions on the binding interface might be correlated to the amino acid charged on the tRNA. The entropy analysis of the full tRNA evolutionary profile (Fig. 3a) shows that a number of bases in tRNA are completely conserved, near zero entropy (see Methods). Of these bases, the CCA hairpin (not shown), G53, and U54 interact with the EF-Tu. In addition to these bases, there are a few bases that are more conserved in tRNAs belonging to a given specificity. These bases may play a role in tuning the binding affinity of aa-tRNA to EF-Tu and, possibly, the ribosome.

The aa-tRNAs interact with the ribosome and the aa-tRNA synthetases (AARSs) in addition to interacting with EF-Tu. Each tRNA is unique in that it contains nucleotides which are intrinsic to its amino acid specificity. These nucleotides are known as identity elements for the AARSs. The elements usually consist of the anticodon at positions 34, 35, and 36 and base 73, also called the discriminator base. U73 is almost completely conserved across all tRNA^{Cys} because it is necessary for recognition by the cysteinyl-tRNA synthetase⁴⁷. Differences in sequence entropy indicate residues specific to that tRNA (Fig. 3b). Since the discriminator base covaries with tRNA specificity, it could also play a role in tuning the affinity of aa-tRNA with EF-Tu. In addition, chemically similar amino acids often have the same discriminator base, indicating that this base may allow detuning between similar amino acids charged on the same tRNA species. The discriminator base shows the highest correlation to amino acid specificity and is conserved in eleven specificities: A for tRNA^{Ala}, tRNA^{Ile}, tRNA^{Leu}, tRNA^{Met}, tRNA^{Phe}, tRNA^{Tyr}, and tRNA^{Val}; U for tRNA^{Cys}; G for tRNA^{Asn} and tRNA^{Asp}; and A or G (purine) for tRNA^{Ser}.

Base pairs on the acceptor stem that interact with EF-Tu also show high correlation to amino acid specificity. Base pair 2·71 is conserved in seven specificities: G·C or C·G in tRNA^{Leu}, and tRNA^{Met}; G·C in tRNA^{Ala}, tRNA^{Gln}, and tRNA^{Pro}; and C·G in tRNA^{Gly} (Fig. 3c), and tRNA^{Thr}. Base pair 3·70 is also conserved in seven specificities: C·G or U·A in tRNA^{Asn}, tRNA^{Asp}, and tRNA^{Thr}; G·C or A·U in tRNA^{Gly}, tRNA^{Pro}, and tRNA^{Trp}; and G·U in tRNA^{Ala}. Finally, the common arm base pair 51·63 is highly conserved in tRNA^{Gly} (G·C) (Fig. 3c), tRNA^{Ser} (G·C or A·U), tRNA^{Asp} (G·C), and tRNA^{Glu} (G·C) and shows higher correlation than the average. In general, the stronger binding tRNAs⁴⁸ have a larger number of conserved bases at the binding interface indicating that these base positions could be tuning the binding affinity of the aa-tRNA to EF-Tu.

Molecular Dynamics Analysis

To investigate the dynamics of the ternary complex, especially at the protein-RNA interface, eleven separate MD simulations were run under various ionic conditions and states of modification. For reference, each simulation is named in Table 1 and ion conditions are shown in Table 2. The first is a control run using the 1B23 crystal structure including the three crystal Mg²⁺ ions. Six trajectories were computed with the system at different Mg²⁺ concentrations and with modified or unmodified bases. Two simulations examined tRNA^{Cys} misacylated with D-Cys and Sep amino acids, and the last two are different variations of the MMC system. Each simulation was run for 20ns.

Structural Dynamics

In these simulations, EF-Tu is very stable with an RMSD of 1.4–1.6 Å from the crystal structure (data not shown). Aside from fluctuation in the small, floppy effector loop, there is little protein motion in any of the simulations. This is expected because the three domains interact extensively with each other while EF-Tu is in the compact GTP form.

In systems with lower Mg²⁺ concentrations (control, LUC, LMC) there is generally more structural motion in the tRNA. It is especially pronounced in the anticodon arm, which lacks

coordination with Mg^{2+} ions that are needed to stabilize the double helix by counteracting the repulsion of backbone phosphates on either side of the deep groove. In setting up the systems at low Mg^{2+} concentration, no Mg^{2+} ions were placed in the anticodon stem because other regions were more electrostatically favorable. While K^+ filled the role of charge compensation, its lower charge relative to Mg^{2+} and the transient nature of its primary solvation shell allowed the anticodon to unravel as shown in Figure 4. In the control run, which had only the two crystallographic Mg^{2+} ions associated with Cys-tRNA^{Cys}, the anticodon arm uncoiled, and all of the base pairs in the anticodon stem eventually broke apart. The uncoiling motions are similar to those observed in the low frequency modes of the complex calculated using the anisotropic network model⁴⁹ in which the anticodon arm moves relative to the common arm (data not shown).

While Mg^{2+} ions help stabilize the tRNA molecule, they also have a powerful local effect on backbone shape and can perturb regions where they have access to two or more phosphates. An abundance of Mg^{2+} , while holding the sides of the tRNA deep groove together, can produce kinks in the backbone where phosphates occupy the Mg^{2+} secondary solvation shell.

As expected, throughout all of the simulations the coaxially stacked acceptor and common stems are very stable since they participate in extensive contacts with EF-Tu as well as harbor multiple Mg^{2+} binding sites. At higher Mg^{2+} concentrations, large-scale motion of the coaxially stacked helices with respect to each other gives rise to deviations in overall tRNA backbone RMSD. Local perturbations of the GG and anticodon arms are caused by interactions with specific Mg^{2+} ions. In HUC, the drastic effect of a single Mg^{2+} on bases in the anticodon loop may be due to the absence of the modified base Mia37. In HMC, Mia37 orients its large hydrophobic isopentenyl group towards the loop interior and does not interact directly with any of the freely diffusing hexahydrated Mg^{2+} ions. The unmodified A37 in HUC exposes an extra hydrogen bonding partner to the loop interior which interacts with nearby Mg^{2+} . Ions can enter the anticodon loop, disrupting the U-turn motif involving U33 and A36. As a result, U33 intermittently hydrogen bonds to A37 before forming a Watson Crick pair with A38 after 15ns. The appearance of this base pair shows the importance of Mia37 for maintaining the 7-nucleotide anticodon loop, in agreement with NMR experiments¹⁷. As shown by the NMR data, unmodified A37 produces conformational flexibility in the anticodon arm allowing extra base pairs to form, and in the most extreme variation, base pairs 32·38 and 33·37 reduce the anticodon loop to the three anticodon nucleotides. Trajectories longer than 20 ns may discover more base pairing interactions in the anticodon loops of tRNAs with unmodified bases. Also, in LUC, MUC, and HUC, hydrogen bonds form between U32 and A38, but these interactions do not appear in the modified systems because the Ψ 32 base is flipped with respect to U32.

Relative to unmodified U, modified bases 4su and D do not appear to have a pronounced effect on the structure and dynamics of the aa-tRNA^{Cys}. In agreement with previous MD simulations of Ψ in the tRNA anticodon loop¹⁴, Ψ is seen to coordinate a highly resident water molecule between O4 and a phosphate oxygen on the same nucleotide. In preliminary simulations using the standard Ψ tautomer, the O4 of Ψ 32 quickly made direct contact with the phosphate of A31. Once formed, this contact remained stable and resulted in the removal of Ψ 32 from base stacking contacts with its neighbors in the anticodon loop. This base stacking disruption then led to the eventual unraveling of the anticodon arm even in high Mg^{2+} simulations. Previous, 500-ps simulations did not show this interaction¹⁴, but in our simulations the contact formed after 500 ps. Correct tautomer determination is especially important for MD simulations in which nucleotides move significantly from their original positions and contact partners. In our modified systems, stable base stacking was achieved by switching Ψ 32 to a different tautomer in which a proton is moved from N3 to O4 where it

hydrogen bonds with the ribose 4'-oxygen (See Methods). It has been suggested previously that U may take different tautomeric states depending on its biophysical context⁵⁰. While the energy states of isolated tautomers have been determined for the standard bases, similar calculations for the different tautomers of modified bases in various environments still need to be performed. As both Ψ39 and Ψ55 maintained their initial contacts during the preliminary simulations, they were kept in the standard tautomer state. Unlike Ψ32 which is part of a loop, Ψ39 is base paired to A31, and Ψ55 participates in a U-turn with A58. T54 and Ψ55 help stabilize the common loop which leads to a change in the binding interface with tighter binding to T54 and looser binding to C51.

Local Nonbonded Interaction Energies

The local nonbonded interaction energies between EF-Tu and Cys-tRNA^{Cys} in each of the simulations were calculated to ascertain the molecular details of their recognition and to determine whether modified bases and ionic concentrations affect the interface contacts. VdW and Coulombic energies were calculated with a cutoff of 21 Å to focus on the binding interface. The cutoff allowed both sides of the acceptor stem and the common arm to be completely included. Since both molecules have a net negative charge, increasing the cutoff distance results in repulsive electrostatic energies as seen in the complete free energy analysis (see MM-PBSA), but trends in the importance of specific contacts remain the same.

As seen in Figure 5a there are almost as many significant repulsive as attractive interactions at the interface between the aa-tRNA^{Cys} and EF-Tu though the attractive interactions dominate. Attraction between EF-Tu and aa-tRNA^{Cys} is mediated almost entirely by positively charged amino acids making direct contact with negatively charged phosphates along the tRNA backbone. Repulsion is accomplished by negatively charged EF-Tu residues positioned towards the tRNA backbone, but the interactions are more diffuse, less specific, and less powerful than the attractive contacts. With the exception of Asp227, all residues responsible for repulsive interactions occur near extremely attractive residues providing a counteraction that fine-tunes the binding between the two molecules. Without such “negative” design, the balance between docking and undocking would be skewed towards the former, requiring a more drastic release mechanism upon GTP hydrolysis.

As previously discussed, many amino acids at the EF-Tu-aa-tRNA interface are well conserved across all three domains of life (See Evolutionary Analyses of EF-Tu and tRNA). Aside from Lys52 and Lys376, residues labeled in Figure 5a are conserved across *Bacteria*. It is notable that while there are significant contributions to the interface energy in the system studied, only about half are strongly conserved across all three domains of life (Fig. 5b). Glu390 conservation has been added as a dashed line to Glu390/Asp390 because in *Archaea* and *Eucarya* the amino acid changes to Asp so the contacts and interaction energy should be conserved. Although Glu390 forms a hydrogen bond with the exocyclic base amine of G63, the contact appears repulsive because Glu390 is being held in the shallow groove between nearby phosphates. With a shorter electrostatic cutoff distance, the attractive energy of this and other hydrogen bonds would be apparent. This nonbonded energy analysis emphasizes the strong electrostatic interactions, but there are previously mentioned, conserved, uncharged amino acids at the interface forming hydrogen bonds or hydrophobic contacts with tRNA^{Cys}. In some cases, *Eucarya* and *Archaea* have residues in the corresponding regions that could be interacting with the tRNA in a different fashion. For example, Arg339 is often replaced by Ala, Ser, and Thr in *Archaea* and *Eucarya*, and these residues could interact with phosphates in the tRNA backbone.

Figure 6a shows tRNA^{Cys} nucleotides with strong nonbonded interactions to EF-Tu. The EF-Tu binding interface is highly conserved at the two extremes of the bound double helix: G1-C72 and the CCA end on the acceptor arm, and nucleotides 52 to 54 on the common arm

(Fig. 6b). G1 makes contact with Lys90 and Arg300, and Asp227 sits between G1 and the CCA end, but there are no other strong interactions between conserved elements of each molecule. Nucleotides 63–67 are highly variable even across bacterial tRNA^{Cys}, so this attractive region mostly disappears when the interaction energy data is scaled by sequence identity. As discussed previously, nucleotides in the base pair 51-63 are conserved in some specificities. Both appear in Figure 6a, but the proximity of bacterially conserved Arg339 to the phosphate of C51 creates a much stronger attraction than that between EF-Tu and G63.

There is a difference in interface contacts between runs with modified or unmodified bases due to changes in interactions of C51 and T54 with domain 3 of EF-Tu. In the common loop, bases T54 and Ψ55 show less motion than the unmodified U54 and U55 because of steric hindrance provided by the T54 methyl group and a resident water bridging the Ψ55 base to its phosphate and/or the T54 phosphate. The additional stability of the common loop with modified bases results in a tighter interaction between the T54 phosphate and His331. As a result, Arg339 associates less frequently with the ribose 2'-hydroxyls of C50 and C51. With unmodified bases, the common arm rocks away from His331 and towards Arg339. The different contacts with unmodified tRNA produce an overall less attractive nonbonded interaction energy at the interface.

A comparison of interface residue conformations revealed that although the RMSD of the protein backbone is small over the entire protein, the sidechains of Glu55 and His331 did not maintain their original contacts to the tRNA. Glu55 moved away from the G1 exocyclic amine and became completely hydrated, and His331 flipped its ring to form a hydrogen bond with the phosphate of T54. The motion of these sidechains indicates a possible discrepancy between the local electrostatic environment in the simulations and that in the crystal structure. A variation of the MMC simulation was run in which the Glu55 sidechain was protonated and the His331 proton was moved from the δ to the ε nitrogen. The tautomeric state of neutral His331 does not appreciably affect the longterm interface interactions. Neutralization of Glu55 causes a number of nearby contacts present in the crystal structure to be lost suggesting that the conventional protonation state is correct. The close association of Glu55 to G1.C72 in the crystal structure may be an artifact of high salt conditions screening repulsive interactions between Glu55 and the tRNA backbone phosphates. Motion near Glu55 may also be due to steric clashes between Thr232 and C75 in the crystal structure which were resolved upon minimization.

The proposed peptide transferase reaction on the ribosome requires the de-protonation of the amine on the amino acid⁵¹. To examine whether the amine would be deprotonated in the ternary complex, we ran simulation MMC_C in which the amine on the cysteine of Cys-tRNA^{Cys} was neutralized. In the other runs, cysteine accounted for approximately 20 % of the local nonbonded interaction energy, but the neutral cysteine accounted for only 4% of this binding energy indicating that the protonated form of cysteine is more stable inside the EF-Tu (data not shown).

Misacylation with O-phosphoserine and D-cysteine

To determine the degree to which EF-Tu participates in proofreading and would allow the incorporation of nonstandard amino acids into nascent protein chains, simulations with Sep (LUS) and D-Cys (MMD) were run. Sep-tRNA synthetase is known to load O-phosphoserine onto tRNA^{Cys} in many methanogenic *Archaea*. This enzyme, along with EF-Tu, could be used to accommodate Sep-tRNA^{Cys} into the ribosome so that phosphorylated serine is added to signalling proteins. D-amino acids could also possibly be incorporated into proteins through protein engineering. This would allow the creation of novel protein structures, like left-handed α-helices, because of the opposite chirality of D-amino acids.

In the system comparable to the wild type ternary complex with L-Cys-tRNA^{Cys}, MMC shown in Figure 7a, Glu226 attracts the L-Cys sulfur while the sidechain of Asn285 and the backbone carboxyl groups of Met272, His273, Asn285, and Val286 attract the positively charged L-Cys backbone amine. Arg274, near the L-Cys backbone carboxyl, provides by far the largest repulsive interaction due to its positive charge and proximity to the backbone amine. L-Cys is also attracted to other, more distant, negatively charged residues: Glu69, Asp227, Glu271, and Asp284. While all of these residues except His273 are conserved across *Bacteria*, most are universally conserved across all three domains of life: Asp227, Met272, Glu271, Asp284, Asn285, and Val286 (Ile in *Archaea*). His273 is conserved in *Archaea* and *Eucarya* but is typically Phe in *Bacteria*.

When tRNA^{Cys} is misacylated with D-Cys, most of the interaction partners of the amino acid with EF-Tu remain the same (Fig. 7b). The change in stereochemistry shifts some of the contacts so that the specific interactions are different: the thiol moves closer to Glu226, Arg274 makes contact with the thiol which pulls Arg274 closer to the backbone amine, and the backbone amine is positioned away from His273. The largest energy changes are a stronger attractive interaction of Glu226 with the D-Cys backbone amine, a stronger repulsive interaction between Arg274 and the backbone amine, and the disappearance of interaction with His273. At the beginning of the 20-ns trajectory, contacts with Met272 are also absent, but after 7 ns of simulation the backbone amine of D-Cys is in a favorable position to form a hydrogen bond with the backbone of residue Met272. Through hydrogen bonds to the D-Cys thiol, bacterially conserved His67 contributes a new attractive interaction although it is weaker than the others mentioned. While His67 hydrogen bonds with Glu226 in both L-Cys and D-Cys systems, the His ring is oriented with a face toward L-Cys so no hydrogen bonds form. In addition, the sidechain of Arg274 forms a hydrogen bond with the ribose of A76. Because of the D-Cys sidechain conformation, Arg274 also lies across D-Cys to associate much more closely with the sulfur. This close contact results in the higher electrostatic repulsion noted above as well as more attractive VdW interactions. The overall structural effect is that the amino acid binding pocket is more compact about D-Cys. The conformational difference in EF-Tu caused by binding D-Cys also changes the interactions with the tRNA CCA end, allowing tighter contacts between tRNA^{Cys} and EF-Tu. This, in turn, helps stabilize the entire binding interface resulting in more attractive VdW and nonpolar solvation energies for MMD than for MMC (Table 4, see MM-PBSA).

With Sep, shown in Figure 7c, the difference in local contacts is much more drastic. The only similar interaction partner is with Asn285, which attracts the Sep backbone amine although less strongly than it does with L-Cys. Because of the -2 charge on the Sep phosphate, Glu226 and Asp227 provide repulsive contacts, as opposed to their attractive roles with L-Cys. As with D-Cys, Arg274 interacts with the Sep sidechain as well as the A76 phosphate, but the negative charge on the Sep phosphate overwhelms any repulsion from its backbone amine. As a result, Arg274 has the most attractive interaction with Sep. Universally conserved Asp51 and Asp81 provide repulsive interactions despite not making direct contact with Sep. Lys45 and Lys52, neither of which is highly conserved across *Bacteria*, are also strongly attractive although they do not contact Sep directly. With regard to structural changes in the amino acid binding pocket, Arg274 closes the pocket as it does with D-Cys, but Glu226 is repelled by the negative Sep causing the side of the binding pocket to open slightly. This is reflected in somewhat more attractive VdW energies for LUS than for MMC (Table 4).

Mg²⁺ and K⁺ Residence Times and Probabilities

As the ternary complex has an overall charge of -87 , it attracts a cloud of counterions. The cation cloud is not uniform because the shape of tRNA concentrates negative charge and the

associated counterions in the junction of the common and GG arms. Most ions diffuse about fairly rapidly while a few remain in specific positions for significant timespans (Fig. 8a and b). These resident ions affect the structure and electrostatic field of the ternary complex, especially the aa-tRNA^{Cys}.

An ion residency site was defined as a location from which a specific ion moved no further than 3 Å for a continuous timespan of at least 2 ns. In the case of Mg²⁺, most of the sites containing highly resident ions correspond to locations of crystallographic Mg²⁺ either in this system or seen in other tRNA crystal structures. In the system setup, the electrostatic potential placed hydrated counterions near these positions, and they move only less than 3 Å from them during the simulation. Residency sites for more mobile Mg²⁺ ions are located in the common loop as well as the deep grooves of the tRNA^{Cys}. K⁺ atoms are found near the GG loop, in the common stem, associated with the phosphates on the bound GTP, and near the effector loop (Fig. 8b). K⁺ ions also have a coordination number of 6, but water exchange happens quickly enough ($4 \cdot 10^{12} \text{ s}^{-1}$ as opposed to $2 \cdot 10^6 \text{ s}^{-1}$ for Mg²⁺⁵²) that K⁺ and its first solvation shell do not hold together and diffuse as one unit like hexahydrated Mg²⁺ does. As a result, the residency time of K⁺ tends to be less than that of Mg²⁺ in the same region.

Although K⁺ ions only have long residency times under the effector loop and near the GTP, the ion probability density or occupancy maps reveal the presence of both K⁺ and Mg²⁺ throughout the tRNA deep groove (data not shown). When the ratio of K⁺ to Mg²⁺ changes, there is a corresponding shift of the ions in the deep groove occupancy sites, especially in the acceptor stem and the anticodon arm. At low Mg²⁺ concentration, K⁺ ions replace Mg²⁺ ions almost entirely in the acceptor stem and anticodon arm (Fig. 8c), while the opposite is true in systems with high Mg²⁺ concentration (Fig. 8d). There are relatively subtle changes in ion occupancy patterns between systems with modified vs. unmodified bases. In all systems there are 5 % probability density regions for Mg²⁺ ions near the GG loop D20/U20 and D21/U21 as well as the common loop T54/U54 and Ψ55/U55. There are no Mg²⁺ 5% probability densities in the anticodon loops of LUC and LMC (Fig. 8c). Mg²⁺ densities are present in the anticodon arm for systems with medium and high Mg²⁺ concentrations (Fig. 8d), but they are closer to A37 in MUC and HUC than they are to the corresponding Mia37 in MMC and HMC. A37 interacts more with Mg²⁺ ions than Mia37 as discussed in the Structural Dynamics section. Crystal structures of tRNA bound to spermine (PDB IDs 1EVV, 1TN2, 2TRA) show an overlap in position between spermine and resident Mg²⁺ in the common loop as well as general agreement between the three spermine binding sites and ion occupancy sites.

Electrostatic potential is an important long-distance molecular recognition mechanism used to aid the docking of complexes such as EF-Tu-GTP-aa-tRNA. Previously, the $-1k_B T/e$ isosurface obtained from a crystal structure was shown to entirely envelop the tRNA due to the extreme electronegativity of the molecule⁵³. However, this picture ignored the contribution of highly resident positive ions to the local charge density around tRNA. As expected, the electrostatic potential map of tRNA^{Cys} is predominantly negative, but the deep grooves harbor positive regions where Mg²⁺ and K⁺ ions interact with the negative phosphate charges (Fig. 9). Time averaging of the system with the associated cations produces isosurfaces with more convolutions and structure than those produced from a single frame. With ions included, a small, positive region appears between the 5' and 3' ends of the tRNA. This area docks to the negative region on the EF-Tu where Glu226 and Asp227 are located. Thus, the electrostatic potential provides a long-distance discrimination mechanism for the correct aa-tRNA orientation. Specifically, the bacterial conservation of both Glu226 and Asp227 may be necessary to insure that the CCA end docks to the left of the positive region and does not insert further into the EF-Tu structure.

MM-PBSA

Average binding free energies for many of the ternary complex simulations were calculated using the MM-PBSA method over the last 5 ns of the trajectories. A detailed breakdown of contributions to binding energy for MMC can be found in Table 3. MMC is taken as a reference because its modified bases and Mg^{2+} concentration set this system closest to physiological conditions⁵⁴. The MM-PBSA binding free energies provide a useful link to experimental results and help determine the relative importance of the various physical terms included in the final free energy values. VdW and electrostatics terms are nonbonded interaction energies calculated easily from the CHARMM27 force field. The polar solvation term, analogous to the Born solvation energy of an ion, is the free energy required to move a molecule from a medium with the same dielectric as the molecular interior into water. Nonpolar solvation free energy is the energy needed to create a water cavity in which to place the molecule. It is directly related to the molecule's surface area. Entropies for the molecules were calculated according to the Schlitter formula⁵⁵ using principal components analysis over the MD trajectories (see Methods).

The factors leading to stabilization of the ternary complex are relatively consistent across the different simulations (Table 4). For example, all terms between simulations MMC and MMD are comparable with the exception of $\langle \Delta E_{\text{VdW}} \rangle$ which is more attractive for MMD. The two systems differ only by the chirality of the charged Cys. The different orientation of D-Cys within the amino acid binding pocket produces compensating electrostatic contacts such that the electrostatic and polar solvation terms are comparable between MMD and MMC (see Misacylation with O-phosphoserine and D-cysteine and Fig. 7). But, the pocket encloses D-Cys more tightly, resulting in a more stable complex.

Comparisons of free energies for LMC, MMC, and HMC reveal a destabilizing trend due almost entirely to changes in $\langle \Delta E_{\text{VdW}} \rangle$ as Mg^{2+} concentration increases. As hexahydrated Mg^{2+} ions move into the acceptor stem and common arm deep grooves, the tRNA constricts around them causing it to pull slightly away from the EF-Tu binding interface. This is also reflected in $\langle \Delta E_{\text{elec}} \rangle$ which is counteracted by compensating $\langle \Delta G_{\text{polar}} \rangle$. A similar trend does not appear in the systems with unmodified bases due in part to the large entropy difference in MUC discussed below.

A difference in binding free energy between systems with modified and unmodified bases is most evident when comparing MMC and MUC, but this does not generalize to LUC and HUC. The VdW and Coulombic terms compensate for each other such that the main difference between MMC and MUC is the reduction in entropy when unmodified bases are used. MUC has by far the least entropy difference, which required further explanation. The entropies were calculated with Schlitter's formula⁵⁵ using the principal components analysis of the MD trajectory (see Methods). The motion from the three most dominant principal components, which influence the entropy values the most, showed highly correlated motion between the anticodon arm and EF-Tu residues in the effector loop and at the interface of the complex for all the runs. Figure 10 shows EF-Tu and Cys-tRNA^{Cys} from MMC colored by mean squared deviation according to the first principal component. Anticodon motion is readily apparent as are regions of the EF-Tu responsible for communication across the protein-tRNA interface. Contributions from the first three principal components for MMC and MUC are shown in Figure 11 in the supplementary information.

The motion from the lowest principal component accounts for more than half of the covariance data in each of the runs except for MUC where it explains only ~40 % of the covariance in the motion of the backbone atoms (Fig. 12 in supplementary information). This could explain why MUC showed the greatest deviation from the average $-\Delta S$ values. We examined the entropic contribution in all the simulations without the tRNA anticodon

arm included to determine whether this motion was responsible for the discrepancy. Without the anticodon arm, $-T\Delta S$ for MMC is 25.2 kcal/mol, for MUC is almost the same at 25.6 kcal/mol, and for the rest of the runs lies within 25.2–27.5 kcal/mol. The entropies calculated on the full tRNA may be taking into account spurious correlations between fluctuations of the anticodon and EF-Tu. The binding entropy values measured using this approach are large in comparison to the experimentally measured value of ~ 10 kcal/mol⁵⁶, but they are similar in magnitude to protein-RNA binding entropies reported in other computational studies³⁵.

It is worthwhile to compare MMC, the system with physiological conditions, to HUC, which has conditions similar to typical *in vitro* experiments. Both have similar binding free energies, so two system changes, modified vs. unmodified bases and medium vs. high Mg^{2+} concentration, result in small changes to $\langle \Delta G_{\text{binding}} \rangle$. This masks an underlying change in the EF-Tu-Cys-tRNA^{Cys} interface. In HUC, the contact gained between C51 and Arg339 results in a tighter VdW interaction but loss of the contact between His331 and the U55 phosphate creates a more repulsive electrostatic interaction (see Local Non-bonded Interaction Energies). Since these charged groups are now more free to interact with solvent, the polar solvation term becomes more attractive, nullifying the difference in electrostatic interaction.

As discussed earlier, the effect of misacylating tRNA^{Cys} with D-Cys is to increase the VdW attraction producing a more stable complex. It is still an open question whether the D-Cys with its different orientation could successfully dock into the amino acid binding pocket in the first place or whether there would be a steric clash which would not allow this complex to form. In contrast, misacylation with Sep produced the complex with by far the least stable binding free energy. In LUS, $\langle \Delta G_{\text{polar}} \rangle$ does not counteract $\langle \Delta E_{\text{elec}} \rangle$ to the same extent as in the other cases because the extra -2 charge is buried in the amino acid binding pocket and not fully exposed to solvent. The lower stability of this complex is also manifested in the higher amplitude of motion due to principal component 1 in LUS as compared to LUC, especially in the regions of the acceptor and common arms close to the interface (Fig. 13 in supplementary information). Although local interactions with Arg274 are attractive (Fig. 7c), the extra -2 charge on Sep creates an unfavorably high electrostatic repulsion with the rest of the protein. From the free energy analysis, Sep-tRNA^{Cys} appears unlikely to be bound by EF-Tu for transport to the ribosome. This holds for natural systems, such as *Methanosarcinales* species which use Sep-tRNA^{Cys} as an intermediate in the production of Cys-tRNA^{Cys}. Also, in order to engineer a translation system that can incorporate Sep into growing polypeptides, the EF-Tu binding pocket will need to be mutated, possibly requiring the creation of a Sep-specific EF-Tu analogous to SelB, the selenocysteine-specific elongation factor. There is, however, the possibility that cations associated with the Sep phosphate could be present in the amino acid binding pocket stabilizing Sep binding.

Evolutionary analysis showed significant conservation of the interface base pair C51.G63 for some tRNA specificities (see Evolutionary Analyses of EF-Tu and tRNA). This base pair may be used to tune the binding of aa-tRNA to EF-Tu. The exocyclic base amine of G63 typically forms a hydrogen bond with the Glu390 sidechain, and this interaction disappears with U63. To study the importance of this base pair on binding free energy, MMC_{AU} was run with C51.G63 replaced by A51.U63. The effect of this double mutation on binding affinities has been studied for *T. thermophilus* EF-Tu and unmodified *E. coli* tRNA^{Leu10}. Experimentally, $\Delta \Delta G_{\text{binding}}$ of 1.4 kcal/mol was seen between the standard tRNA and the double mutant. Here, the trend remains the same, but the difference between MMC and MMC_{AU} is larger. In the simulation, EF-Tu pulls back from the Cys-tRNA^{Cys} as Glu390 is repelled by tRNA backbone phosphates. This creates an opening in the interface exposing more of the complex to solvent as seen in the larger nonpolar solvation term. Since water

surrounds the area near Glu390, nearby residues Arg330, His331, and Arg393 are solvated, disrupting interactions with one another. This can be seen in the corresponding weakening of $\langle E_{\text{elec}} \rangle$ and strengthening of $\langle G_{\text{polar}} \rangle$ for EF-Tu, but these two changes cancel each other out in the final analysis. The primary contribution to the binding free energy difference between MMC_{AU} and MMC comes from $\langle \Delta E_{\text{VdW}} \rangle$ which shows the wild type complex to be more stable when the G63 exocyclic base amine provides a hydrogen bond donor to Glu390 which is missing in U63. This effect comes mostly from the weakened interface VdW attraction due to Glu390 and the surrounding EF-Tu residues separating from Cys-tRNA^{Cys}.

Despite choosing the interior dielectric to place the binding free energy of MMC near physiological values, the spectrum of binding free energies appears exaggerated. If separate trajectories had been created for lone EF-Tu and aa-tRNA^{Cys} molecules, the adaptation energy of forming the complex could have been taken into account. As reported by Kollman and co-workers, binding free energies calculated solely from trajectories of a complex are more favorable than those that use separate trajectories for the unbound molecules³⁸. This is because the bound molecules are already in conformations favorable to the state of the complex so the reorganization energy associated with conformational changes required to move from unbound to bound structures is ignored.

Conclusion

The dynamics and energetics of EF-Tu.aa-tRNA binding depend on many factors, such as the amino acids and nucleotides present at the binding interface, cation concentrations, modified tRNA nucleosides, and the particular amino acid with which the tRNA is charged. Our extensive evolutionary and energetic analyses revealed evolutionarily conserved features important for EF-Tu.aa-tRNA binding (Fig. 5, 6). Highly conserved, charged EF-Tu residues interact with nucleotides at both ends of the coaxially stacked acceptor and common arm domain, and these nucleotides are well conserved across different tRNA specificities. The discriminator base as well as base pairs 2·71, 3·70, and 51·63 show correlation with tRNA specificity for several cases and may be responsible for tuning the EF-Tu-tRNA binding interface energies (Fig. 3).

Many *in vitro* experiments with tRNA use extremely high Mg²⁺ concentrations and transcribed tRNA. The MD simulations here showed that both of these conditions can affect tRNA structure, dynamics, and binding to EF-Tu. Higher Mg²⁺ concentrations stabilize tRNA^{Cys} structure but destabilize EF-Tu.aa-tRNA^{Cys} binding in systems with modified bases. Modified nucleosides are important for tRNA structural stability; for example, the 7-nucleotide anticodon loop is kept open by Mia37 which is also less attractive to Mg²⁺ ions than A37. Modified bases T54 and Ψ55 cause a shift in contacts at the EF-Tu binding interface relative to the transcribed system.

EF-Tu binding to misacylated tRNA^{Cys} appears energetically feasible for D-cysteine but infeasible for O-phosphoserine due mainly to its high negative charge (Table 4). Protonation of Sep may improve its binding energy, but to date only amino acid precursors or unnatural amino acids with large, neutral or positive functional groups have been successfully introduced into the genetic code using the cellular translational machinery. Our two misacylation simulations support the hypothesis that EF-Tu discriminates against tRNAs misacylated with amino acids carrying extra negative charge while same charge misacylations are readily accepted.

Trends shown in MM-PBSA analyses of the ternary complex obtained from averages over 20-ns MD simulations and the observed structural variations in them are consistent with

previously reported experimental results. Our study provides detailed mechanisms for: fluctuations in anticodon loop base pair formation with unmodified A37, lower binding free energy for tRNA A51·U63 mutants, and destabilization of tRNA structure at low Mg^{2+} concentration. Longer simulations may reveal more base pairing in the anticodon loops of unmodified tRNAs. Calculation of ion residency times should also improve with much longer simulation times that allow the unbound ions to achieve ergodicity. While the energetic components observed in the MM-PBSA study help to elucidate the underlying phenomena responsible for the free energy difference, the entropy obtained when the full tRNA is included seem to be too large.

The combination of evolutionary analyses and MD focused this study on important aspects of EF-Tu-aa-tRNA binding that have been maintained since the last universal common ancestral state as well as features of the binding interface common to all three domains of life. Evolutionary analysis of the binding interface suggests ways to understand mechanisms for aa-tRNA discrimination. The evolutionary interpretation predicts changes in binding free energy with the mutation of nucleotide 73 in Cys-tRNA^{Cys} as well as base pairs 2·71 and 3·70 in Gly-tRNA^{Gly} (Fig. 3) in addition to the modifications of 51·63 already studied.

Materials and Methods

Bioinformatics

EF-Tu/EF1A sequences were obtained through a BLAST⁵⁷ search against SWISS-PROT⁵⁸ and the National Center for Biotechnology Information⁵⁹ databases, and structures were retrieved from the Protein Data Bank⁶⁰. The 338 sequences obtained using BLAST were aligned with ClustalW⁶¹ and improved manually. All of the above procedures were carried out using Multi-Seq⁴⁰.

EPs are statistically well balanced profiles that span the evolutionary space of a given set of sequences and can be derived using the sequence-based⁴² or structure-based⁶² QR factorization. These algorithms perform multidimensional QR factorization of a multiple sequence or structure alignment and order the sequences according to increasing linear dependence, which corresponds to redundancy in the alignment. As EF-Tu/EF1A is highly conserved, sequence QR was used to obtain an EP with 18 sequences representing all three domains of life: 3 bacteria, 9 archaea, and 6 eukaryotes. The most diverse sequences are shown in Figure 2, but the conservation coloring is based on the entire set. This EP was generated using a sequence identity cutoff of 75 % and seeded with the *T. aquaticus*, *Sulfolobus solfataricus*, and *Saccharomyces cerevisiae*, for which structures are available. Another EP with five sequences representing only Bacteria was derived using a sequence identity cutoff of 75% and seeded with *T. aquaticus* so that the structure could be colored by sequence identity.

Sequences from all twenty tRNA specificities were obtained from the tRNA Compilation 2000⁴⁵, the Integrated Microbial Genome database at the Joint Genome Institute⁶³, and the Genomic tRNA database⁶⁴. These were separated by domain of life and anticodon, and each group was aligned using ClustalW⁶¹ and improved manually. Because tRNA is short and difficult to correctly identify in genomic data, a list of universally conserved bases was established as quality control and used to remove sequences which were not tRNA or were initiator or mitochondrial tRNA. These bases are G18, G19, U33, G53, U54, U55, and C56. Further, only complete sequences were considered, complete indicating that the ends of the tRNA were intact and no deletions or additions were present in the sequence except in the variable regions in the GG arm and the variable loop. Introns were excised based on comparison to the rest of the sequences in the group and the tRNA Compilation 2000 sequences, which already had introns removed. Using the standalone program seqqr 1, all

tRNA EPs were constructed at a QR cutoff of 85% using norm order = 3 and gap penalty = 0.5 with the *E. coli*, *Haloferax volcanii*, or *S. cerevisiae* sequence of each tRNA specificity being used to seed the profile. These organisms were chosen because the tRNA Compilation 2000 database contains sequences with modified nucleosides for all tRNAs within these organisms. Consensus sequences were created from these EPs and aligned together to create a list of highly conserved nucleotides within each isoacceptor of each specificity for each domain of life⁴³.

Sequence Entropy

The amount of sequence variability for a particular column (*i*) of the alignment is measured

by sequence entropy (S_i) which is expressed as $S_i = - \sum_{j=1}^4 P_{ij} \ln P_{ij}$, where P_{ij} is the probability of finding nucleotide *j* in column *i* in the alignment. P_{ij} is measured using the

expression: $P_{ij} = \frac{N_{ij}}{N_{seq}}$, where N_{ij} is the number of sequences with nucleotide *i* in column *j* in the alignment, and N_{seq} is the number of sequences in the alignment. Conserved columns and columns dominated by gaps (such as insertions in the GG loop and variable arm of tRNA) have low sequence entropy while variable columns have high entropy. If sequence tuning occurs in tRNA, nucleotides important for tuning may be conserved within a single specificity across all organisms. For a particular nucleotide position, tuning is established based upon two criteria: high entropy (low conservation) over all species and all tRNA specificities, and low entropy (high conservation) over all species within a given specificity. The difference between these two entropies provides a quick method to identify potential tuning elements for a given specificity. Hence, the correlation of the base in column *i* to the amino acid specificity is high when $\Delta S_i = S_i - S_{iAA}$ is large, where S_i is the entropy at position *i* for the full alignment and S_{iAA} is the entropy for the subset specific to a particular amino acid (AA) specificity.

Molecular Dynamics Analysis

General System Setup—Three main parameters were varied between the simulations: Mg^{2+} ion concentration (low, medium, high), modified vs. unmodified nucleosides, and the amino acid attached to tRNA^{Cys} (Cys, D-Cys, Sep). tRNA^{Cys} in structure 1B23 contains modified bases so for systems run with transcribed tRNA^{Cys}, the corresponding bases were mutated back to their unmodified state. The conformations of these unmodified bases were checked against the structure of transcribed tRNA^{Cys} complexed with CysRS in 1U0B⁶⁵. Each system was neutralized through the addition of Mg^{2+} and K^+ ions dependent on the corresponding concentration. Hydrogens were added with psfgen (VMD) using CHARMM27⁶⁶ topologies for proteins and nucleic acids with custom topologies for the modified bases and the 3' aminoacylated A76 (explained below). The systems were solvated using the solvate package within VMD⁴¹. All simulation was done with periodic boundary conditions using the NPT ensemble with pressure set to 1.01325 bar. Timesteps used for force calculations were: bonded 1 fs, VdW 2fs, and electrostatic 4fs. VdW forces were calculated with a switching distance of 10 Å and a cutoff of 12 Å. Electrostatics were calculated with the Particle Mesh Ewald method⁶⁷. The MD program NAMD2 was used to generate all of the trajectories⁶⁸.

Throughout most of minimization and equilibration, the water molecules and ions were allowed to associate with aa-tRNA^{Cys} while constraints were applied to the macromolecules. For minimization, heavy atoms were fixed for 2000 steps. Then waters were freed but ions

¹<http://www.scs.uiuc.edu/schulten/software/index.html/>

kept fixed for 3000 steps. Ions and sidechains were then freed while holding protein and nucleic acid backbones fixed for the next 5000 steps. Finally, everything was set free for the last 20000 steps of minimization.

For equilibration, a temperature jump protocol similar to that used by Auffinger and Westhof¹⁴ was employed to allow water and ions to slowly diffuse into the tRNA structure. The initial temperature was set to 100 K, and ions and heavy atoms in the protein and nucleic acid chains were harmonically constrained for the first 25000 fs. Then the temperature was raised to 200 K, and backbone atoms were harmonically constrained for 25000 fs. Force constants for all harmonic constraints were set to 1 kcalmol⁻¹ Å². Finally, the temperature was raised to 298.15 K, and all atoms were freed for the next 3.9 ns. After this 4 ns equilibration, each system was run for 16 ns.

Initial Ion Placement—Both Mg²⁺ and K⁺ ions were placed by the program *ionize*.² It creates a 3D grid across the system, computes Coulombic interaction energies for the addition of an ion at each gridpoint, and chooses the gridpoint with the most attractive energy. An ion is then placed at the chosen point, and the process is repeated for each successive ion. The Mg²⁺ ions were included to provide a more realistic ionic environment in which to simulate the aa-tRNA^{Cys}. Stability provided by associated Mg²⁺ is important to prevent RNA unfolding during longer MD simulations.

The close Mg²⁺ placement protocol was tested against the tRNA structure from 1FFY⁶⁹ which contains 10 resolved Mg²⁺ ions. All water and ions were removed from the structure, and *ionize* was used to place 10 Mg²⁺ ions around the protein-tRNA complex. The first few were placed within the elbow region of the tRNA near identified ions, but later placements became less accurate. Only a few (1–6) Mg²⁺ ions have been observed to bind directly to tRNAs^{70,71}. A final ion placement protocol was established. Including crystal Mg²⁺ ions, four Mg²⁺ ions were allowed near enough to the tRNA to make direct contact. These Mg²⁺ ions were placed 2 Å away from the tRNA, and all remaining Mg²⁺ and K⁺ ions were placed 6.5 Å away. After ion placement, all Mg²⁺ ions were solvated using a script which checks each Mg²⁺ ion for nearby oxygen atoms and adds water oxygens until the ion is fully solvated with octahedral coordination in its first solvation shell. In preliminary MD runs, Mg²⁺ ions placed 6.5 Å away from tRNA^{Cys} sometimes managed to bind directly to the phosphate backbone before acquiring all six primary solvation shell waters, and since first solvation shell water molecules exchange at timescales (~10 μs) longer than the MD trajectories, Mg²⁺ solvation was necessary to prevent unrealistic contacts. Three Mg²⁺ concentrations were considered: high (26 atoms), medium (18 atoms), and low (11 atoms) (see Table 2). Medium Mg²⁺ concentration was chosen to be close to cellular ion conditions (30 mM)⁵⁴, acknowledging that Mg²⁺ concentration is higher near nucleic acids due to their tendency to bind Mg²⁺.

Force Field Parameters—The structure of the ternary complex used (PDB ID 1B23) is from a crystal with GDPNP, a non-hydrolyzable GTP analogue, which was replaced by GTP in all simulations. The partial atomic charges for GTP were derived based on ATP (available in CHARMM27⁶⁶), and no additional parameters were required. Parameters for the bond between A76 and Sep were previously developed by analogy²⁴, and those for the bond between A76 and Cys/D-Cys were taken by analogy from the Sep system. For D-Cys the regular Cys parameters were used but the chirality of the molecule was changed by swapping the amine and hydrogen around the C_α.

²<http://www.ks.uiuc.edu/Development/MDTools/ionize/>

Modified Bases: The parameters for the modified bases were derived by analogy to parameters of the respective transcribed bases and all charges and equilibrium bond lengths and angles were verified using *ab initio* quantum chemistry calculations with the 6–31G* basis set in Spartan04 (Wavefunction, Inc., Irvine, CA).

Dihydrouridine and 4-thiouridine: The parameters for D were available as a patch in the Unified Atom force field⁷², which is distributed as a stream file with the CHARMM27 force field. Hydrogens were added to this patch by analogy with uridine and cyclopentane. The parameters for 4su were derived by analogy with uridine.

Pseudouridine: The parameters for the standard Ψ tautomer were developed by analogy with cytidine and mostly agreed with preexisting parameters provided by Lennart Nilsson. These were used for Ψ 39 and Ψ 55. A different tautomer was deemed necessary for Ψ 32 because after roughly 1ns of simulation, the base destacked from the anticodon loop and formed a hydrogen bond with the A31 phosphate. The anticodon arm would then unravel regardless of the Mg^{2+} concentration. Similar activity was achieved in the unmodified system by flipping U32 to expose the base amine. The Ψ 32 tautomer had the N3 hydrogen moved to O4, forming a hydrogen bond with its own phosphate, and retaining the crystal structure conformation. The parameters were derived by analogy to uridine and cytidine as well as a guanine tautomer in a stream file.

2-methylthio- N^6 -isopentenyladenosine: The parameters for Mia were derived by analogy with adenine, ethyl methyl sulfide, ethenyl methyl thioether, 2-butene, dimethyl amine, and isopentenyl amine. All quantum chemistry calculations for the parameterization step for Mia were performed with GAMESS⁷³ using the MP2/Hartree Fock methods with the 6–31G* basis set. The force constants for all the bond parameters (bond stretching, angle bending, and torsion parameters) were obtained by scaling the corresponding terms from the quantum chemistry calculations with similar CHARMM27 parameters from the molecules listed above. The equilibrium values for these bonding terms were obtained from the optimization step of the quantum chemistry calculations.

All parameters were tested using a short minimization and equilibration with NAMD2. The parameters were also checked by comparing the conformations of the corresponding nucleotides in the tRNA simulation with those in the crystal structure.

Electrostatic Calculations—Electrostatic potential maps were calculated by solving the nonlinear Poisson-Boltzmann equation (NPBE) with the APBS (Adaptive Poisson-Boltzman Solver) software package⁷⁴, available within VMD⁴¹. NPBE solutions were calculated at 298.15K, using 150mM KCl with solute and solvent dielectric constants of 2.0 and 78.54, respectively, and both atomic radii and partial atomic charges from the CHARMM27 force field⁶⁶.

An electrostatic potential map was calculated for every 20 ps of the MMC trajectory, considering: the ternary complex plus the ions within 5 Å, EF-Tu-GTP plus the ions within 5 Å, and aa-tRNA and ions within 5 Å. Maps corresponding to the last 5 ns of the simulation were averaged. Compared to isopotentials from a single frame, time averaging over the last 5 ns did not change the EF-Tu ± 1 k_BT/e surfaces significantly.

Ion Residency and Occupancy—Before the calculation of residency and occupancy sites, each frame of the trajectory (last 16 ns of equilibration sampled every 40 ps) was aligned to the first frame using the tRNA^{Cys} backbone. Residency sites were identified by tracking the locations of ions over time. When an ion strayed 3 Å from its starting location,

its residency time was recorded, and a new starting location was noted with residency time reset to 0 ns.

Occupancy maps were generated using the VolMap plugin to VMD. Average occupancy was calculated at 1 Å resolution for Mg^{2+} and K^+ ions separately.

Free Energy Calculations—Free energies of binding for the ternary complex were determined using the MM-PBSA method^{34,35,38,36,39,37}. The form of the free energy is $\langle \Delta G_{\text{binding}} \rangle = \langle \Delta G_{\text{polar}} + \Delta G_{\text{nonpolar}} + \Delta E_{\text{elec}} + \Delta E_{\text{vdW}} - T\Delta S \rangle$. Each term on the right side of the above equation is a difference between the value for the ternary complex and the sum of values for the EF-Tu-GTP and tRNA $\langle \Delta G_{\text{binding}} \rangle = \langle G_{\text{complex}} - (G_{\text{EF-Tu-GTP}} + G_{\text{aa-tRNA}^{\text{Cys}}}) \rangle$. Bound Mg^{2+} ions along with their primary shell waters were included: 1 bridging EF-Tu to GTP and 4 directly bound to the aa-tRNA^{Cys}. These hydrated cations were included because they are stable components of the system. Their overall effect on $\langle \Delta G_{\text{binding}} \rangle$ is to shift it 20–25 kcal/mol closer to 0 kcal/mol primarily through the ΔE_{elec} and ΔG_{polar} terms. Values were obtained for 500 frames (10,000 frames for $T\Delta S$) across the last 5 ns of the 20 ns trajectories. This time window and resolution was chosen because it minimized both the error in free energy and computation time. Processing more frames did not lower the error bounds appreciably. G_{polar} , the polar solvation energy, is the energy required to move the system from a dielectric of $\epsilon_{\text{in}} = 1.6$ to $\epsilon_{\text{out}} = 78.54$. ϵ_{in} was set to 1.6 so that $\langle \Delta G_{\text{binding}} \rangle$ for MMC would be close to −10 kcal/mol, a physiologically reasonable quantity. The overall trends in binding free energies between different simulations were the same as those seen with $\epsilon_{\text{in}} = 2$. G_{polar} terms were obtained by numerically solving the NPBE with APBS⁷⁴. Solutions were calculated at 298.15K, using 117mM KCl and 38mM MgCl_2 , and both atomic radii and partial atomic charges from the CHARMM27 force field⁶⁶ on a grid of 129 points per dimension with automated focussing. The nonpolar solvation energy, G_{nonpolar} , is the energy required to make a cavity in the solvent for a given system. $G_{\text{nonpolar}} = \gamma \text{SASA} + b$ where $\gamma = 0.00542 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$, SASA is the solvent accessible surface area, and $b = 0.92 \text{ kcal/mol}$. SASA values were generated using a solvent radius of 1.4 Å. E_{elec} is the sum of all pair-wise Coulombic interactions in a system with the dielectric $\epsilon_{\text{in}} = 1$. G_{nonpolar} and E_{elec} terms were calculated by the acc and coulomb programs bundled with APBS⁷⁴. EVdW is the sum of pairwise van der Waals interaction energies. EVdW terms were calculated using the NAMDenergy plugin to VMD^{41,68} with VdW parameters from the CHARMM27 force field⁶⁶.

The absolute entropies of the complex, the protein, and the tRNA molecules were calculated using Schlitter's formula^{55,75,76} based upon the covariance matrix of fluctuations in atomic positions. For our entropy estimate, 10,000 frames were used for the covariance matrix calculation from the last 5 ns of the 20-ns simulation of the equilibration of the complex under the specified conditions. Schlitter showed that the entropy (S) of a quantum mechanical particle with one degree of freedom (x) is smaller than the entropy of a simple harmonic oscillator (S_{sho}) which is:

$$S \leq S_{\text{sho}} = \frac{k_B \alpha}{e^\alpha - 1} - k_B \ln[1 - e^{-\alpha}], \quad (1)$$

where $\alpha = \hbar\omega/k_B T$, ω is the frequency of the oscillator, $\hbar = h/(2\pi)$ and h is Planck's constant. In the classical approximation, the frequency ω of the oscillator is taken to be proportional to the inverse variance of the system from molecular dynamics simulations ($\langle x^2 \rangle_c$) and is solved using the equipartition theorem: $m\omega^2 \langle x^2 \rangle_c = k_B T$. The entropy is then solved by substituting for ω in equation 1 and using the approximation:

$$S \leq S_{sho} \leq S' = \frac{k_B \ln(1 + \frac{k_B T e^2}{\hbar^2} m \langle x^2 \rangle_c)}{2}. \quad (2)$$

Schlitter approximated the above formula for a system with many degrees of freedom using the mass-weighted covariance matrix σ' of the atomic fluctuations which is defined as:

$\sigma'_{ij} = \sqrt{m_i m_j} \langle (x_i - \langle x_i \rangle)(x_j - \langle x_j \rangle) \rangle$, where x_i refers to one of the Cartesian coordinates of an atom and m_i refers to its mass. The mass-weighted covariance matrix can also be written in terms of the covariance matrix (σ) as: $\sigma' = M^{1/2} \sigma M^{1/2}$, where M refers to the mass matrix in which the diagonal elements hold the masses corresponding to each degree of freedom. The mass-weighted covariance matrix is diagonalized to give a new set of uncorrelated mass-weighted coordinates q_i , and the eigenvalues (λ_i) of this matrix represent the variance ($\langle q_{ii}^2 \rangle_c$) of each degree of freedom in the new coordinate system. The entropy can be calculated as:

$$S \leq S' = \frac{k_B \sum_{i=1}^{3N} \ln(1 + \frac{k_B T e^2}{\hbar^2} \lambda_i)}{2}. \quad (3)$$

Since the matrix is diagonalized, the product of the diagonal entries is equal to the determinant of the covariance matrix. This allows the entropy to be calculated using the determinant of the covariance matrix before the transformations. Entropy was calculated in this work using the backbone C_α and phosphorous atoms. We use the program Carma⁷⁷ to diagonalize the mass-weighted covariance matrix using principal components analysis, and the eigenvalues of the covariance matrix are used in the above for λ_i . The entropy calculations measured using Schlitter's formula were found to be comparable to the entropy change using ideal gas approximation for translational and rotational entropy change and using anisotropic network analysis⁴⁹ for vibrational entropy change.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to thank Elijah Roberts for providing the Tcl script for calculating ion residency times. Lennart Nilsson kindly provided us with pseudouridine parameters. We thank Ivet Bahar for sending us her anisotropic network analysis program for protein-RNA complexes. We also thank Nathan Baker and Olke Uhlenbeck for helpful discussions. Funding and supercomputer time was provided by: NSF (MCB04-46227), NIH Molecular Biophysics Training Grant (PHS5T32GM08276), NIH Chemical Biology Training Grant (5T32GM070421), NCSA LRAC (MCA03T027), NIH (P41-RR05969), and NSF CRIF (0541659).

References

1. Bosch L, Kraal B, Van der Meide PH, Duisterwinkel FJ, Van Noort JM. The elongation factor EF-Tu and its two encoding genes. *Prog. Nucl. Acid Res. Mol. Res.* 1983; 30:91–9126.
2. Theobald-Dietrich A, Frugier M, Giege R, Rudinger-Thirion J. Atypical archaeal tRNA pyrrolysine transcript behaves towards EF-Tu as a typical elongator tRNA. *Nucl. Acids Res.* 2004; 32(3):1091–1096. [PubMed: 14872064]
3. Nissen P, Kjeldgaard M, Thirup S, Polekhina G, Reshetnikova L, Clark BFC, Nyborg J. Crystal structure of the ternary complex of Phe-tRNA^{Phe}, EF-Tu, and a GTP analog. *Science.* 1995; 270:1464–1472. [PubMed: 7491491]

4. Nissen P, Thirup S, Kjeldgaard M, Nyborg J. The crystal structure of Cys-tRNA^{Cys}-EF-Tu-GDPNP reveals general and specific features in the ternary complex in the tRNA. *Structure*. 1999; 7:143–156. [PubMed: 10368282]
5. Gupta R, Woese CR. Unusual modification patterns in the transfer ribonucleic acids of archaeobacteria. *Curr. Microbiol.* 1980; 4(4):245–249.
6. Iwabe N, Kuma K, Hasegawa M, Osawa S, Miyata T. Evolutionary relationship of archaeobacteria, eubacteria, and eukaryotes inferred from phylogenetic trees of duplicated genes. *Proc. Natl. Acad. Sci. USA*. 1989; 86:9355–9359. [PubMed: 2531898]
7. Perona J, Hou Y. Indirect Readout of tRNA for Aminoacylation. *Biochemistry*. 2007; 46(37):10419–10432. [PubMed: 17718520]
8. LaRiviere FJ, Wolfson AD, Uhlenbeck OC. Uniform binding of aminoacyl-tRNAs to elongation factor Tu by thermodynamic compensation. *Science*. 2001; 294(5540):165–168. [PubMed: 11588263]
9. Dale T, Sanderson LE, Uhlenbeck OC. The affinity of elongation factor Tu for an aminoacyl-tRNA is modulated by the esterified amino acid. *Biochemistry*. 2004; 43(20):6159–6166. [PubMed: 15147200]
10. Sanderson LE, Uhlenbeck OC. The 51–63 base pair of tRNA confers specificity for binding by EF-Tu. *RNA*. 2007; 13(6):835–840. [PubMed: 17449728]
11. Helm M. Post-transcriptional nucleotide modification and alternative folding of RNA. *Nucl. Acids Res.* 2006; 34(2):721–733. [PubMed: 16452298]
12. Thomas G, Favre A. 4-Thiouridine triggers both growth delay induced by near-ultraviolet light and photoprotection. *Eur. J. Biochem.* 1980; 113(1):67–74. [PubMed: 7007049]
13. Jones CI, Spencer AC, Hsu JL, Spemulli LL, Martinis SA, DeRider M, Agris PF. A counterintuitive Mg²⁺-dependent and modification-assisted functional folding of mitochondrial tRNAs. *J. Mol. Biol.* 2006; 362(4):771–786. [PubMed: 16949614]
14. Auffinger P, Westhof E. RNA hydration: three nanoseconds of multiple molecular dynamics simulations of the solvated tRNA(Asp) anticodon hairpin. *J. Mol. Biol.* 1997; 269(3):326–341. [PubMed: 9199403]
15. Charette M, Gray MW. Pseudouridine in RNA: what, where, how, and why. *IUBMB Life*. 2000; 49(5):341–351. [PubMed: 10902565]
16. Davis DR. Stabilization of RNA stacking by pseudouridine. *Nucl. Acids Res.* 1995; 23(24):5020–5026. [PubMed: 8559660]
17. Cabello-Villegas J, Winkler ME, Nikonowicz EP. Solution conformations of unmodified and A(37)N(6)-dimethylallyl modified anticodon stem-loops of *Escherichia coli* tRNA(Phe). *J. Mol. Biol.* 2002; 319(5):1015–1034. [PubMed: 12079344]
18. Wang L, Brock A, Herberich B, Schultz PG. Expanding the genetic code of *Escherichia coli*. *Science*. 2001; 292(5516):498–500. [PubMed: 11313494]
19. Stanzel M, Schon A, Sprinzl M. Discrimination against misacylated tRNA by chloroplast elongation factor Tu. *Eur. J. Biochem.* 1994; 219(1–2):435–439. [PubMed: 8307009]
20. Becker HD, Kern D. *Thermus thermophilus*: a link in evolution of the tRNA-dependent amino acid amidation pathways. *Proc. Natl. Acad. Sci. USA*. 1998; 95(22):12832–12837. [PubMed: 9789000]
21. Roy H, Becker HD, Mazauric M-H, Kern D. Structural elements defining elongation factor Tu mediated suppression of codon ambiguity. *Nucleic Acids Res.* 2007; 35(10):3420–3430. [PubMed: 17478519]
22. Ibba M, Söll D. Quality control mechanisms during translation. *Science*. 1999; 286:1893–1897. [PubMed: 10583945]
23. Sauerwald A, Zhu W, Major TA, Roy H, Palioura S, Jahn D, Whitman WB, Yates JR 3rd, Ibba M, Söll D. RNA-dependent cysteine biosynthesis in archaea. *Science*. 2005; 307(5717):1969–1972. [PubMed: 15790858]
24. O'Donoghue P, Sethi A, Woese CR, Luthey-Schulten ZA. The evolutionary history of Cys-tRNA^{Cys} formation. *Proc. Natl. Acad. Sci. USA*. 2005; 102(52):19003–19008. [PubMed: 16380427]
25. Ma J, Karplus M. Molecular switch in signal transduction: reaction paths of the conformational changes in ras p21. *Proc. Natl. Acad. Sci. USA*. 1997; 94(22):11905–11910. [PubMed: 9342335]

26. Kosztin I, Bruinsma R, O'Lague P, Schulten K. Mechanical force generation by G proteins. *Proc. Natl. Acad. Sci. USA*. 2002; 99(6):3575–3580. [PubMed: 11904419]
27. Prabhakaran M, Harvey SC, McCammon JA. Molecular-dynamics simulation of phenylalanine transfer RNA. II. Amplitudes, anisotropies, and anharmonicities of atomic motions. *Biopolymers*. 1985; 24(7):1189–1204. [PubMed: 2992621]
28. Auffinger, P.; Vaiana, A. Molecular Dynamics Simulations of RNA Systems, Ch. 34. In: Hartmann, RK.; Bindereif, A.; Schöen, A.; Westhof, E., editors. *Handbook of RNA Biochemistry*. Weinheim: Wiley-VCH Verlag; 2005. p. 560-576.
29. Norberg J, Nilsson L. Molecular dynamics applied to nucleic acids. *Acc. Chem. Res.* 2002; 35(6): 465–472. [PubMed: 12069632]
30. Cheatham TE 3rd, Kollman PA. Molecular dynamics simulation of nucleic acids. *Ann. Rev. Phys. Chem.* 2000; 51:435–471. [PubMed: 11031289]
31. Auffinger P, Louise-May S, Westhof E. Molecular dynamics simulations of solvated yeast tRNA(Asp). *Biophys. J.* 1999; 76(1 Pt 1):50–64. [PubMed: 9876122]
32. Draper DE. A guide to ions and RNA structure. *RNA*. 2004; 10(3):335–343. [PubMed: 14970378]
33. Hermann T, Westhof E. Exploration of metal ion binding sites in RNA folds by Brownian-dynamics simulations. *Structure*. 1998; 6(10):1303–1314. [PubMed: 9782053]
34. Froloff N, Windemuth A, Honig B. On the calculation of binding free energies using continuum methods: application to MHC class I protein-peptide interactions. *Prot. Sci.* 1997; 6(6):1293–1301.
35. Kollman PA, Massova I, Reyes C, Kuhn B, Huo S, Chong L, Lee M, Lee T, Duan Y, Wang W, Donini O, Cieplak P, Srinivasan J, Case DA, Cheatham TE 3rd. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc. Chem. Res.* 2000; 33(12):889–897. [PubMed: 11123888]
36. Rocchia W, Alexov E, Honig B. Extending the Applicability of the Nonlinear Poisson-Boltzmann Equation: Multiple Dielectric Constants and Multivalent Ions. *J. Phys. Chem. B*. 2001; 105:6507–6514.
37. Pogorelov TV, Autenrieth F, Roberts E, Luthey-Schulten ZA. Cytochrome *c*₂ Exit Strategy: Dissociation Studies and Evolutionary Implications. *J. Phys. Chem. B*. 2007; 111(3):618–634. [PubMed: 17228920]
38. Reyes CM, Kollman PA. Structure and thermodynamics of RNA-protein binding: using molecular dynamics and free energy analyses to calculate the free energies of binding and conformational change. *J. Mol. Biol.* 2000; 297(5):1145–1158. [PubMed: 10764579]
39. Gohlke H, Kiel C, Case DA. Insights into protein-protein binding by binding free energy calculation and free energy decomposition for the Ras-Raf and Ras-RalGDS complexes. *J. Mol. Biol.* 2003; 330(4):891–913. [PubMed: 12850155]
40. Roberts E, Eargle J, Wright D, Luthey-Schulten Z. MultiSeq: Unifying sequence and structure data for evolutionary analysis. *BMC Bioinformatics*. 2006; 7(1):382. [PubMed: 16914055]
41. Humphrey W, Dalke A, Schulten K. VMD – Visual Molecular Dynamics. *J. Mol. Graphics*. 1996; 14:33–38.
42. Sethi A, O'Donoghue P, Luthey-Schulten Z. Evolutionary profiles from the QR factorization of multiple sequence alignments. *Proc. Natl. Acad. Sci. USA*. 2005; 102:4045–4050. [PubMed: 15741270]
43. Full tRNA evolutionary profile available upon request.
44. Rodnina MV, Gromadski KB, Kothe U, Wieden H-J. Recognition and selection of tRNA in translation. *FEBS Lett.* 2005; 579(4):938–942. [PubMed: 15680978]
45. Sprinzl M, Horn C, Brown M, Ioudovitch A, Steinberg S. Compilation of tRNA sequences and sequences of tRNA genes. *Nucl. Acids Res.* 1998; 26(1):148–153. [PubMed: 9399820]
46. Gupta R. *Halobacterium volcanii* tRNAs. Identification of 41 tRNAs covering all amino acids, and the sequences of 33 class I tRNAs. *J. Biol. Chem.* 1984; 259(15):9461–9471. [PubMed: 6746655]
47. Komatsoulis GA, Abelson J. Recognition of tRNA(Cys) by *Escherichia coli* cysteinyl-tRNA synthetase. *Biochemistry*. 1993; 32(29):7435–7444. [PubMed: 8338841]
48. Asahara H, Uhlenbeck OC. The tRNA specificity of *Thermus thermophilus* EF-Tu. *Proc. Natl. Acad. Sci. USA*. 2002; 99(6):3499–3504. [PubMed: 11891293]

49. Wang Y, Rader AJ, Bahar I, Jernigan RL. Global ribosome motions revealed with elastic network model. *J. Struct. Biol.* 2004; 147:302–314. [PubMed: 15450299]
50. Weixlbaumer A, Murphy FV 4th, Dziergowska A, Malkiewicz A, Vendeix FAP, Agris PF, Ramakrishnan V. Mechanism for expanding the decoding capacity of transfer RNAs by modification of uridines. 2007; 14(6):498–502.
51. Nissen P, Hansen J, Ban N, Moore PB, Steitz TA. The structural basis of ribosome activity in peptide bond synthesis. *Science.* 2000; 289:920–930. [PubMed: 10937990]
52. Ohtaki H. Ionic Solvation in Aqueous and Nonaqueous Solutions. *Monatsh. Chem.* 2001; 132:1237–1268.
53. Chin K, Sharp KA, Honig B, Pyle AM. Calculating the electrostatic properties of RNA provides new insights into molecular interactions and function. *Nature Struct. Biol.* 1999; 6(11):1055–1061. [PubMed: 10542099]
54. Cowan, JA. *The Biological Chemistry of Magnesium*. New York: VCH Publishers, Inc; 1995.
55. Schlitter J. Estimation of absolute and relative entropies of macro-molecules using the covariance matrix. *Chem. Phys. Lett.* 1993; 215:617–621.
56. Pingoud A, Urbanke C, Krauss G, Peters F, Maass G. Ternary complex formation between elongation factor Tu, GTP and aminoacyl-tRNA: an equilibrium study. *Eur J Biochem.* 1977; 78(2):403–409. [PubMed: 334538]
57. McGinnis S, Madden TL. BLAST: at the core of a powerful and diverse set of sequence analysis tools. *Nucl. Acids Res.* 2004; 32:W20–W25. [PubMed: 15215342]
58. Boeckmann B, Bairoch A, Apweiler R, Blatter M-C, Estreicher A, Gasteiger E, Martin MJ, Michoud K, O'Donovan C, Phan I, Pilboud S, Schneider M. The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucl. Acids Res.* 2003; 31:365–370. [PubMed: 12520024]
59. Wheeler DL, Chappey C, Lash AE, Leipe DD, Madden TL, Schuler GD, Tatusova TA, Rapp BA. Database resources of the National Center for Biotechnology Information. *Nucl. Acids Res.* 2000; 28:10–14. [PubMed: 10592169]
60. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The protein data bank. *Nucl. Acids Res.* 2000; 28:235–242. [PubMed: 10592235]
61. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucl. Acids Res.* 1994; 22:4673–4680. [PubMed: 7984417]
62. O'Donoghue P, Luthey-Schulten Z. Evolutionary profiles derived from the QR factorization of multiple structural alignments gives an economy of information. *J. Mol. Biol.* 2005; 346(3):875–894. [PubMed: 15713469]
63. Markowitz VM, Korzeniewski F, Palaniappan K, Szeto E, Werner G, Padki A, Zhao X, Dubchak I, Hugenholtz P, Anderson I, Lykidis A, Mavromatis K, Ivanova N, Kyrpides NC. The integrated microbial genomes (img) database. *Nucl. Acids Res.* 2006; 34:D344–D348. [PubMed: 16381883]
64. Lowe TM, Eddy SR. tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. *Nucl. Acids Res.* 1997; 25:955–964. [PubMed: 9023104]
65. Hauenstein S, Zhang C-M, Hou Y-M, Perona JJ. Shape-selective RNA recognition by cysteinyl-tRNA synthetase. 2004; 11(11):1134–1141.
66. Foloppe N, MacKerell AD Jr. All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data. *J. Comp. Chem.* 2000; 21:86–104.
67. Darden T, York D, Pedersen L. Particle mesh Ewald: An N-log(N) method for Ewald sums in large systems. *J. Chem. Phys.* 1993; 89(12):10089–10092.
68. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K. Scalable molecular dynamics with NAMD. *J. Comp. Chem.* 2005; 26(16):1781–1802. [PubMed: 16222654]
69. Silvian LF, Wang J, Steitz TA. Insights into editing from an ile-tRNA synthetase structure with tRNA^{ile} and mupirocin. *Science.* 1999; 285:1074–1077. [PubMed: 10446055]

70. Quigley GJ, Teeter MM, Rich A. Structural analysis of spermine and magnesium ion binding to yeast phenylalanine transfer RNA. *Proc. Natl. Acad. Sci. USA*. 1978; 75(1):64–68. [PubMed: 343112]
71. Römer R, Hach R. tRNA conformation and magnesium binding. A study of a yeast phenylalanine-specific tRNA by a fluorescent indicator and differential melting curves. *Eur. J. Biochem*. 1975; 55(1):271–284. [PubMed: 1100382]
72. Nilsson L, Karplus M. Empirical energy functions for energy minimization and dynamics of nucleic acids. *J. Comp. Chem*. 1986; 7:591–616.
73. Schmidt MW, Baldrige KK, Boatz JA, Elbert ST, Gordon MS, Jensen JH, Koseki S, Matsunaga N, Nguyen KA, Su S, Windus TL, Dupuis M, Montgomery JA. General atomic and molecular electronic structure system. *J. Comp. Chem*. 1993; 14:1347–1363.
74. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. USA*. 2001; 98(18):10037–10041. [PubMed: 11517324]
75. Schafer A, Mark AE, van Gunsteren WF. Absolute entropies from molecular dynamics simulation trajectories. *J. Chem. Phys*. 2000; 113:7809–7817.
76. Anicricioaei I, Karplus M. On the calculation of entropy from covariance matrices of the atomic fluctuations. *J. Chem. Phys*. 2001; 115:6289–6292.
77. Glykos NM. Software news and updates. Carma: a molecular dynamics analysis program. *J. Comp. Chem*. 2006; 27(14):1765–1768. [PubMed: 16917862]

Fig. 2.

Structure of EF-Tu and sequence alignment colored by sequence identity across all three domains of life. Structure coloring is the same as the *T. aquaticus* line in the alignment. Blue is most conserved, and red is least conserved. The alignment shows a subset of sequences from the evolutionary profile constructed with QR cutoff 75 % (See Methods). Residues labeled on the EF-Tu structure are marked by “*” in the alignment with the exception of Glu55 which is blocking the view of Glu56. These residues are universally conserved and contribute substantially to interface energy shown in Fig. 5b. Glu390 is only conserved in *Bacteria* and is replaced by similarly charged Asp in *Archaea* and *Eucarya*. Other residues important for tRNA binding (Fig. 5a) are marked with “+”. Other conserved interface residues are marked with a dot.

Fig. 3.

Sequence Entropy and tuning of tRNA sequences. a) The sequence entropy per residue of the full evolutionary profile (all domains of life and all specificities) is plotted. Residues highly conserved across all tRNAs (U8, A14, G18, G19, U33, G53, U54, U55, C56, A58, and C61) have sequence entropy close to zero. Correlation of the tRNA sequence to amino acid specificity is measured by the difference in sequence entropy for all tRNAs (shown in (a)) with sequence entropy for a given specificity (see Methods). Entropy differences are shown for b) tRNA^{Cys} and c) tRNA^{Gly}. The nucleotides interacting with EF-Tu in the tRNA^{Cys} system are marked by black lines while the nucleotide that forms a base pair with one of the nucleotides on the interface is shown by thatched lines. The positions that are conserved and show tuning for tRNA^{Gly} and tRNA^{Cys} are marked with “*”. The bases corresponding to acceptor stem (AS), GG arm, anticodon arm (AC), and common arm (C) are marked. The CCA hairpin is universally conserved and not included in the graph. The last position shown corresponds to base 73.

Fig. 4. Comparison of RMSD plots for tRNA in complexes containing modified (right) and unmodified (left) bases with different Mg^{2+} concentrations. RMSD was calculated for the tRNA^{Cys} backbone and for subsections of the tRNA^{Cys} backbone. For each frame of the trajectory, the set of atoms under consideration was structurally aligned to the same set from the crystal structure.

Fig. 5.

a) Mean local nonbonded interaction energies of EF-Tu residues with Cys-tRNA^{Cys} averaged over the last 16 ns sampled every 20 ps. b) The same energies multiplied by sequence identity across all three domains of life. Only the strongest interactions are labeled

Fig. 6.

a) Mean local nonbonded interaction energies of tRNA nucleotides with EF-Tu. Averaging is the same as in Figure 5a. b) The same energies multiplied by sequence identity across all three domains of life and all tRNA specificities. Since nucleotides 17 and 47 are absent from this tRNA, values at those positions have been set to 0. The “x” at nucleotide 76 shows the interaction energy of A76 and the charged Cys together against EF-Tu. Since the second graph applies to all tRNA specificities, they do not share a common charged amino acid, and the “x” is not included.

Fig. 7.

a) Mean local nonbonded interaction energies of EF-Tu residues with the amino acid carried by tRNA^{Cys} for a) MMC (L-Cys), b) MMD (D-Cys), and c) LUS (Sep). Averaging is the same as in Figure 5a

Fig. 8. Residence of a) Mg^{2+} and b) K^{+} ions in MMC. Spheres mark starting locations from which an ion moved no further than $3\text{\AA}$ within a given period of time: red (2–4 ns), green (4–8 ns), blue (>8 ns). Overlapping spheres show residency sites that are the result of non-overlapping time windows. Occupancy of Mg^{2+} (red) and K^{+} (blue) is shown in c) LMC and d) HMC. The surfaces enclose volume elements occupied by ions for 5 % of the trajectory.

Fig. 9. Electrostatic isosurfaces generated by APBS and averaged over the last 5 ns of their respective trajectories. a) $\pm 1k_B T/e$ isosurfaces for a run with EF-Tu-GTP. b) $\pm 2k_B T/e$ isosurfaces for a run with tRNA^{Cys}. A white circle shows the positive region in b that corresponds to the circled negative region in a.

Fig. 10.
The ternary complex colored by mean squared deviation corresponding to the first principal component. The color scales are different for protein (0–0.5 Å²) and tRNA (0–3 Å²).

Table 1

Abbreviated names for all eleven simulations listed with their specific conditions.

	Name	Ion Conditions	Nucleobase Modification	Charged Amino Acid	Extra Modification
	control	crystal	modified	L-Cys	
	LUC	low			
	MUC	medium	unmodified	L-Cys	
	HUC	high			
0.8	LMC	low			
	MMC	medium	modified	L-Cys	
	HMC	high			
	LUS	low	unmodified	Sep	
	MMD	medium	modified	D-Cys	
	MMC _C	medium			Neutral Cys-tRNA ^{Cys} Cys amine
	MMC _{AU}	medium	modified	L-Cys	A51.U63 mutant

Table 2

Ionic conditions for all eleven simulations.

Run	Ionic Str (mM.e ⁻)	Mg ²⁺ Number	Mg ²⁺ Conc. (mM)	K ⁺ Number	K ⁺ Conc. (mM)
control	104.21	3	6.45	85	182.61
LUC	118.20	11	23.01	69	144.35
MUC	132.40	18	37.53	55	114.68
HUC	154.50	26	56.18	39	84.27
0.8 LMC	120.72	11	23.50	69	147.43
MMC	135.43	18	38.39	55	117.31
HMC	153.89	26	55.96	39	83.94
LUS	122.79	11	23.49	71	151.62
MMD	135.51	18	38.41	55	117.37
MMC _C	137.67	18	38.72	56	120.45
MMC _{AU}	135.51	18	38.41	55	117.37

Detailed MM-PBSA binding free energy values for MMC. The 95% confidence interval range for each value is $\pm$ the number shown below in parentheses.

Table 3

MMC	Complex	EF-Tu	tRNA	Difference
$\langle E_{vdw} \rangle$	-1714.51 (1.79)	-1827.27 (1.29)	244.67 (1.15)	-131.91 (0.35)
$\langle E_{elec} \rangle$	-101272.26 (8.40)	-66810.25 (3.55)	-37890.85 (6.68)	3428.85 (3.39)
$\langle G_{polar} \rangle$	-29448.89 (8.47)	-4568.79 (3.18)	-21553.33 (6.04)	-3326.77 (3.76)
0.8 $\langle E_{elec} + G_{polar} \rangle$	-130721.15 (6.31)	-71379.04 (1.74)	-59444.18 (5.46)	102.07 (1.33)
$\langle G_{nonpolar} \rangle$	151.93 (0.06)	96.73 (0.04)	73.59 (0.04)	-18.39 (0.02)
-TS	-1126.80	-846.30	-314.50	34.00
$\langle G_{binding} \rangle$	-119677.65 (7.24)	-65981.92 (3.31)	-53681.50 (5.78)	-14.23 (1.27)

MM-PBSA free energy differences in kcal/mol. The 95 % confidence interval range for each value is $\pm$ the number shown below in parentheses. Standard deviations for the $\langle \Delta G_{\text{binding}} \rangle$ were all 27–32 kcal/mol.

Table 4

	LUS	MMC _{AU}	HMC	MMC	HUC	LMC	MMD	MUC
$\langle \Delta E_{\text{vdw}} \rangle$	-136.93 (0.31)	-121.20 (0.34)	-126.36 (0.29)	-131.91 (0.35)	-134.61 (0.35)	-138.60 (0.31)	-141.85 (0.34)	-142.30 (0.33)
$\langle \Delta E_{\text{elec}} \rangle$	3822.72 (3.44)	3701.69 (4.09)	3470.30 (3.26)	3428.85 (3.39)	3647.85 (2.77)	3416.21 (2.87)	3469.12 (3.85)	3604.61 (3.06)
$\langle \Delta G_{\text{polar}} \rangle$	-3674.53 (3.33)	-3597.38 (4.23)	-3366.62 (3.40)	-3326.77 (3.76)	-3543.96 (3.01)	-3310.34 (3.00)	-3368.37 (4.13)	-3494.03 (3.16)
0.8								
$\langle \Delta E_{\text{elec}} + \Delta G_{\text{polar}} \rangle$	148.19 (1.41)	104.31 (1.27)	103.69 (1.28)	102.07 (1.33)	103.89 (1.22)	105.87 (1.26)	100.75 (4.13)	110.58 (1.28)
$\langle \Delta G_{\text{nonpolar}} \rangle$	-18.76 (0.02)	-17.61 (0.03)	-18.28 (0.02)	-18.39 (0.02)	-18.33 (0.02)	-17.92 (0.02)	-18.86 (0.02)	-18.38 (0.02)
-TΔS	33.20	34.50	33.50	34.00	33.50	34.20	34.70	23.50
$\langle \Delta G_{\text{binding}} \rangle$	25.70 (1.36)	0.01 (1.22)	-7.46 (1.25)	-14.23 (1.27)	-15.54 (1.19)	-16.45 (1.23)	-25.27 (1.23)	-26.60 (1.25)