



MASTERARBEIT | MASTER'S THESIS

Titel | Title

Characterization of molecular determinants of tRNAPhe
aminoacyl isomer recognition by EF-Tu using molecular
dynamics simulations

verfasst von | submitted by

Benjamin Bernrieder B.Sc.

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of
Master of Science (MSc)

Wien | Vienna, 2024

Studienkennzahl lt. Studienblatt | Degree
programme code as it appears on the
student record sheet:

UA 066 834

Studienrichtung lt. Studienblatt | Degree
programme as it appears on the student
record sheet:

Masterstudium Molekulare Biologie

Betreut von | Supervisor:

Dr. Anton Polyansky

Table of Contents

1. Acknowledgments	3
2. Abstract	5
3. Zusammenfassung	7
4. Introduction	9
4.1 Translation/Elongation cycle	9
4.2 Aminoacylation	10
4.3 EF-Tu/EF-1A	11
4.4 Binding affinity compensation	12
4.5 Molecular dynamics simulations	13
4.6 Information theoretic approaches to allostery	14
4.7 Configurational entropy	16
5. Aims and motivation	19
6. Methods	21
6.1 Molecular dynamics simulations	21
6.2 Trajectory analysis	22
6.3 Configurational entropy calculations	23
6.4 Mutual information analysis	23
7. Results	25
7.1 Trajectory overview	25
7.2 Clustering	26
7.3 Solvent accessible surface area of the amino acid carbonyl	27
7.4 Interactions between tRNA 5'- and 3'-ends	28
7.4.1 Hydrogen bonding	28
7.4.2 G1-A76 base stacking	31
7.4.3 Acceptor stem strand separation	32
7.5 Conformational dynamics	32
7.5.1 tRNA	33
7.5.2 EF-Tu	34
7.6 Mutual information	36
7.6.1 Mutual information from MD	37
7.6.2 Mutual information from MSA	43
8. Discussion	45
8.1 Impact of attachment isomerism on tRNA and EF-Tu	45
8.2 Limitations and outlook	49
9. References	51
10. Supplementary material	59

1. Acknowledgments

I would like to sincerely thank Bojan Žagrović for giving me the opportunity to study and do research in the laboratory for molecular biophysics over the course of my thesis work. His teaching was my first exposure to and started my interest in computational structural biology and molecular dynamics simulations. I am grateful for the opportunity to intern in his group leading to a position as a masters student and his feedback always opened new perspectives and considerations for my understanding. Special thanks to Anton Polyansky for supervising my work, who was always available to lend a hand when problems arose, and whose feedback and support has been instrumental in shaping not only this work but my understanding, appreciation and passion for molecular and computational biology. His patience, support and encouragement were invaluable and I am honored by his trust in me, to take me on as his student. His wealth of knowledge and the skills he taught me helped immensely in overcoming any difficulties that came up during my thesis and furthering my understanding of a multitude of topics, related and unrelated. Further I would like to thank all members of the Žagrović group for the welcoming and collaborative atmosphere I encountered in my time there. I would specifically like to thank Thomas Kapral for many rounds of table tennis, fun and interesting discussions as well as chess lessons and for making my time here that much more enjoyable. I would like to thank friends and family who were always supportive and encouraging during my studies and whose company always lifted my mood when pressure was mounting and deadlines were approaching. Special thanks to Jakob Sussmann and Fabian Seidlmeier, whose valuable feedback helped to considerably improve this thesis. I would like to thank the VBC5 and University of Vienna staff, especially Ivan Yudushkin, Iris Gruber and Renate Fauland for their patience, kindness and assistance in dealing with my questions and guidance in administrative matters. It has been a joy and a privilege being able to work on and learn about interesting and challenging questions surrounded by people of such intelligence, kindness and curiosity.

2. Abstract

Delivery of tRNA to the ribosome and accurate incorporation of amino acids in the nascent peptide chain are a cornerstone of translation in all organisms. In bacteria, all species of tRNA are bound by the GTPase elongation factor thermo unstable (EF-Tu) and delivered to the ribosome where EF-Tu hydrolyses GTP and dissociates upon recognition of cognate codon/anticodon pairs. The binding of EF-Tu and tRNA therefore plays an important part in translational fidelity. EF-Tu encounters a mixture of tRNA attachment isomers made by different classes of aminoacyl synthetases and acts as an isomerase unifying tRNA populations towards the 3'-aa-tRNA isomer. This work investigates the impact of the tRNA aminoacylation isomerism on conformation and dynamics of tRNA^{Phe} and the respective ternary complex with EF-Tu and GTP with atomistic microsecond molecular dynamics (MD) simulation. Interactions between the 3'- and 5'-end within the tRNA acceptor stem are found to be markedly different between attachment isomers resulting in varying levels of strand separation in the acceptor stem and solvent accessibility of the aminoacyl carbonyl group. Mutual information (MI) between degrees of freedom calculated from all atom MD simulations is used to quantify conformational coupling between amino acids in EF-Tu and bases in tRNA^{Phe}. MI between different residues in tRNA and EF-Tu from MD is compared to sequence-based co-variation from multiple sequence alignments. Clustering of MI between amino acids is able to recover functional domains in EF-Tu and suggests a clear distinction in how EF-Tu integrates information about differently aminoacylated bound tRNAs into its intramolecular communication.

3. Zusammenfassung

Der Transport von tRNA zum Ribosom und der korrekte Einbau von Aminosäuren in die entstehende Peptidkette ist ein fundamentaler Bestandteil der Translation in allen Organismen. In Bakterien werden alle aminosäurespezifischen tRNA Moleküle von der GTPase elongation factor thermo unstable (EF-Tu) gebunden und an das Ribosom transportiert. Dort wird, nach erfolgreicher Erkennung von passendem Codon und Anticodon, GTP abgespalten und EF-Tu dissoziiert. Das Binden von tRNA und EF-Tu spielt deshalb eine wichtige Rolle bei der Gewährleistung der Genauigkeit der Translation. Ursprünglich begegnet EF-Tu einer Mischung von tRNA Aminoacylierungsisomeren und agiert als Isomerase um dieses Verhältnis zugunsten des 3'-aa-tRNA^{aa} Isomers zu vereinheitlichen. Die vorliegende Arbeit untersucht die Auswirkungen des natürlich vorkommenden tRNA Beladungsisomerismus auf die Konformation und Dynamik von tRNA^{Phe} und den entsprechenden Komplexen mit EF-Tu mithilfe von Simulationen der molekularen Dynamik in Mikrosekundenlänge. Interaktionen zwischen den 3'- und 5'-Enden des tRNA Akzeptorstammes zeigen deutliche Unterschiede und resultieren in unterschiedlicher Trennung der Strangenden und der Zugänglichkeit des Lösungsmittels für die Aminoacyl-Carbonyl-Gruppe. Transinformation zwischen atomaren Freiheitsgraden, berechnet aus Molekulardynamik Simulationen wurden verwendet um die Kopplung von Basen in tRNA^{Phe} und Aminosäuren in EF-Tu zu quantifizieren. Transinformation zwischen Positionen in Sequenzalignments wurden mit den Kopplungen, die in den Molekulardynamik Simulationen beobachtet wurden, verglichen. Die Clusteranalyse der Transinformation zwischen den EF-Tu Aminosäuren ist in der Lage funktionelle Domänen des Proteins wiederzugeben und impliziert einen erkennbaren Unterschied in der Art wie EF-Tu Informationen über gebundene tRNA Aminoacylisomere in seine intermolekulare Kommunikation integriert.

4. Introduction

4.1 Translation/Elongation cycle

To accomplish a wide range of catalytic, regulatory, signaling, and structural tasks, genetic material is constantly transcribed into mRNA and subsequently translated into polypeptides. Accuracy and speed of this process are crucial for all organisms and dysregulation or disruption can easily lead to non recoverable reduction in cellular functionality and fitness. This, in fact is the mechanism of many antibiotic agents and a hallmark of cancer and some neurological conditions¹. A complex cellular machinery achieves this feat leveraging a host of components. Central to translation is the ribosome, a two subunit complex comprising ribosomal RNA and ribosomal proteins. The small subunit binds the messenger RNA (mRNA) transcript and mainly facilitates the translation of triplet codons into a polypeptide chain. The large subunit performs the catalytic functions while also binding aminoacylated transfer RNAs (tRNAs). tRNAs connect the triplet code of the mRNA to their corresponding amino acid and deliver it to the ribosomal A-site and subsequently to the nascent peptide chain in the ribosomal P-site². During protein synthesis, aminoacyl tRNAs cross the entire ribosome by binding to the three primary interaction sites (A, P and E), together facilitating the elongation of a peptide chain³. As the link between codon and amino acid, tRNA's conformational flexibility and dynamic behavior have been shown in recent experimental and computational work to play a vital role in all main steps of translation^{4–11}. During the initiation phase of 70s complexes, the initiator tRNA is held in its characteristic place by interactions between the P-site of the small subunit and the anticodon stem, and another between the acceptor stem and the initiation factor 2 (IF2)^{4,12,13}. Comparison of tRNA structures from different eukaryotic initiation complexes has revealed a high degree of flexibility of the acceptor arm in particular and may point to a regulatory aspect of tRNA flexibility on assembly of ribosomal sub-units and incorporation into the P site^{4,14}. For the next major step of translation, incorporation of the correct elongator aa-tRNA^{aa} is crucial. This step is regulated via a kinetic discrimination mechanism¹⁵, where upon codon anticodon binding, conformational changes in the 30S subunit accelerate the forward rates of GTP hydrolysis in EF-Tu, leading to accommodation only of cognate elongator species¹⁶. Analysis of aa-tRNA cryo-EM structures in complex with GTPase activated EF-Tu proxies (antibiotic kirromycin leads to EF-Tu mimicking the activated configuration⁵) have detected anticodon and D-stem loop region distortions in the tRNA^{6–8}. Together with the mandatory presence of an intact aa-tRNA to trigger the GTPase activation⁹, these results point to an active role for aa-tRNA in tuning the GTPase activation by EF-Tu ribosome interaction. FRET studies involving pre-translocation complex analogs and either wildtype or mutagenized tRNA, discovered specific tRNA-ribosome interactions regulating the translocation intermediate complex formation. This

suggests a crucial role of tRNA conformational flexibility in pre-translocation complex dynamics¹⁰.

4.2 Aminoacylation

An important part of the translation elongation cycle is the charging and subsequent delivery of the aminoacylated elongator tRNA to the ribosome. Aminoacylation is facilitated by aminoacyl-tRNA-synthetases (AaRS). They are structurally divided into two classes I and II and while functionally these classes overlap heavily via their primary site of attachment, there is one exception in PheRS. In general, class I aaRSs transfer the amino acid to the 2'-OH of the ribose of A76^{17,18} while class II aminoacylate the 3'-OH¹⁹. PheRS, a class II aaRS however attaches at the 2'-hydroxyl group of the terminal adenosine²⁰. Charged aa-tRNA^{aa}'s subsequently form a ternary complex with eefFAA (in eukaryotes) or EF-Tu (bacteria) and GTP and are delivered to the ribosomal A-site²¹. EF-Tu/eefFAA stabilizes the 3'-isomer and shuttles it to the ribosome exclusively in this form^{22,23}. The 3'- and 2'-hydroxyl groups are involved in many different functions across the whole translation cycle including peptide bond formation, proofreading and tRNA charging and the translation process at many steps has been demonstrated to stabilize and specifically exploit either isomer²⁴. The replacement of even the non-connective hydroxyl with either a hydrogen or a fluorine has been demonstrated to completely abolish the deacylation of mischarged tRNA^{ILE}²⁵. Importantly, the 2'- and 3'-stereoisomers have been shown to interconvert via an orthoester (OE) intermediate (Figure 1A) in water with the 3'-isomer being ~3 times more common²⁶. The majority of recent structural data supports a stabilization of the 3'-isomer by EF-Tu with some evidence from NMR-studies reporting the presence of the OE-conformation in the EF-Tu bound state for valyl-tRNA²⁷. While both 2'- and 3'-isomers are capable of binding EF-Tu^{28,29}, they do so with radically different affinities and hydrogen bonding between the 2'-OH and EF-Tu might be required to stimulate GTPase activity³⁰. Analysis of aminoacyl bond half-lives showed a marked difference in the stability of the ester bond between attached amino acids. This difference was overridden by EF-Tu protecting the linkage against hydrolysis³¹. Previous molecular dynamics (MD) studies on aa-ester bond exposure from charged acceptor triplets implicate the nature of the attachment site to be potentially important for the hydrolytic stability of the aa-ester bond. This study observed the invariant adoption of clear extended and collapsed states for trinucleotides in both isomers as well as a systematically higher solvent exposure of the aminoacylester bond in the 3'-attached systems³². This effect is studied in the context of full length tRNA^{Phe} in the present work with the inclusion of the intermediate OE.

4.3 EF-Tu/EF-1A

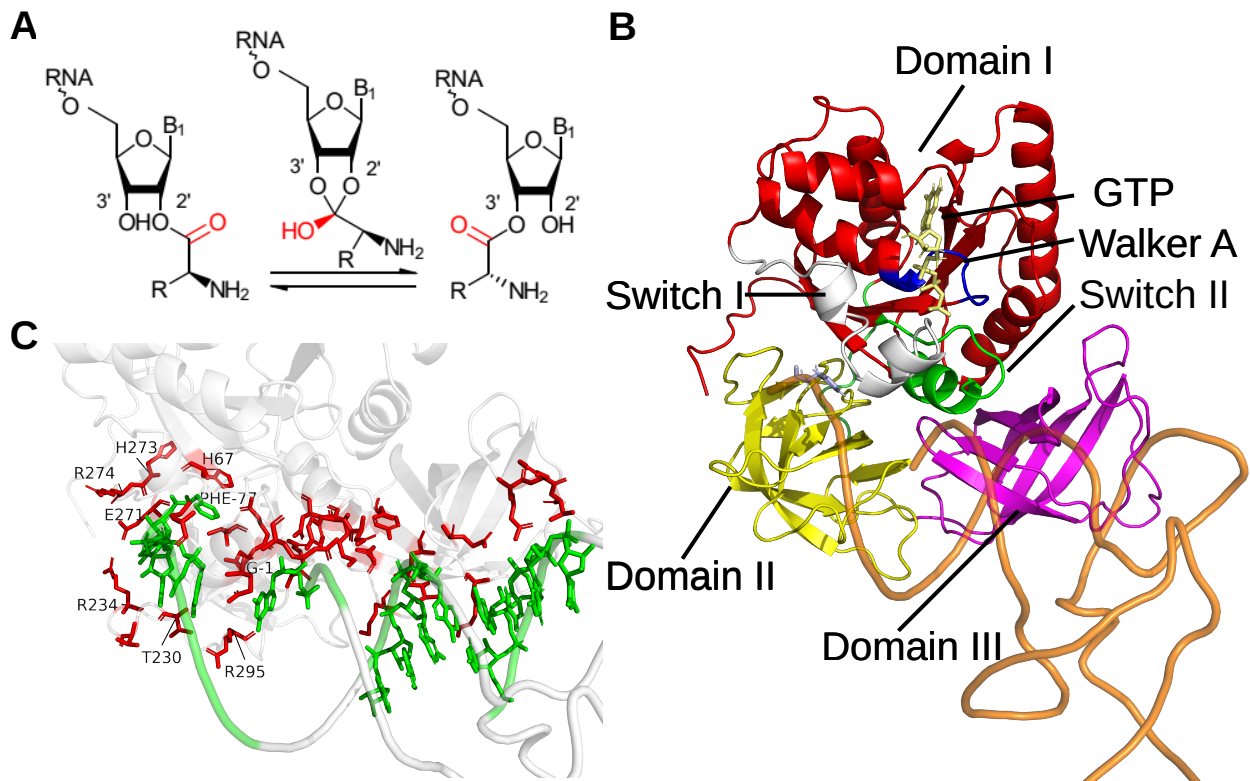


Figure 1: tRNA charging isomerism and structure of the Phe-tRNA^{Phe}.EF-Tu-GTP complex. **A** Charging isomerism at the ribose of the terminal adenosine in tRNA with the intermediate OE. Carbonyl/carboxyl group colored red. **B** EF-Tu domains. EF-Tu comprises three domains with domain I binding GTP and incorporating classical nucleotide binding motifs such as the walker-A/P-loop motif. Domains II and III are responsible for tRNA binding, ribosomal interactions and binding of nucleotide exchange factors after GTP hydrolysis. Upon codon/anticodon recognition at the ribosome a conformational change is initiated resulting in GTP hydrolysis. Both switch regions are instrumental for this conformational change. **C** Binding interface of *S. cerevisiae* Phe-tRNA^{Phe}, *T. aquaticus* elongation factor EF-Tu-GDPNP ternary complex (PDB code: 1TTT). Proposed interacting EF-Tu residues are colored in red and corresponding tRNA residues in green. Important residues at the acceptor stem interface are labeled.

Translation is dependent on rapid and accurate delivery of aminoacylated tRNA to the ribosome. After recognition and aminoacylation of a diverse set of tRNA species by their cognate tRNA synthetases, delivery of aminoacylated elongator tRNA to the ribosomal A-site is facilitated through the translation factor EF-Tu in bacteria³³ and its eukaryotic homologue EF-1A via a ternary complex comprising the aminoacyl tRNA (aa-tRNA) and the GTP-bound form of EF-Tu/EF-1A. EF-Tu is charged via EF-Ts with GTP³⁴ after which it binds aa-tRNA in a ternary complex. This complex binds the ribosomal A-site during decoding and upon codon/anticodon recognition a conformational change in both the ribosome and the ternary complex results in GTPase activation and subsequent GTP hydrolysis.^{16,35–44} This leads to the full accommodation of aa-tRNA and the dissociation of EF-Tu from the ribosome⁴⁵. EF-Tu can also dissociate reversibly without GTP hydrolysis⁴⁶ and contributes to translational fidelity via a steric component, by reducing free energy barriers associated with early accommodation steps⁴⁷. EF-Tu-GDP then enters the cycle at the beginning and

undergoes EF-T dependent nucleotide exchange. A critical step in the EF-Tu dependent decoding path is the recognition and uniform binding of elongator aa-tRNA's. Regardless of the esterified amino acid, all aa-tRNA's need to be bound tightly enough for secure translocation to the ribosome and weakly enough so as not to interfere with its release from the ternary complex during decoding^{48,49}. All these considerations require a complex network of communication between tRNA, ribosome and EF-Tu to facilitate fast accurate and uniform delivery of the ternary complex to the ribosome and trigger GTP hydrolysis in the presence of cognate codon/anticodon moieties at the ribosomal A-site. The exact mechanism of triggering the necessary conformational changes after decoding is only partially understood. EF-Tu from *T. thermophilus* consists of three domains (Figure 1B): domain I (1-210), a nucleotide binding domain similarly found in all GTPases including a Walker A (P-switch) motif as well as two switch regions responsible for conformational changes upon GTPase activation^{35,50-52}. Domains II (211-310) and III (311-405) make extensive contacts with tRNA, while the C-terminal domain III is responsible for binding EF-T's during nucleotide exchange⁵³. EF-Tu's affinity for GDP is about two orders of magnitude higher than for GTP, a difference that disappears when removing domains II and III implicating them in allosteric regulation of EF-Tu function⁵¹ as they do not directly contact the nucleotide binding site. EF-Tu interactions with tRNA^{Phe} can be divided into three groups: (1) interactions with the tRNA 3'-acceptor triplet and esterified amino acid, (2) interactions with the 5'-end acceptor stem, (3) interactions with the tRNA T-stem (Figure 1C). A pocket between domains I and II fits all 20 amino acids with hydrophobic sidechains lining the pocket and D226 and E227 resulting in a net negative charge of the pocket. All three domains are in contact with the tRNA acceptor arm while except for one contact between the K90 residue of domain I and G65, only domain III interacts with the T-stem. Almost all interactions consist of EF-Tu residues forming hydrogen bonds and ion pairs with 2'-hydroxyl's and phosphates of the tRNA T-stem and acceptor stem regions. Notably, H67 stacks on top of the Phe sidechain ring while E271 stacks to the free 2'-OH of the A76 sugar. K52 from the switch I region contacts the two phosphates at positions 74 and 75 forming the only contact of switch I with the CCA triplet. K90 interacts with phosphates of both acceptor stem base 2 and T-stem base 65. All other contacts between domain I and tRNA are located at the 5'-end of the acceptor stem between tRNA residues 1 and 4³⁵.

4.4 Binding affinity compensation

Although all species of elongator aa-tRNA bind EF-Tu with similar affinities⁵⁴⁻⁵⁷, some small specificity remains. This might be required to account for intracellular concentrations of tRNA and indeed the inverse abundance correlates well with these differences in *E. coli*⁵⁸. Misacylation abolishes the uniform affinity leading to a range of 60 times weaker to 120 times tighter binding⁵⁹. The discrete contribution of the

esterified amino acid is compensated by the tRNA body and systematic misacylation studies show good matching of tight binding amino acids with weakly binding bodies^{60,61}. Binding assays of EF-Tu with alanine-mutated residues at 20 positions interfacing with tRNA revealed 5 positions whose effects on tRNA binding affinity were specific to the tRNA species. 3 of these positions interacted with T-stem regions of variable sequence, while the rest interacted with conserved positions^{61,62}. This suggests both a contribution of sequence specific contacts, as well as of the overall tRNA structure and/or dynamics to modulating binding affinity. Measurements of binding affinities for 42 single base-pair mutations in the acceptor and T-stem of *S. cerevisiae* tRNA^{Phe} to *T. thermophilus* EF-Tu yielded a set of three base pairs (49–65, 50–64 and 51–63) responsible for a large part of the specificity of binding. Inserting these base pairs from several *E. coli* tRNA's into yeast tRNA^{Phe} resulted in EF-Tu binding affinities comparable to those of the respective donor tRNA. Additional experiments with double and triple base-pair mutants suggested additive and independent thermodynamic contributions of each position to the binding, enabling reasonably accurate prediction of binding affinities for all *E. coli* tRNA's⁶³. Mismatches in the acceptor stem 1-72 base pair have been shown to contribute to poor binding of initiator tRNA^{64–66}.

4.5 Molecular dynamics simulations

Common experimental methods for investigating the structural basis of molecular processes are significantly limited when it comes to spatial and temporal resolution. Additionally, dynamic systems are often insufficiently captured by single configurations and better understood at the level of ensembles. Molecular dynamics (MD) simulations are able to provide atomistic detail as well as femtosecond-level temporal resolution in addition to sampling the configurational space. In basic terms, MD simulations predict the movement of atoms based on a model of inter atomic interaction forces⁶⁷. Positions and velocities for each atom at a given time during the MD simulation are recalculated in accordance to the forces acting on each atom. These forces are the result of a predefined potential energy function and are used to compute acceleration of an atom and its resulting velocity (assumed to stay constant during an integration step). Different sets of parameters together with the functional form describing the interaction Hamiltonian, termed force fields, have been developed over the years, fit to different quantum mechanical calculations and experimental data⁶⁸. MD simulations allow for careful control and exact knowledge of parameters such as pressure, temperature, initial conformation, presence of ions, ligands and post translational modifications or mutations. Increased availability of computational resources both in terms of CPU and GPU⁶⁹ computation have made microsecond MD trajectories widely available for systems of moderate size such as tRNA bound to EF-Tu (~30000 solvent atoms and 8882 atoms in the Phe-tRNA^{Phe}·EF-Tu·GTP ternary complex simulation box). The ability of MD simulations to probe equilibrium behavior of such systems in response to subtle modifications such as amino acid attachment to a tRNA, makes them

uniquely suited for investigation of conformational and dynamic properties of this part of the translation machinery. Covalent modifications to bases in tRNA occur with high prevalence⁷⁰ and are implicated in tuning conformation, folding dynamics and translation accuracy^{71–74}. Modified bases are nonstandard residues and as such need to be manually parameterized for MD simulations of systems containing tRNA. The partial charges for nonstandard tRNA^{Phe} residues were derived previously by Zhang et al.⁷¹ via the RESP approach and the missing parameters were manually assigned in analogy to the GAFF force field⁷⁵ as well as from the online AMBER parameter database⁷⁶. Comparison of modified and unmodified tRNAs from MD simulations the influence of modified bases on tRNA flexibility, increasing in some areas and decreasing in others⁷¹. The ternary complex of tRNA·EF-Tu·GTP has been previously investigated via MD simulations. Explicit solvent MD simulations of the ternary complex found a rapid conversion of the EF-Tu switch I region during aa-tRNA accommodation. Switch I converts from an α -helix to a β -hairpin in the GDP conformation of EF-Tu⁷⁷. Additionally, MD simulations revealed the loss of interactions between the nucleotide and the switch I upon GTP hydrolysis to be responsible for the increase in switch I flexibility necessary for this transition. This loss of contact was theorized to enable a rotation of EF-Tu domain I relative to domains II and II, allowing for release of the tRNA CCA tail bound between domains I and II⁴⁸. Tying into previous MD studies of acceptor triplet conformation and aminoacyl ester bond solvent exposure in trinucleotides aminoacylated with several different amino acids at either the 2'- or 3'-OH, this work aims to investigate conformational trends in full length tRNA as well as tRNA in complex with EF-Tu. To further characterize the molecular determinants of amino acid attachment isomerism in tRNA and the tRNA^{Phe}·EF-Tu·GTP ternary complex, the intermediary OE along with both Phe-tRNA^{Phe} charging isomers is additionally investigated in this work. Here, the dynamic properties of these systems are investigated via microsecond MD simulations and further explored using original information theory approaches for configurational entropy estimations and allosteric network analysis (see below).

4.6 Information theoretic approaches to allostery

Information theoretic approaches for analyzing and predicting allosteric networks have been implemented in the last couple of years both from multiple sequence alignments^{78,79} as well as MD simulation^{80–84}. Covariation of residues in multiple sequence alignments (MSA) has been widely used in the field of structure prediction to great success, most recently in machine learning approaches such as AlphaFold⁸⁵. Covariance of positions in MSA comprises several influences such as phylogenetic, structural, functional and random components⁷⁹. In the case of known structures, mutual information (MI) between residues can be analyzed with a focus on functional/allosteric coupling between residues. Using this approach indirect correlations between far apart

residues can be discovered from multidimensional mutual information (mdMI) from sequence data⁸⁶. These couplings can be thought of as arising not from direct contacts or indirect geometric/steric constraints but could instead be indicative of functional or allosteric networks.

As all atomic motions are present in MD trajectories, MI between all degrees of freedom can be calculated and this is in fact a necessary step for the calculation of configurational entropy via the mutual information expansion or maximum information spanning tree approach (see Section 4.7). From this, residue pairs with correlated degrees of freedom can be identified in an equilibrium ensemble. Smaller perturbations to a biomolecule involve the activation of the same fluctuation pathways as ones activated by random collisions present at equilibrium⁸⁷ and correlations observed in this way, may propagate structural or dynamic perturbations from external forces such as ligand or interaction partner binding^{88,89}. The mechanisms through which a perturbation at one site can couple to another include direct steric or electrostatic interactions as well as indirect action such as solvent reordering or networks of correlated conformations⁸⁰. MI values between degrees of freedom calculated from MD simulations suffer from spurious covariation and uneven/incomplete sampling. Additionally the merging of degrees of freedom into biologically more meaningful units (such as amino acids or nucleotides) leads to a small loss of data at the edges of these specified units as well as a relative distortion of the self-coupling due to different unit size. As MI asymptotically approaches its real value when increasing sampling, it is possible to predict some of the systemic noise layered over the true pattern of mutual information at incomplete sampling as it is implemented in the recently developed software ARTEMIS⁹⁰ for the analysis of MI maps (2D pairwise matrices) obtained from MD ensembles. MI maps intuitively can be thought of as a measure of similarity between residues and as such can be used for graph-based clustering/community detection methods when thinking of the molecule as an undirected weighted graph with residues (vertexes) connected by edges of varying weight (MI). This allows for the partitioning of such a graph into communities of residues sharing high internal MI while making (depending on the method used for partitioning) very few assumptions about the size and identity of these communities. One approach is the optimization of modularity in partitioning such graphs as implemented by the Leiden algorithm. This requires only treatment of inconsistent self edges in the graph, arising from collapsing individual degrees of freedom into residues of different atom numbers. Another approach is spectral clustering although in this approach the number of clusters needs to be specified beforehand. Both of these approaches have been applied in this work (see Section 6.4 for details).

4.7 Configurational entropy

Gibbs free energy change in the NPT ensemble comprises both enthalpic and entropic contributions and combines them into the key thermodynamic measure of the likelihood of a reaction. The entropy change in a biomolecular process has been shown to play an important part in some systems^{91–94} and consists of spatial and momentum components, the latter of which is usually neglected at constant temperature⁹⁵. Equally, the rotational and translational degrees of freedom are intentionally not considered as their contribution to the spatial component of entropy is also constant^{95–97}. Calculation of configurational entropy from MD simulations poses a problem as the estimation of the configurational probability density function requires a full coverage of the configurational space, a task computationally intractable with current resources for typical biomolecules. One approach to approximate configurational entropy is the maximum information spanning tree (MIST)^{97,98}. In this approach the amount of coupling between different degrees of freedom (correlation) is expressed as mutual information terms used to correct the uncorrelated entropy value. The formula for the joint entropy of two degrees of freedom X_1 and X_2 :⁹⁹

$$S_N(X_1, \dots, X_n) = \sum_{i=1}^N S_i(X_i | X_1, \dots, X_{i-1}) \quad (1)$$

can be rewritten including MI terms

$$I_N(X_1; \dots; X_N) = \sum_{i=1}^N [S_i(X_i) - S_i(X_1, \dots, X_{i-1})] \quad (2)$$

as¹⁰⁰:

$$S_N(X_1, \dots, X_n) = \sum_{i=1}^N [S_i(X_i) - I_i(X_i; X_1, \dots, X_{i-1})] \quad (3)$$

Computational restrictions on sampling of the probability density function put a limit on the order of included correlations. For this reason only correlations up to second order are considered. Truncated at second order the expression takes the form:

$$S_N(X_1, \dots, X_N) \approx \sum_{i=1}^N [S(X_i) - I_2(X_i; X_{j \in \{1, \dots, i-1\}})] \quad (4)$$

Of note is that this expression offers a choice of MI term. As the second order mutual information term can never be higher than the original term including all the degrees of freedom, this truncation constitutes an upper bound to the exact value¹⁰⁰.

Finding and choosing the lowest value for this upper bound results in the following expression:

$$S_N^{MIST_2}(X_1, \dots, X_N) = \sum_i^N [S(X_i) - \max_{j \in (1, \dots, i-1)} I_2(X_i; X_j)]$$

(5)

with the superscript denoting the second order of the MIST approximation. Notably, this maximization of the second order MI terms eliminates the problem of near-zero terms accumulating and therefore significantly improves the convergence of this approximation¹⁰⁰. This expression is equivalent to finding the minimum information spanning tree and can be implemented using the Prim's algorithm¹⁰¹, with the slight change from a distance to a closeness measure and therefore arriving at the maximum spanning tree. MI between degrees of freedom or residues in biomolecules has been put to use for many different use cases, such as contact prediction analysis of protein dynamics and allosteric prediction as described in the previous section.

5. Aims and motivation

The present work aims to further study the conformational and dynamic implications of the 3’-/2’-aminoacyl attachment isomerism in Phe-tRNA^{Phe} from MD simulations. This question is motivated by earlier work by Polyansky et al., showing differences in the solvent exposure of the aminoacyl ester bond between isomers and conformational preferences for both 2’- and 3’-isomers. As this study investigated these differences in aminoacylated trinucleotides using MD simulations, the natural next step, performed in this work, involves the addition of full length tRNA^{Phe}, the intermediate 2’/3’ orthoester and the protein binding partner of tRNA in early bacterial translation, EF-Tu. Additionally, this work aims to elucidate the dynamical properties of both tRNA^{Phe} and EF-Tu and how the aminoacylation isomerism impacts them. To this end residue level fluctuations as well as configurational entropies are calculated and compared. The third objective is to characterize differences in residue-residue pairwise mutual information (MI) between the simulated systems to qualify and quantify if and how EF-Tu and tRNA^{Phe} dynamics and intermolecular communication are tuned by attachment isomerism. To accomplish this, all atom microsecond MD simulations for all isomers and non-acylated tRNA^{Phe} in the free and EF-Tu bound state were performed. Specifically, conformational clustering was used to identify the molecular determinants of conformational differences and MI between degrees of freedom was calculated to characterize intramolecular couplings in the context of differential communication in tRNA and EF-Tu depending on isomerism and binding states.

6. Methods

6.1 Molecular dynamics simulations

Molecular dynamics (MD) simulations and analysis were performed using the GROMACS¹⁰² 5.1.4 and 2023.2¹⁰³ packages. Initial structures for yeast tRNA^{Phe}, EF-Tu from *T. aquaticus* and 3'-Phe-tRNA^{Phe}·EF-Tu were taken from the PDB with the codes 1EHZ¹⁰⁴, 1EFT¹⁰⁵ and 1TTT³⁵, respectively. After energy minimization, the prepared systems were subjected to equilibration and productive MD runs of 1 μ s, with one frame output every 5 ps. The AMBER99SB-ILDN¹⁰⁶ force field was used for EF-Tu, while an adapted force field from literature⁷¹ was used to model the tRNA^{Phe}, both times in conjunction with the TIP3P¹⁰⁷ water model. Ester bound tRNA bases and the GTP nucleotide were parameterized using the Antechamber module of the AmberTools20 package¹⁰⁸. For 3'-, 2'- and orthoester-bound Phe, geometry optimization and RESP¹⁰⁹ charge calculations were performed with the Gaussian 16 program¹¹⁰ at the B3LYP/6-31+G(d,p) level to be consistent with the employed tRNA force field. Individual charged tRNA^{Phe} molecules were placed inside a cubic simulation box of 10 nm size solvated with ~50000 TIP3P water molecules with 150 mM sodium-chloride concentration. tRNA^{Phe} molecules in complex with EF-Tu were similarly placed in a 12 nm cubic simulation box, filled with ~30000 solvent molecules and identical salt concentration. The timestep for integration was 2 fs. Particle mesh Ewald¹¹¹ was used for electrostatics (real space cutoff 12 and 1.2 Å grid with fourth-order spline interpolation) while a pressure of 1-1.013 atm (Parrinello-Rahman barostat¹¹²) and 310 K (nose-hoover thermostat¹¹³) was maintained with a compressibility of 4.5×10^{-5} bar⁻¹ for the barostat. Temperature and pressure coupling was implemented using 0.5/10 ps time constants respectively. The LINCS algorithm¹¹⁴ was employed to constrain bond lengths. 3D periodic boundary conditions were employed, as well as a twin-range (10/12 Å) spherical cut-off to truncate van der Waals interactions. For EF-Tu as well as the canonically attached 3'-Phe-tRNA^{Phe} systems, few independent MD runs (replicate number denoted in parentheses) were simulated for a total of 3 and 4 μ s of simulated time, respectively. Additionally, two MD trajectories with a timestep of 1 ps were prepared for EF-Tu and 3'-aminoacylated tRNA^{Phe}. An overview of all MD trajectories can be found in Table 1.

Table 1: The simulated systems.

trajectory	EF-Tu present	tRNA present	amino acid	output step	Simulated time
tRNA ^{Phe}	no	yes	/	5 ps	1 μ s
3'Phe-tRNA ^{Phe} (1)	no	yes	3'-attached	5 ps	1 μ s
3'Phe-tRNA ^{Phe} (2)	no	yes	3'-attached	5 ps	1 μ s
3'Phe-tRNA ^{Phe} (3)	no	yes	3'-attached	1 ps	1 μ s
2'Phe-tRNA ^{Phe}	no	yes	2'-attached	5 ps	1 μ s
OE-Phe-tRNA ^{Phe}	no	yes	OE-attached	5 ps	1 μ s
tRNA ^{Phe} .EF-Tu	yes	yes	/	5 ps	1 μ s
3'Phe-tRNA ^{Phe} .EF-Tu	yes	yes	3'-attached	5 ps	1 μ s
2'Phe tRNA ^{Phe} .EF-Tu	yes	yes	2'-attached	5 ps	1 μ s
OE-Phe-tRNA ^{Phe} .EF-Tu	yes	yes	OE-attached	5 ps	1 μ s
EF-Tu (1)	yes	no	/	5 ps	1 μ s
EF-Tu (2)	yes	no	/	5 ps	1 μ s
EF-Tu (3)	yes	no	/	5 ps	1 μ s
EF-Tu (4)	yes	no	/	1 ps	1 μ s

6.2 Trajectory analysis

For calculations of the solvent accessible surface area (SASA) the *sasa* module from GROMACS version 23.2 was used in conjunction with a probe radius of 1.4 Å and 196 dots per sphere. This utility uses an algorithm published by Eisenhaber et al.¹¹⁵ The solvent accessible surface area for the carbonyl group of the amino acid attached to the adenosine 76 of the tRNA^{Phe} was calculated for all Phe-tRNA^{Phe} and Phe-tRNA^{Phe}.EF-Tu acylation-isomer MD simulations (Table 1). The surface calculations were performed every 100 ps of simulation time. The values were divided by the maximum observed value of 42 Å² (3'Phe-tRNA^{Phe}) and plotted in a normalized histogram using *xmgrace*¹¹⁶ and 100 bins per histogram. To check the MD trajectories for major opening or closing events and to gauge the stability of EF-Tu bound tRNA^{Phe} complexes, the radius of gyration (RGYR) was computed. RGYR over the simulation time was calculated using the GROMACS *gyrate* utility for every frame. For calculations of the RGYR, all atoms were considered. The root-mean-square-deviation (RMSD) between pairs of backbone atoms of the average structure and structures over the simulated time was calculated using the GROMACS *rms* tool for both tRNA and EF-Tu components of a system. Fitting was performed accordingly over backbone atoms of the respective molecule. The calculations were performed every 5 ps of simulated time. The root-mean-square fluctuation (RMSF) around the starting structure along the MD trajectories was calculated for each residue as an average of its backbone atoms. Calculations were performed with the GROMACS *rmsf* module. MD conformational ensembles were clustered using single linkage clustering approach implemented in the GROMACS *cluster* utility adding structures to a cluster when its RMSD is within a cutoff to any element in a cluster. Conformational clustering was performed for tRNA backbone atoms (cutoff 1.85 Å) and residues G1, A76 and amino acid backbone atoms (cutoff 0.5 Å). A time step of 500 ps was employed resulting in 2000 frames considered per trajectory. tRNA backbone RMSD from the cryo-EM reference states (Loveland et al. 2017⁴⁴) was calculated for every frame of full MD trajectories and for every frame

contained in the three dominant clusters obtained by clustering. H-bonds presence was computed between the amino group of the amino acid and the phosphate group of residue G1 and non-connective hydroxyl groups of A76 (both hydroxyl groups in case of non-acylated tRNA^{Phe}) and phosphate G1. For this purpose the *hbond* utility was used with a hydrogen-donor-acceptor angle cutoff of 30° (0° representing an extended arrangement) and a distance cutoff donor-acceptor of 3.5 Å. The percentage of hydrogen bonding was calculated for all MD trajectories and three dominant clusters. Stacking interactions between the base rings of A67 and G1 were quantified using a distance cutoff between ring planes and an angle range between ring plane normal vectors. Distances were calculated using the GROMACS *pairdist* tool with a cutoff of 4 Å. Normal vector angles were calculated using the *gangle* utility and applying a cutoff range of 165°-180° and 0°-15°. Frames satisfying both criteria were counted as stacking and percentage of stacking interactions during a trajectory was computed for all MD trajectories and three dominant clusters. Minimum distances between the G1 phosphate group and the amino group and the A76 hydroxyl groups respectively were calculated using the GROMACS *pairdist* module. Morphs between central structures of the three dominant conformational clusters from clustering by G1, A76 and amino acid backbone to their respective EF-Tu bound counterpart were produced using the RigiMol plugin of PyMol¹¹⁷ with 100 states and 3 refinement cycles. All structure visualizations were rendered using PyMol.

6.3 Configurational entropy calculations

Configurational entropy calculations were performed using the maximum information spanning tree (MIST) based approach implemented in the PARENT¹⁰⁰ software package for all MD trajectories and the tRNA and EF-Tu subsets of atoms for MD trajectories where both were present. For calculations of differences in configurational entropy between uncharged and charged tRNA isomers, entropies were normalized by degrees of freedom (dof) and rescaled to the target system after subtraction:

$$\Delta S = \left(\frac{S_{MIST}^X}{dof^X} - \frac{S_{MIST}^{tRNA}}{dof^{tRNA}} \right) \times dof^X \quad (6)$$

with the superscript X denoting the respective charged tRNA^{Phe} isomer.

6.4 Mutual information analysis

Calculations of mutual information (MI) for all residue pairs truncated at second order were performed using the PARENT software suite and analyzed using the ARTEMIS framework⁹⁰ (manuscript in preparation). The automatic systemic noise reduction feature implemented in ARTEMIS was used on the raw data generated by PARENT with a target time step of ops. The implemented systemic noise reduction is achieved by calculating the ΔMI between each residue pair for a trajectory and its artificially

sampling deficient counterpart and subtract this noise pattern from the original trajectory after rescaling it. Considering a trajectory of 2×10^5 frames, recalculating the total MI for this trajectory with decreasing numbers of frames reveals a dependency of MI on the number of frames (sampling), where fewer frames lead to higher MI as spurious covariation increases. Using this dependency, the difference between two patterns of MI (one from maximum available sampling and one artificially decreased in sampling) can be rescaled and subtracted from the original data to arrive at an approximation of the MI calculated from a perfectly sampled ensemble. MI as a similarity measure was used to construct an undirected, weighted graph and clustered via spectral clustering¹¹⁸ as implemented in the scikit-learn python library¹¹⁹. The number of clusters was determined according to the elbow of the k-means inertia graph. 3 and 4 clusters were chosen for tRNA and EF-Tu MD trajectories respectively while 5 clusters were calculated for full EF-Tu·tRNA^{Phe} systems. The k-means clustering on the graph Laplacian of the first k eigenvectors was performed with 1000 initializations to counteract differences stemming from random seed initialization and the consensus clustering of these seeds was chosen for visualization. Additionally, MI for each residue was normalized to obtain a similarity measure sMI according to (7).

$$sMI_{i,j} = \frac{MI_{i,j}}{\max(MI_{i,i}, MI_{j,j})}$$

(7)

This transformation bounds the MI values and treats inconsistencies in the MI of a residue with it self, as the grouping of degrees of freedom into residues creates differing values for self loops in a graph constructed from them. The resulting similarity matrix was used to construct an undirected graph that was subsequently clustered using the Leiden clustering algorithm¹²⁰ implementing strict modularity and 1000 iterations (Supplement 1, Supplement 2). Pairwise MI maps from multiple sequence alignments (MSA) were calculated using the MSavolve matlab toolbox⁷⁹. The MSA for tRNA was created using the Sprinzl et al. tRNAdb¹²¹ using all organisms and tRNA species. A custom numerical representation of modified bases was prepared for use with the MSavolve toolbox. The *T. aquaticus* sequence for EF-Tu was used as a query sequence for the deepMSA2 server¹²² as well as mmseqs2¹²³ search and alignment. The EF-Tu alignments were post processed to only include segments aligned to the reference sequence with no overhang at both ends to ensure identical length of the MSA. 3dMI was calculated from both MSAs without gap correction.

7. Results

7.1 Trajectory overview

Root-mean-square deviation (RMSD) analysis over complete trajectory timelines reveal no major structural fluctuations in RMSD for all tRNA^{Phe} molecules. Molecular dynamics (MD) trajectories are stable after 300 ns of simulation time and RMSD to starting structures does not reach more than 7 Å (Figure 2A, Table 2). All MD trajectories display a stable behavior of radius of gyration (RGYR) throughout with no major opening or closing events. Notably, MD trajectories of tRNA^{Phe}.EF-Tu complexes fluctuate less in their RGYR than single tRNA^{Phe} systems (Figure 2B, Table 2).

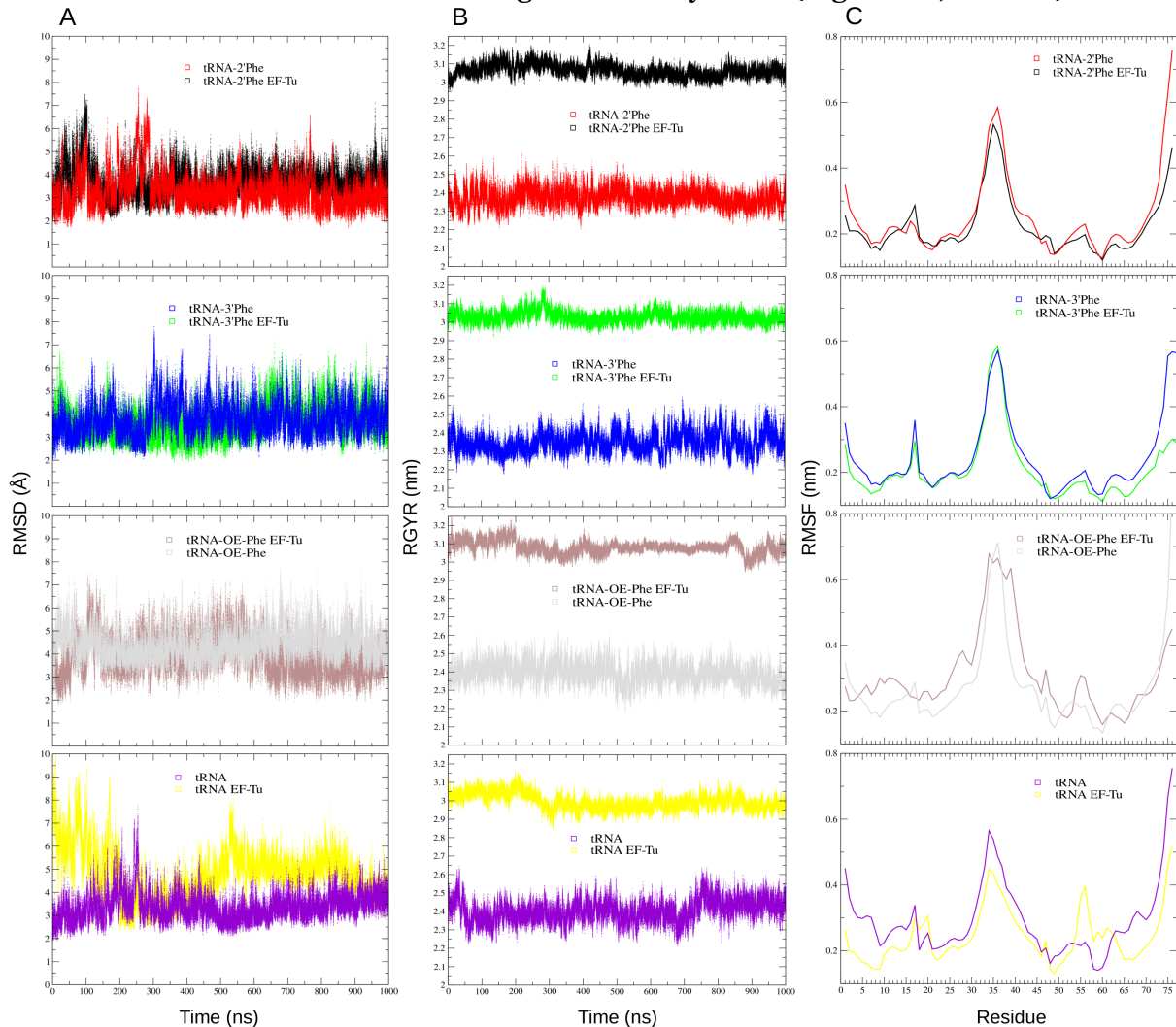


Figure 2: Overview of tRNA^{Phe} isomer MD trajectories for all systems. **A** tRNA^{Phe} backbone RMSD to average structure over MD trajectories. **B** RGYR for system over MD trajectories. **C** RMSF per residue (average of backbone atoms) for tRNA^{Phe} calculated for the complete MD trajectories.

The profiles of average root-mean-square fluctuation (RMSF) per tRNA backbone residue all display a similar pattern with a highly fluctuating anticodon as well as

acceptor stem and residue H2U17 (5,6-dihydrouridine). In complex with EF-Tu all systems fluctuate less at the acceptor stem although the orthoester (OE) and 3'-aminoacylated tRNA show the greatest decrease in acceptor stem dynamics. The 3'-attached tRNA already displayed a lower level of fluctuation at acceptor stem in the free tRNA system compared to the 2'-isomer. The systems with the OE-attached amino acid additionally exhibit the highest RMSF at the anticodon, peaking at around 7 nm average backbone RMSF, both for single tRNA and in complex with EF-Tu (Figure 2C). The uncharged systems show lesser dynamic range in complex with EF-Tu resulting in a less pronounced profile.

Table 2: Mean RMSD and RGYR with standard deviation over the whole MD trajectories.

mean	3'Phe-tRNA	2'Phe-tRNA	OE-Phe-tRNA	3'Phe-tRNA EF-Tu	2'Phe-tRNA EF-Tu	OE-Phe-tRNA EF-Tu	tRNA EF-Tu	tRNA
RMSD (Å)	3.65	3.69	3.95	3.31	3.39	4.78	3.78	4.40
±st.dev.	± 0.70	± 0.645	± 0.81	± 0.738	± 0.58	± 1.09	± 0.69	± 0.61
RGYR (nm)	2.59	2.37	2.39	3.02	3.06	3.08	2.99	2.40
±st.dev.	± 0.052	± 0.046	± 0.052	± 0.030	± 0.033	± 0.034	± 0.04	± 0.05

7.2 Clustering

All MD conformational ensembles were clustered using the single linkage method with a predefined RMSD cutoff for several different components of the tRNA, namely the tRNA backbone atoms and the backbone of the acceptor stem residues G1, A76 and if present in the system, the attached Phe. The resulting clustering was used to manually inspect dominant conformational aspects of the acceptor stem across aminoacylation isomers. The central structure from clustering by acceptor stem residues were used to identify possible interactions between tRNA strand ends while clustering by complete tRNA backbone atoms was inspected for major shifts in global conformation. No obvious impact of aminoacylation site on global structure could be identified from clustering when visually inspecting the central structures of the highest occupied clusters of tRNA^{Phe} backbone conformations, but clustering of the acceptor stem backbone revealed the hydrogen bonding and base stacking interactions between residues 1 and 76 (see Section 7.4) to be different across central structures of dominant clusters between aminoacylation isomers. Conformational variety of the acceptor stem ends generally decreases in systems where tRNA is bound to EF-Tu (Table 3). These systems possess one or two clusters containing the vast majority of the conformations considered for clustering. The OE- and 3'-attachment isomers without EF-Tu display the biggest variety in G1, A76 and Phe backbone conformations, with only 12% and 24% of clustered structures in the three most populated clusters. This points to greater diversity in acceptor stem conformations compared to MD simulations of non-acylated and 2'-aminoacylated tRNA^{Phe}, where the combined occupancy of the top three clusters rises to 55.5% and 73.1% respectively.

Table 3: Occupancies of top three clusters when clustering by A76, G1 and Phe backbone using single linkage clustering (RMSD cutoff 0.5 Å).

system	occupancy cluster 1	occupancy cluster 2	occupancy cluster 3
tRNA	53.42%	1.10%	1.00%
tRNA·EF-Tu	87.21%	6.80%	0.40%
3'Phe-tRNA	18.04%	4.10%	2.00%
3'Phe-tRNA·EF-Tu	50.32%	48.68%	0.10%
2'Phe-tRNA	36.63%	35.68%	0.80%
2'Phe-tRNA·EF-Tu	76.76%	10.64%	1.30%
OE-Phe-tRNA	6.95%	4.00%	1.80%
OE-Phe-tRNA·EF-Tu	74.36%	3.25%	3.25%

7.3 Solvent accessible surface area of the amino acid carbonyl

As a proxy for hydrolytic stability of the aminoacyl ester bond, the solvent accessible surface area (SASA) for the carbonyl group of the aminoacyl ester bond was computed for all aminoacylated tRNA^{Phe} systems. This was done in analogy to previous computational work by Polyansky et al. (see Section 4.2) where systematically higher exposure to the solvent was found for 3'-aminoacylated trinucleotides. The mean exposure of the carbonyl group is highest in the single OE-aminoacylated tRNA^{Phe} system followed by the 3'-aminoacylated tRNA^{Phe} in complex with EF-Tu (Table 4). The highest value of exposure was calculated from the single 3'-aminoacylated tRNA^{Phe} trajectory at 42 Å². The distributions normalized by this value show a generally buried carbonyl group with the only exceptions being the aforementioned MD trajectories. If taken as a proxy for hydrolytic stability, generally only the 3'-aminoacylated tRNA^{Phe} exposes the ester bond in complex with EF-Tu while all free tRNA's shield the bond from solvent exposure except for the OE-aminoacylated tRNA^{Phe} (Figure 3).

Table 4: Aminoacyl ester carbonyl exposure SASA in Å² with standard deviation over complete MD trajectories.

	3'Phe-tRNA	2'Phe-tRNA	OE-Phe-tRNA	3'Phe-tRNA EF-Tu	2'Phe-tRNA EF-Tu	OE-Phe-tRNA EF-Tu
SASA mean (Å ²) ± st.dev	9.32 ± 10.4	7.79 ± 6.1	14.86 ± 5.9	14.63 ± 5.6	5.33 ± 5.7	4.42 ± 4.6
relative exposure	22.19%	18.54%	35.38%	35.38%	12.69%	10.5%

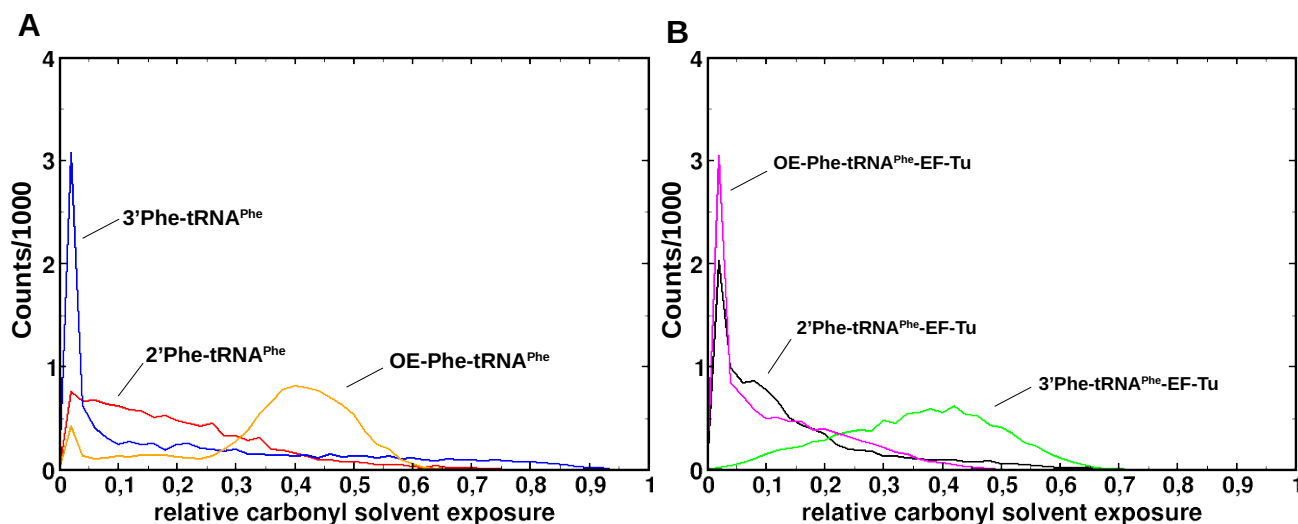


Figure 3: Exposure of the amino acid carbonyl group (SASA rescaled by the observed maximum value). **A** Distribution of relative carbonyl exposure for three charging isomer MD trajectories without EF-Tu. **B** Distribution of relative carbonyl exposure for three charging isomer MD trajectories of tRNA^{Phe} bound to EF-Tu. All SASA values normalized by the highest exposure value across all MD trajectories (42 Å², 3'-Phe-tRNA^{Phe}).

7.4 Interactions between tRNA 5'- and 3'-ends

While visual inspection of the central structures of dominant tRNA backbone clusters did not give any indication of obvious and consistent differences in global tRNA^{Phe} backbone conformation between amino acid attachment isomers, clustering of tRNA backbone atom conformations of the acceptor stem pointed to a difference in interactions occurring between the G1 and A76 ends of tRNA joined at the acceptor stem. The importance of the A76 non-connective hydroxyl group established from literature (see Section 4.2) further bolstered the need for quantifying the observed interactions to better understand the implications of amino acid attachment for tRNA^{Phe} conformation at the acceptor stem. The main interactions identified to differentially occur between aminoacylation isomers can be divided into two classes: hydrogen bonding between 3'- and 5'-strand ends and stacking between sidechain base rings.

7.4.1 Hydrogen bonding

Hydrogen bonding between the phosphate group of the G1 residue to either the amino group of the attached amino acid or the non connective hydroxyl group of the terminal adenosine of the CCA acceptor triplet (Table 5) was drastically influenced by the attachment site of the amino acid. Naturally the OE-attached amino acid prevented hydrogen bonding to either of the A76 hydroxyl groups while the absence of the aminoacyl ester bond in uncharged tRNA^{Phe} allowed for frequent interactions with both hydroxyl groups (Figure 4B). Hydrogen bonding was observed about five times as often to the 3'-compared to the 2'-hydroxyl group. Interestingly the difference in attachment

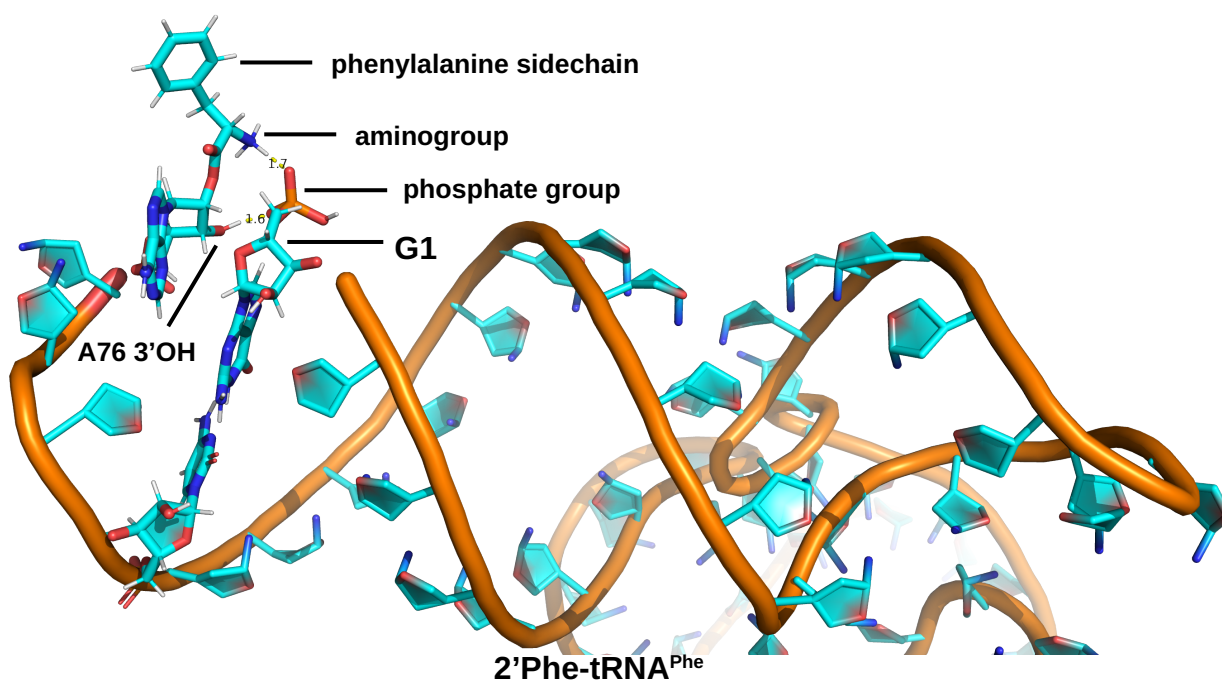
site of Phe at the 2'- or 3'-hydroxyl position lead to an even starker imbalance in observed hydrogen bonding. The 3'-aminoacylated tRNA^{Phe} satisfied the constraints for detected hydrogen bonding only in ~4% of MD simulation frames, while about 37% of frames in the 2'-attached tRNA^{Phe} system displayed a hydrogen bond between one of the G1 phosphate oxygen atoms and the 3'-hydroxyl hydrogen. Hydrogen bonding between the amino group of the attached Phe (Figure 4A) was overall less commonly observed and occurred about twice as often in the 2'-Phe-tRNA^{Phe} than in its 3'-aminoacylated counterpart.

Table 5: Percentage of frames with identified hydrogen bond between G1 phosphate and A76 non connective hydroxyl or aminogroup of the attached Phe amino acid.

h-bond partners	3'-Phe-tRNA	2'-Phe-tRNA	OE-Phe-tRNA	tRNA
G1 - A76 2'-/3'-OH	4.15%	36.71%	/	60.34%
G1 - Phe aminogroup	13.99%	27.92%	2.60%	NA

Barely any interactions between the 5'-end of the tRNA molecule and the amino group of the attached amino acid could be observed in the OE-Phe-tRNA^{Phe} system with less than 3% of frames satisfying the distance and angle constraints set. No hydrogen bonding between tRNA strand ends was observed in EF-Tu bound systems and minimum distances between the 3'- and 5'-ends of tRNA were too high to facilitate such interactions over the course of all MD trajectories containing tRNA EF-Tu complexes (Figure 6).

A



B

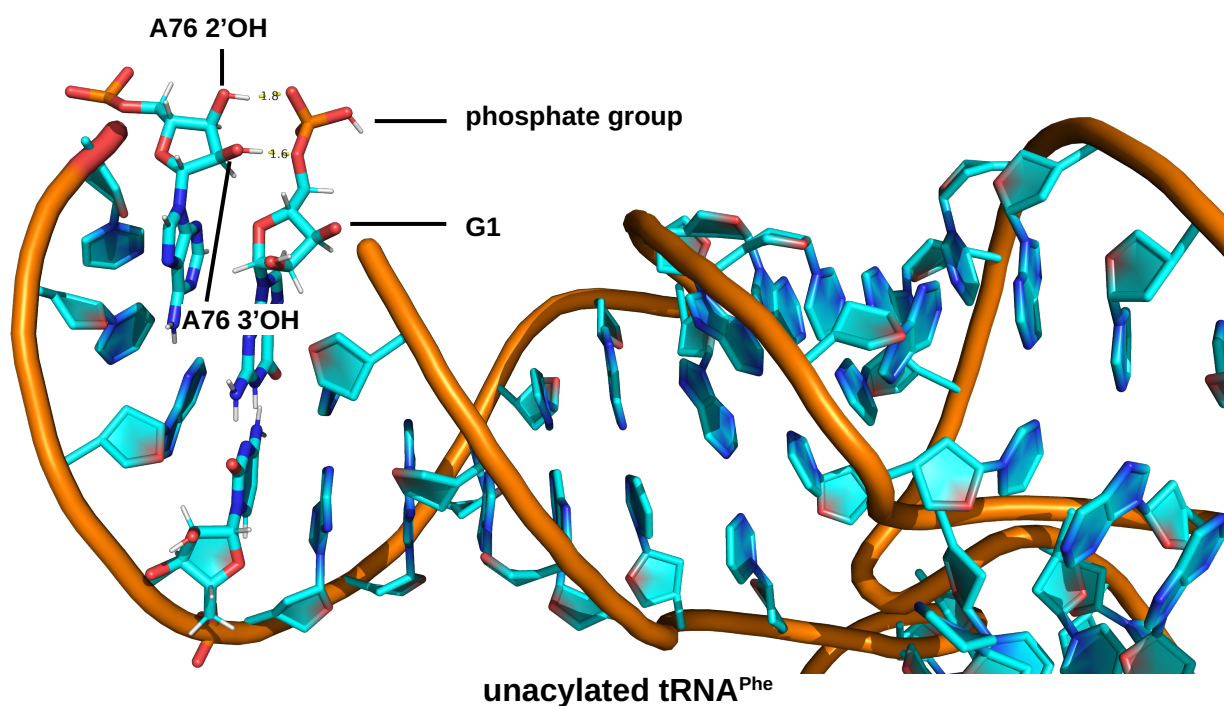


Figure 4: Examples of interactions between G1 and the tRNA acceptor triplet identified with conformational clustering by G1, A76 and amino acid backbone. **A** Hydrogen bonds between the G1 phosphate group and the A76 3'-OH group as well as the Phe aminogroup. Central structure of cluster 1 (see Table 3), 2'Phe-tRNA^{Phe} system. **B** Hydrogen bonds identified in non-acylated tRNA^{Phe} between A76 hydroxyl groups and G1 phosphate. Central structure of cluster 1 (see Table 3). Distances are given in Å, tRNA shown as cartoon, A76, G1 and Phe rendered with stick representation.

7.4.2 G1-A76 base stacking

Another interaction between tRNA ends was the stacking of the sidechain base rings of residues A76 and G1 (Figure 5). According to the stacking criteria (given in Methods), ~70% of frames of the uncharged tRNA^{Phe} exhibited stacking. About 61% of 2'-aminoacylated tRNA^{Phe} MD simulation frames also satisfied the criteria while the base stacking interaction was much less commonly observed in the 3'- and OE-aminoacylated MD trajectories at around 18% and 6% respectively (Table 6). The distribution of angles between ring-plane normal vectors for all MD trajectories clearly displays a preference for the uncharged and 2'-aminoacylated tRNA^{Phe} systems for planar base ring stacking compared to a peak at ~125 degrees for the 3'-Phe-tRNA^{Phe} system. Additionally, a small percentage of possible T-stacking between bases in non-acylated tRNA^{Phe} can be observed (Supplement 3).

Table 6: Percentage of frames with identified base stacking between A76 and G1.

	3'-Phe-tRNA	2'-Phe-tRNA	OE-Phe-tRNA	tRNA
% of frames	18.46%	16.13%	6.00%	69.91%

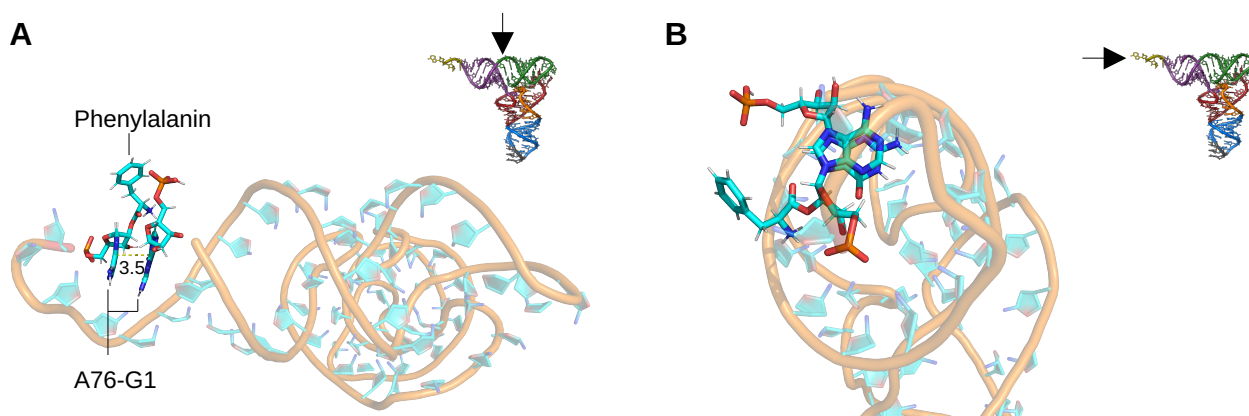


Figure 5: Stacking interactions between G1 and A76 for the central structure of second most occupied cluster for 2'-Phe-tRNA^{Phe} MD ensemble clustered by G1, A67 and Phe backbone. **A** Top view with distance between base planes in Å, marked by yellow dashed lines. **B** Front view. G1 and A67 shown with sticks, tRNA body shown in cartoon representation.

7.4.3 Acceptor stem strand separation

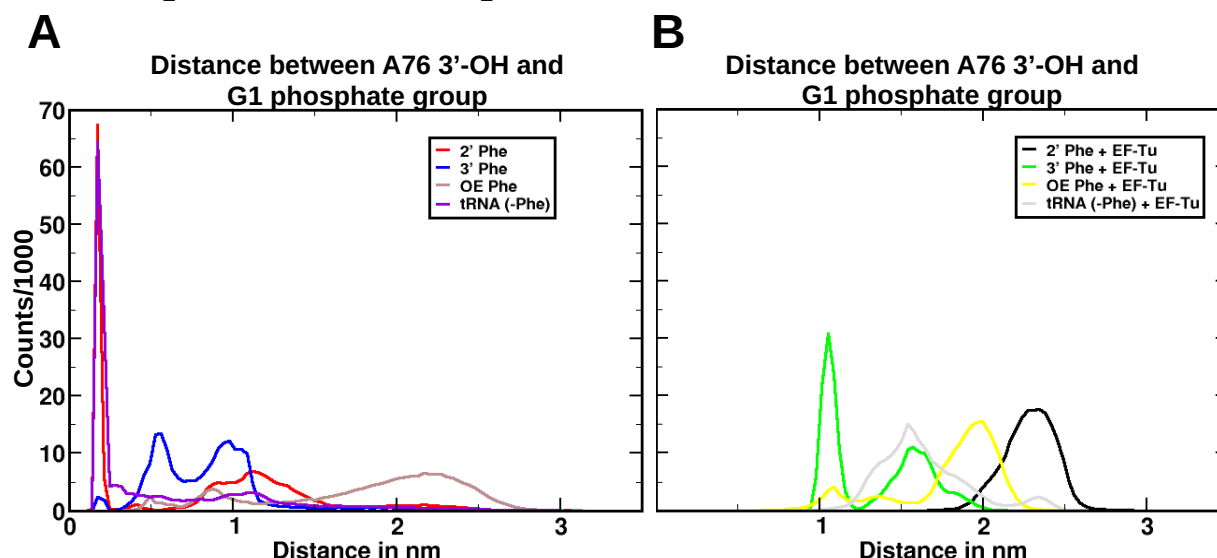


Figure 6: Distribution of distance between A76 3'-OH and G1 phosphate groups.
A Free tRNA^{Phe} isomers. B EF-Tu bound tRNA^{Phe} isomers.

The distance between the 5'- and 3'-ends of the tRNA molecule comprising the 5'-terminal acceptor stem guanosine base (G1) and the 3'-terminal adenosine (A76) of the acceptor triplet CCA, are measured throughout MD trajectories by the minimum distance between the A76 3'-OH and the G1 phosphate group. The distributions show a clear difference in strand separation between the unbound (Figure 6A) and bound tRNA (Figure 6B) as well as a clear preference for close association of both strands in the uncharged and 2'-aminoacylated tRNA in free tRNA systems. The OE-aminoacylated tRNA displays the largest separation between strands in free tRNA (Table 7).

Table 7: Mean distance between G1 phosphate group and 3'-oxygen of A76 over complete MD trajectories.

	3'Phe-tRNA	2'Phe-tRNA	OE-Phe-tRNA	3'Phe-tRNA EF-Tu	2'Phe-tRNA EF-Tu	OE-Phe-tRNA EF-Tu	tRNA EF-Tu	tRNA
distance								
3'-OH-G1P	0.82	0.84	1.79	1.32	2.27	1.78	1.62	0.57
(nm)	±0.32	±0.62	±0.59	±0.29	±0.15	±0.33	±0.28	±0.57
± st.dev								

7.5 Conformational dynamics

After investigation of the aforementioned geometric aspects of the tRNA charging isomerism and implications for tRNA conformation and internal interactions, the dynamical behavior of the different simulated systems was characterized. For this purpose, RMSF of backbone atoms averaged over residues and bases as well as configurational entropy calculations were performed on both tRNA and EF-Tu.

7.5.1 tRNA

From the guaranteed upper bound in configurational entropy provided by the MIST approach implemented in PARENT (see Section 4.7), the charging of tRNA appears to be entropically unfavorable for all isomers. Both 2'- and OE-attachment of Phe incur a penalty to the entropic contribution to free energy ($T\Delta S$) of ~ 40 kJ/mol compared to the uncharged tRNA.

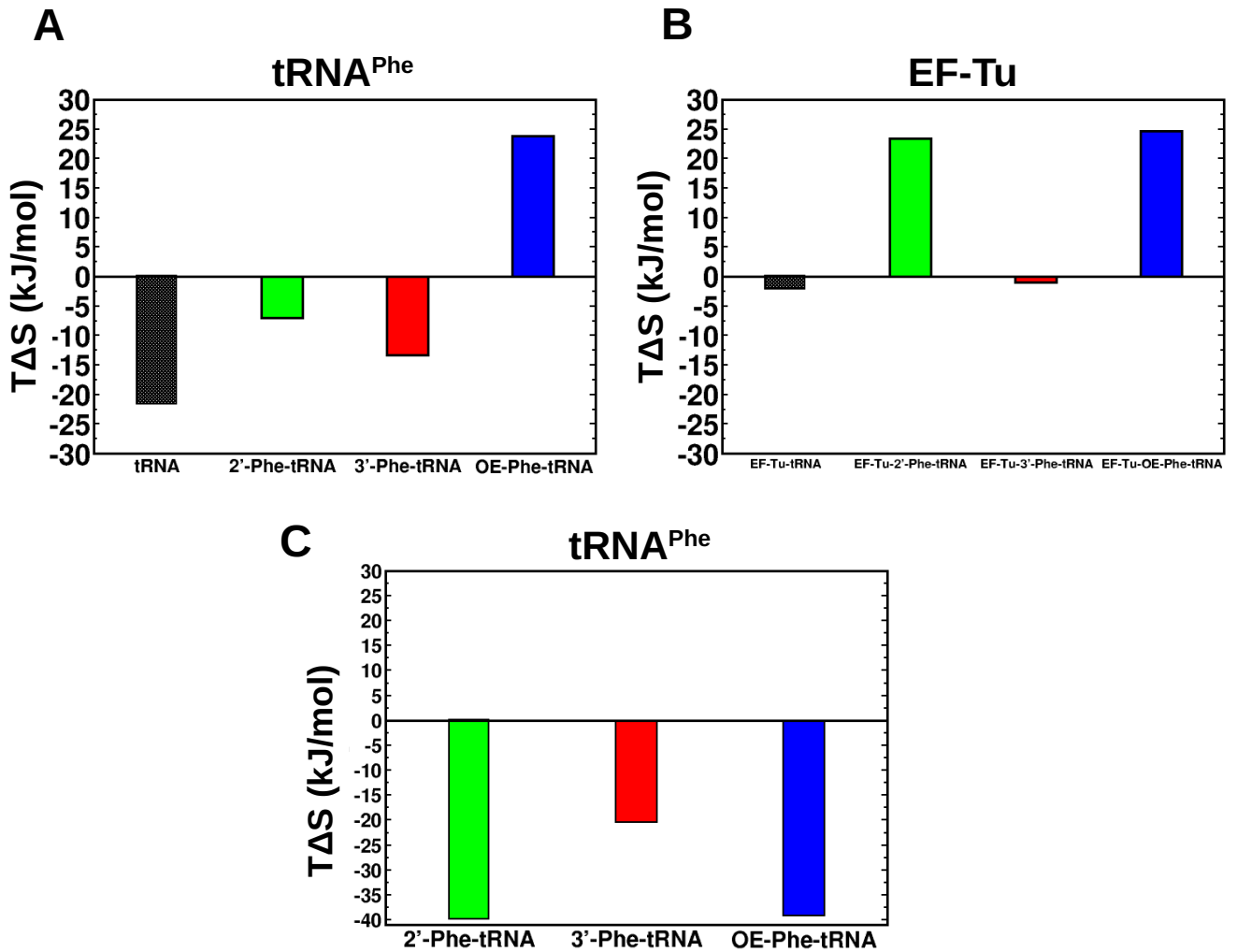


Figure 7: Change in configurational entropy ($T\Delta S$) for ternary complexes and tRNA^{Phe} charging. **A** $T\Delta S$ of tRNA^{Phe} between bound and unbound tRNA isomers. **B** $T\Delta S$ of EF-Tu between EF-Tu in complex with different tRNA^{Phe} isomers and free EF-Tu. **C** $T\Delta S$ of tRNA^{Phe} between charged tRNA isomers and uncharged tRNA^{Phe}. (Differences rescaled by degrees of freedom, see Methods).

The 3'-attachment is slightly less unfavorable with a reduction in $T\Delta S$ of ~ 20 kJ/mol (Figure 7C). Overall, uncharged tRNA has the highest calculated configurational entropy while the tRNA molecules with 2'- or 3'-attached Phe bound to EF-Tu have the lowest (Supplement 4).

Configurational entropy for the OE-isomers is higher than for the differently charged tRNA^{Phe} isomers in complex and even slightly higher than the 2'- and 3'-aminoacylated unbound tRNA^{Phe} systems. The impact of configurational entropy on the free energy of EF-Tu binding the different tRNA molecules is shown to be unfavorable for uncharged and 2'- or 3'-aminoacylated tRNA^{Phe} and slightly favorable (~25 kJ/mol) for the OE-isomer. Uncharged tRNA^{Phe} displays the largest reduction in TΔS in the EF-Tu bound state (Figure 7A).

7.5.2 EF-Tu

The dynamics of EF-Tu was analyzed by average RMSF of backbone atoms per residue (RMSF_{res}). RMSF_{res} was evaluated for each of the three EF-Tu domains across simulated systems (EF-Tu single, bound to all aminoacylation isomers). In domain I, the nucleotide binding domain containing one half of the interface to the tRNA acceptor stem in between domain I and domain II, a spike in fluctuation for the stretch of residues 52-62, in the single EF-Tu and 2'-Phe-tRNA^{Phe}·EF-Tu MD trajectories. This patch lies entirely in the switch I region (41-63) of EF-Tu. Moreover, residues K52 and E55 directly contact the 3'-strand of the acceptor stem at nucleotides A73 and C74 (the first residue of the universal CCA acceptor triplet). The other notable difference in RMSF_{res} profiles between systems is an increased fluctuation at residue G165 in the trajectory containing uncharged tRNA^{Phe} bound to EF-Tu (Figure 8A). Domain II in which the majority of tRNA acceptor stem binding interacting residues are located, displays two principal differences in RMSF_{res} between systems. The first is an increase average fluctuation of residue T232 in 2'-Phe-tRNA^{Phe}·EF-Tu as well as a decrease in dynamics between residues 264 and 278 in 2'-Phe-tRNA^{Phe}·EF-Tu compared to other systems. T232 forms hydrogen bonds to the C75 phosphate in tRNA^{Thr} and possibly tRNA^{Phe}⁴⁹. Both E272 and R274 interact with the 3'-terminal adenosine A76 in the crystal structure of tRNA^{Phe} ternary complex¹⁰⁵. R274 interacts with the phosphate backbone while E271 stacks on top of the adenosine. H273 interacts with the amino acid directly via ionic interactions with the carboxyl oxygen (Figure 8B). Domain III of EF-Tu contains almost exclusively the residues of the tRNA·EF-Tu interface involved in binding the T-stem region of tRNA. From the RMSF_{res} profile, F344 to T346 fluctuates more strongly in free EF-Tu MD simulations than in systems where tRNA is present. The proximal W343 possibly interacts with the phosphate backbone of C65 in both tRNA^{Thr} and tRNA^{Phe} as do K90 and R369, indicating a crucial hub of the EF-Tu:T-stem interface at this junction. The second discrepancy in RMSF_{res} for domain III is located at the residue E390 which is found to be less fluctuating in EF-Tu bound to 3'-aminoacylated tRNA^{Phe} (Figure 8C). E390 is thought to be interacting with the N2 atom of G51 in the *T. thermophilus* EF-Tu yeast Phe-tRNA^{Phe} complex⁴⁹.

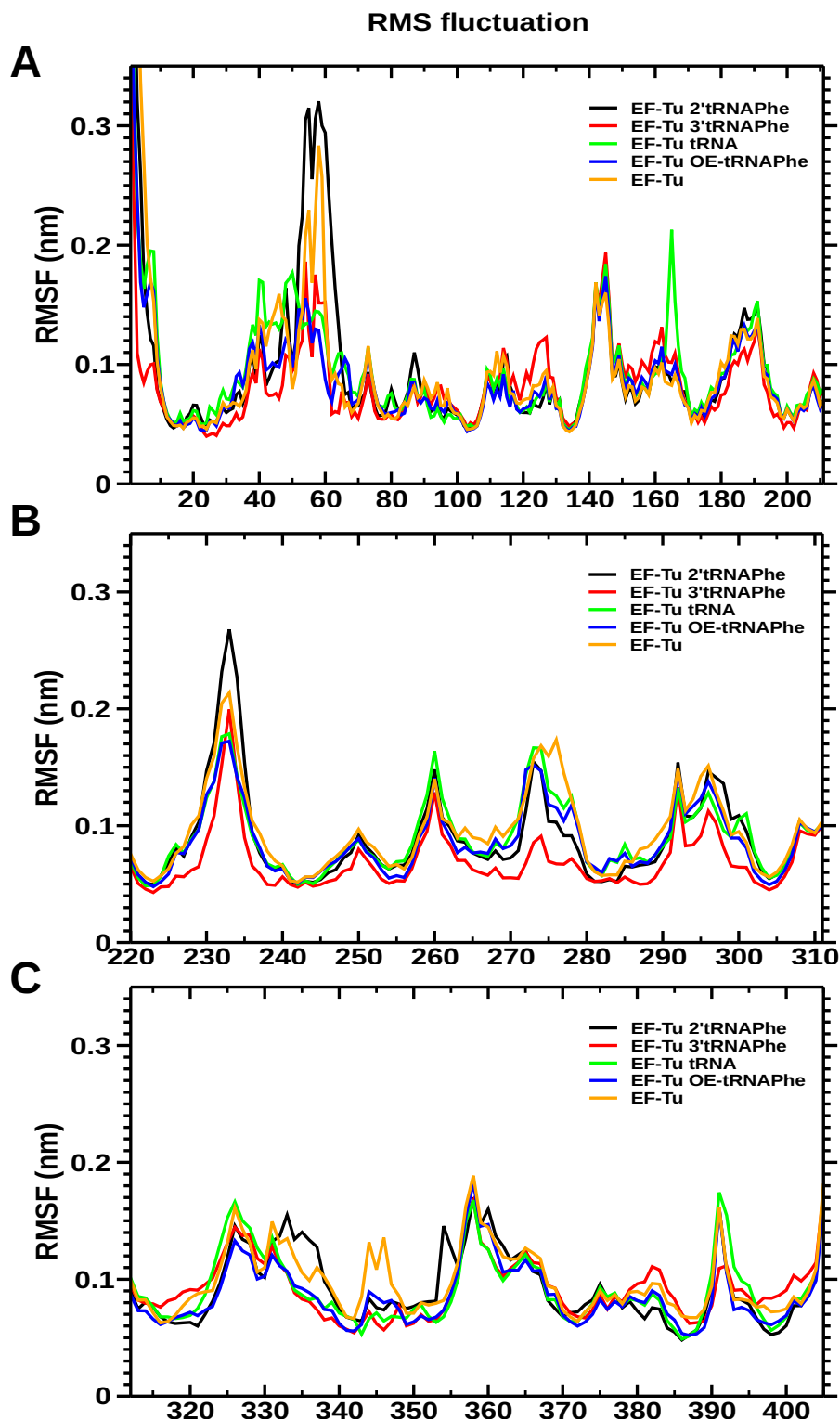


Figure 8: Average backbone atom RMSF per residue in EF-Tu for all systems.
A Domain I. B Domain II. C Domain III.

TΔS for the EF-Tu protein bound to differently aminoacylated tRNA^{Phe}'s is close to zero when binding non-acylated or 2'-aminoacylated tRNA^{Phe}, while binding 3'-or OE-aminoacylated tRNA^{Phe} raises the entropic contribution to free energy to around 25 kJ/mol (Figure 7B). When subtracting the configurational entropies of free EF-Tu and

the respective charging isomers from that of the respective complexes and multiplying by the temperature, $T\Delta S$ for the OE-Phe-tRNA^{Phe} EF-Tu and 2'-Phe-tRNA^{Phe} EF-Tu ternary complexes are highest at 36.5 kJ/mol and 67.8 kJ/mol. The complexes with non-acylated and the 3'-Phe-tRNA^{Phe} are both similarly unaffected in their $T\Delta S$. The change in configurational entropy upon binding is close to 0 with a $T\Delta S$ of -5 kJ/mol and 5 kJ/mol (Figure 9) as compared to their individual components not in complex.

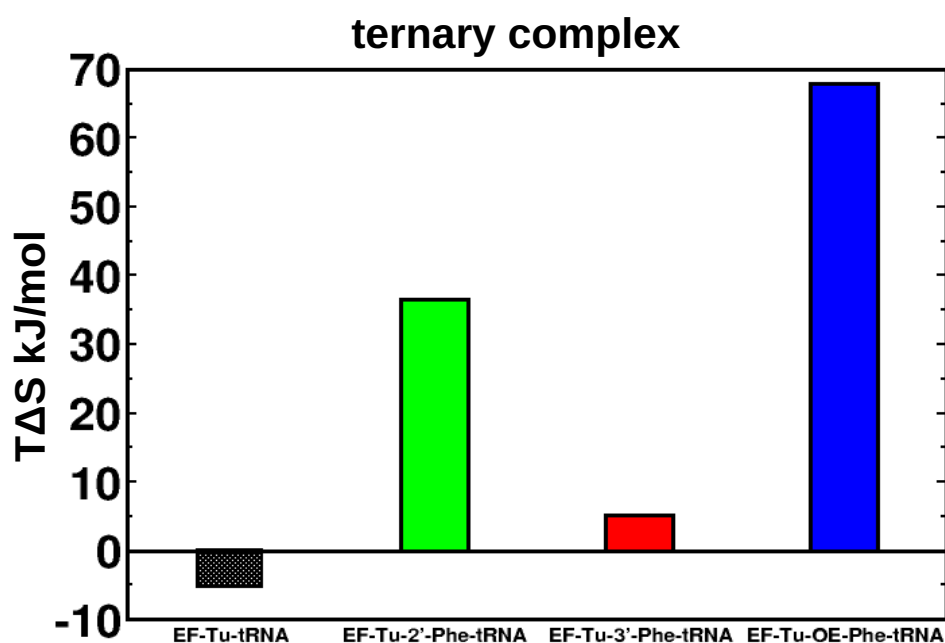


Figure 9: Configurational entropy change upon binding ($T\Delta S$) for ternary complexes compared to free EF-Tu and respective free tRNA^{Phe} charging isomers.

7.6 Mutual information

A necessary step in calculating the configurational entropy from MD trajectories with the approach outlined in the introduction is the calculation of mutual information (MI) between degrees of freedom. These MI values can be combined into a one value for a pair of amino acids or nucleotides by adding up the MI between all degrees of freedom associated with such units. These values can be interpreted as a measure of how much two residues are coupled in their conformational dynamics during equilibrium MD simulations. The inclusion of higher order correlations between degrees of freedom in MI measures can be used to identify pathways of correlated conformation from equilibrium MD simulations and infer pathways activated by perturbations such as ligand or protein protein/RNA interactions. Instead of correlated conformations, co variation in residues across evolving homologous protein sequences can be computed in terms of MI as well (see Section 4.6). This approach is used here to compare coupling between residues arising completely from conformational dynamics in MD simulations to those containing additional phylogenetic, structural and random influences present in the alignments of sequence data. By contrasting MI from these different datasets with

differing sources of systemic and random noise, true functional coupling can be more clearly delineated.

7.6.1 Mutual information from MD

As the calculation of meaningful MI values is very dependent on complete sampling of the conformational space, two MD trajectories with one million frames were prepared for the closer inspection of MI between residues/nucleotides of single EF-Tu and 3'-aminoacylated tRNA^{Phe}. All pairwise MI maps were preprocessed using ARTEMIS (see Section 6.4, Supplement 5). To qualitatively analyze regions of high internal coupling via MI, community detection was employed to cluster protein residues/tRNA nucleotides using graph-based approaches (see Section 6.4). All free tRNA^{Phe} MD trajectories display high MI between the anticodon loop residues in their pairwise MI maps. Consistently, the anticodon and acceptor stem regions share higher than background MI across all aminoacylation states. Both T-stems are additionally coupled to the anticodon as well as the D-arm region at residue numbers 15-20. This basic pattern in shared MI varies in strength between aminoacylation types (Figure 10). 2'-Phe-tRNA^{Phe} has relatively lower internal coupling between anticodon loop residues while the 3'-Phe-tRNA^{Phe} trajectory displays the strongest coupling between the anticodon and acceptor stem. The non-acylated tRNA^{Phe} system shows an additional coupling pattern of acceptor stem residues A66, A67, U68 with the anticodon and the D-arm. These patterns generally persist in complex with EF-Tu with the exception of 3'-attached Phe-tRNA^{Phe}. There, both the MI between anticodon and acceptor stem as well as the internal coupling of the anticodon are greatly reduced (Supplement 6).

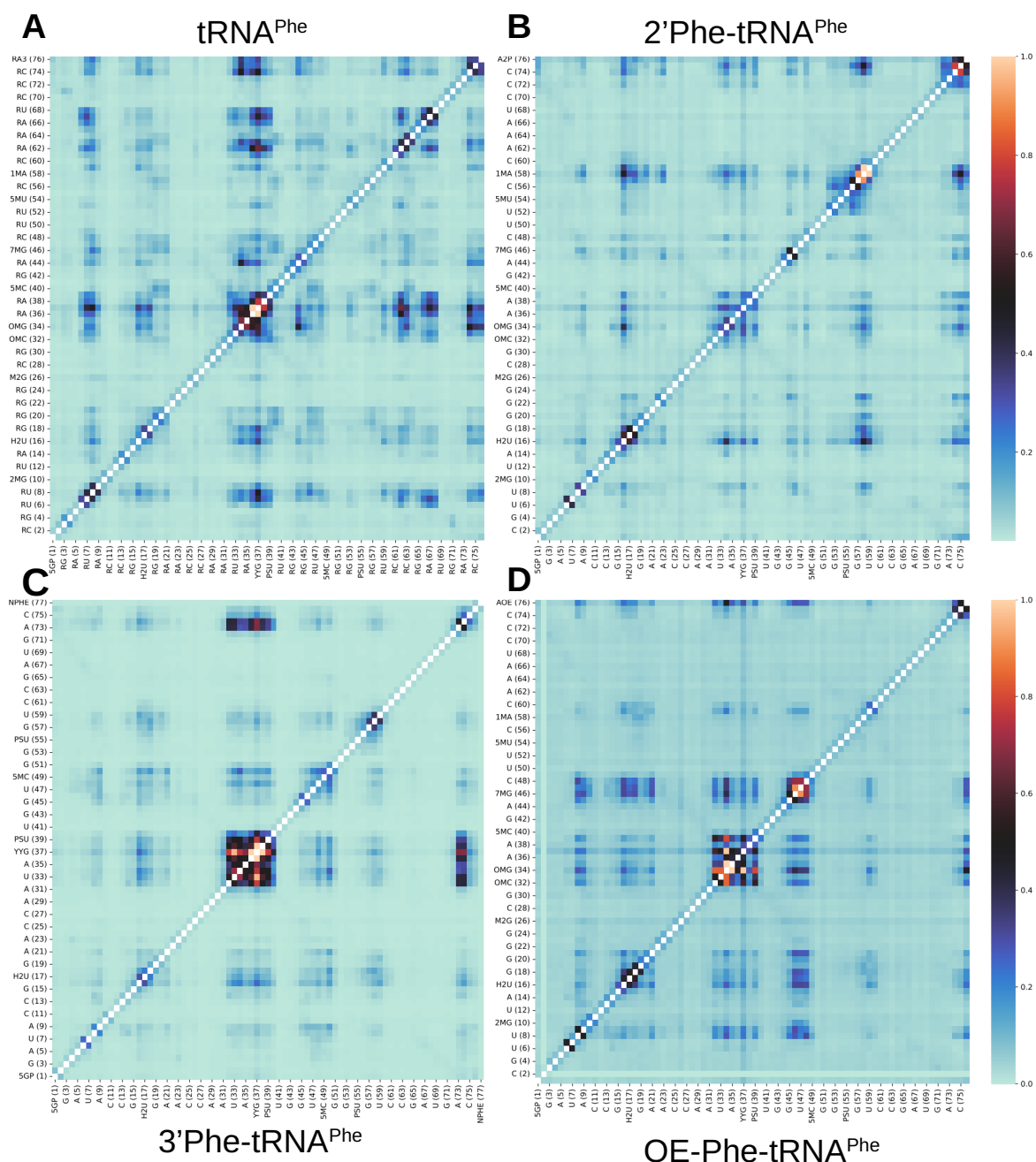


Figure 10: Pairwise MI maps for free $tRNA^{Phe}$ aminoacylation isomer MD simulations. **A** MI map of uncharged $tRNA^{Phe}$ trajectory. **B** MI map of $2'\text{-Phe-}tRNA^{Phe}$ trajectory. **C** MI map of $3'\text{-Phe-}tRNA^{Phe}$ trajectory. **D** MI map of $OE\text{-Phe-}tRNA^{Phe}$ trajectory. All values rescaled by maximum MI value for a given map.

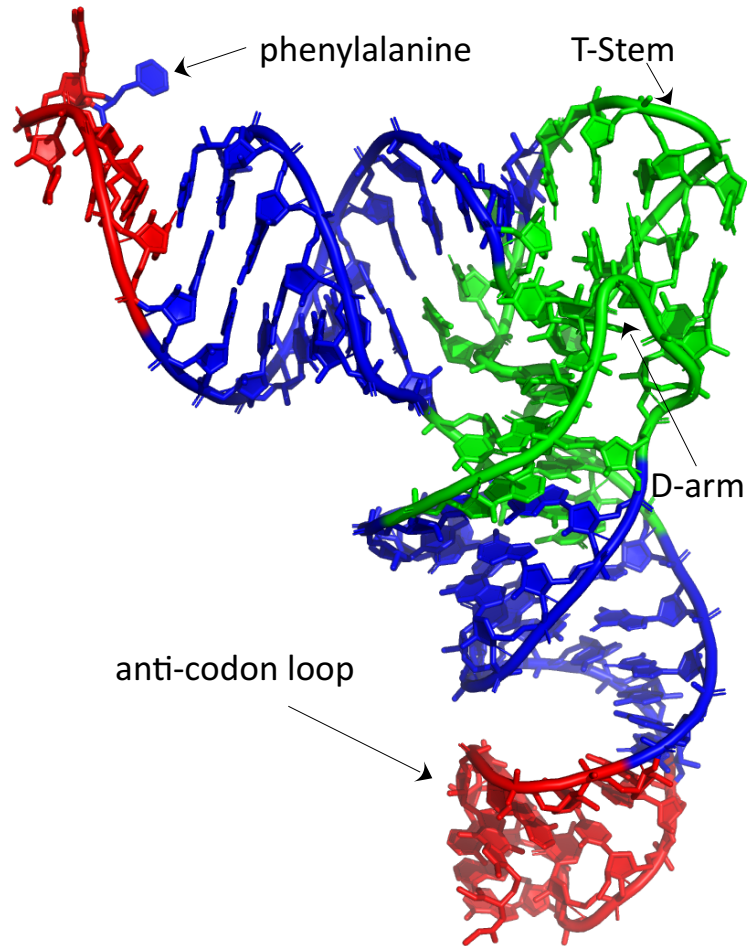


Figure 11: Spectral clustering by MI of 3'Phe-tRNA^{Phe} system (1×10^6 frames trajectory). Nucleotides colored by cluster assignment.

Using the elbow of the k-means inertia graph, the number of clusters was assigned for each type of system (tRNA, EF-Tu and ternary complex). For this purpose the EF-Tu concatenated master trajectory was used and MI from all EF-Tu MD simulations was clustered into 4 subsets. Elbow analysis for the 1×10^6 frames trajectory of 3'Phe-tRNA^{Phe} lead to clustering all free tRNA MD trajectories with 3 clusters. Finally the optimal number of clusters for ternary complexes was determined at 5 clusters from the 3'Phe-tRNA^{Phe}·EF-Tu trajectory. Clustering of the data using spectral clustering with three clusters, groups the anticodon and the acceptor stem as well as the T-arm/D-arm in one cluster for all free tRNA systems except the OE-Phe-tRNA^{Phe} trajectory, where the internal coupling of the anticodon loop overpowers the connection to the acceptor stem which is instead grouped with the T-stem 2 and D-arm. Figure 11 shows the clustering of the 1 million frames trajectory of 3'Phe-tRNA^{Phe}. Clustering of MI for the concatenated EF-Tu master trajectory comprising 4 replicate 1 μ s MD simulations with four clusters recovers the biological domains to a large extent. Additionally a fourth

cluster is present spanning the interface between all domains (Figure 12). This clear cut distinction between functional domains repeated by MI analysis is less pronounced but still visible in the clustering of EF-Tu coupling in the 3'-Phe-tRNA^{Phe} EF-Tu complex (Figure 12B). Other MD trajectories exhibit differing clustering results indicating a change in intramolecular communication communities after binding non canonical tRNA charging isomers or uncharged tRNA (Supplement 7). Mapping the clusters to their majority representative domains of EF-Tu and calculating the change in MI between clusters for EF-Tu systems binding differently aminoacylated tRNA, reveals differences in inter domain communication between the attachment isomers. For each of the three replicates of free EF-Tu MD simulations, the mean MI between the residue subgroups was calculated along with the standard deviation. The standard deviation in percent was calculated for each channel (sum of pairwise MI between subgroups) to establish a threshold for random noise in each communication channel between the identified residue groups. This was used to distinguish the impact of differently aminoacylated tRNA bound to EF-Tu, from noise. Different subgroups display varying degrees of noise in their shared MI (Table 8). Binding of uncharged tRNA^{Phe} does not seem to globally perturb the intramolecular communication of the identified EF-Tu subgroups. Only a slight increase in MI between the domain I and interface group clears the noise threshold established by the replicate EF-Tu MD trajectories (Table 8). In 2'-Phe-tRNA^{Phe} bound to EF-Tu, clusters representing domain II and III share more MI than in single EF-Tu. This is also the case for domain II and the interface cluster as well as the coupling of domain III and the interface subgroup. Both OE- and 3'-aminoacylated tRNA^{Phe} bound to EF-Tu reduces the MI between all subgroups. The biggest reduction is seen between the domain I cluster and the domain III cluster for EF-Tu in complex with OE-Phe-tRNA^{Phe} with a reduction of MI for this channel of 34% of the MI in the unbound EF-Tu MD trajectory (Figure 13).

Table 8: Mean and standard deviation of MI between communication channels from three independent MD-runs for free EF-Tu.

subgroup	mean MI	SD	SD/mean
interface-I	4882.68	216.59	4.44%
interface-II	3452.11	422.09	12.23%
interface-III	2708.24	140.31	5.18%
I-II	2081.82	41.08	1.97%
I-III	1696.51	72.87	4.30%
II-III	1155.20	84.38	7.30%

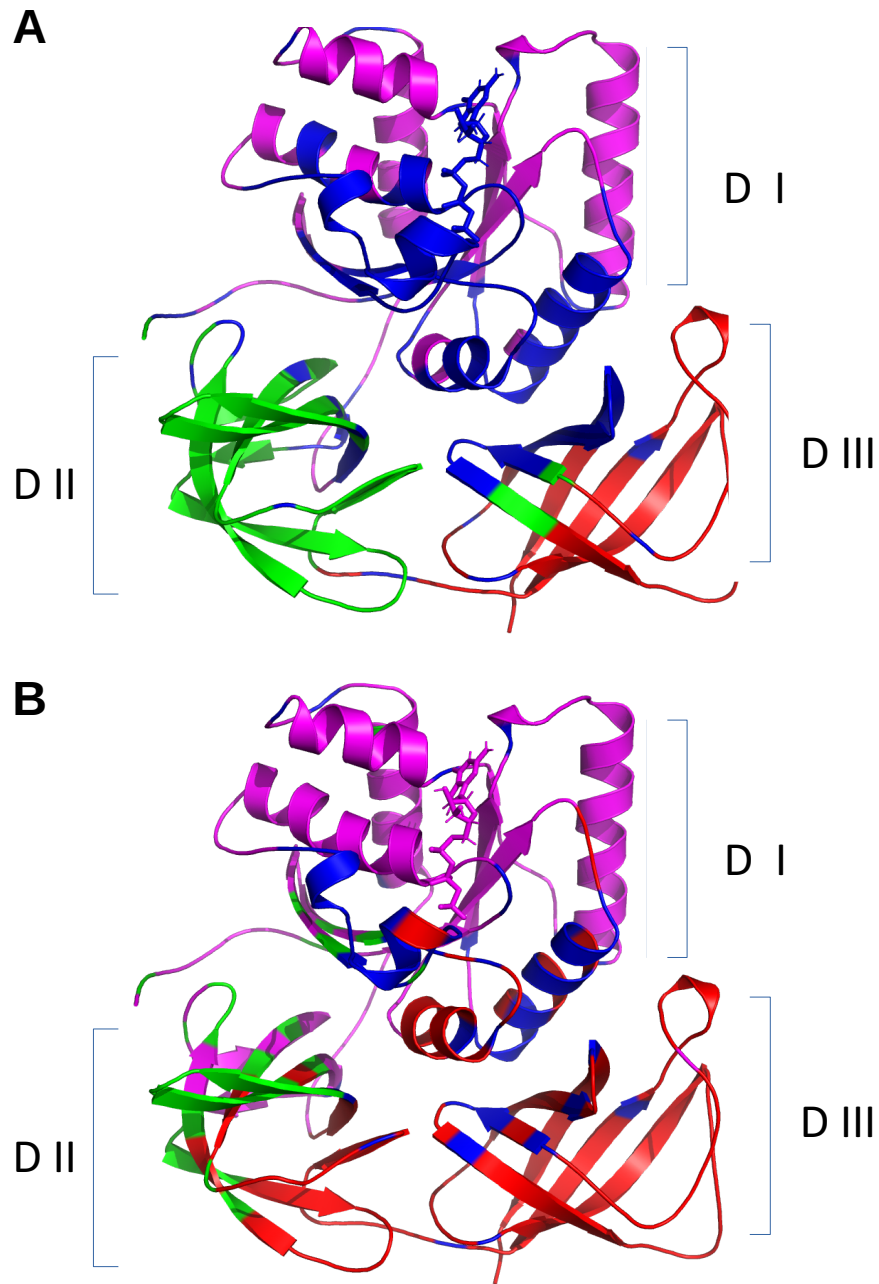


Figure 12: Spectral clustering of MI between EF-Tu amino acids into 4 clusters. Clusters mapped to the functional domains in EF-Tu. **A** Free EF-Tu. **B** EF-Tu from MD trajectory in complex with 3'Phe-tRNA^{Phe}. Protein shown in cartoon representation, GTP as sticks. Domain I in magenta, domain II in green, domain III in red, interface in blue.

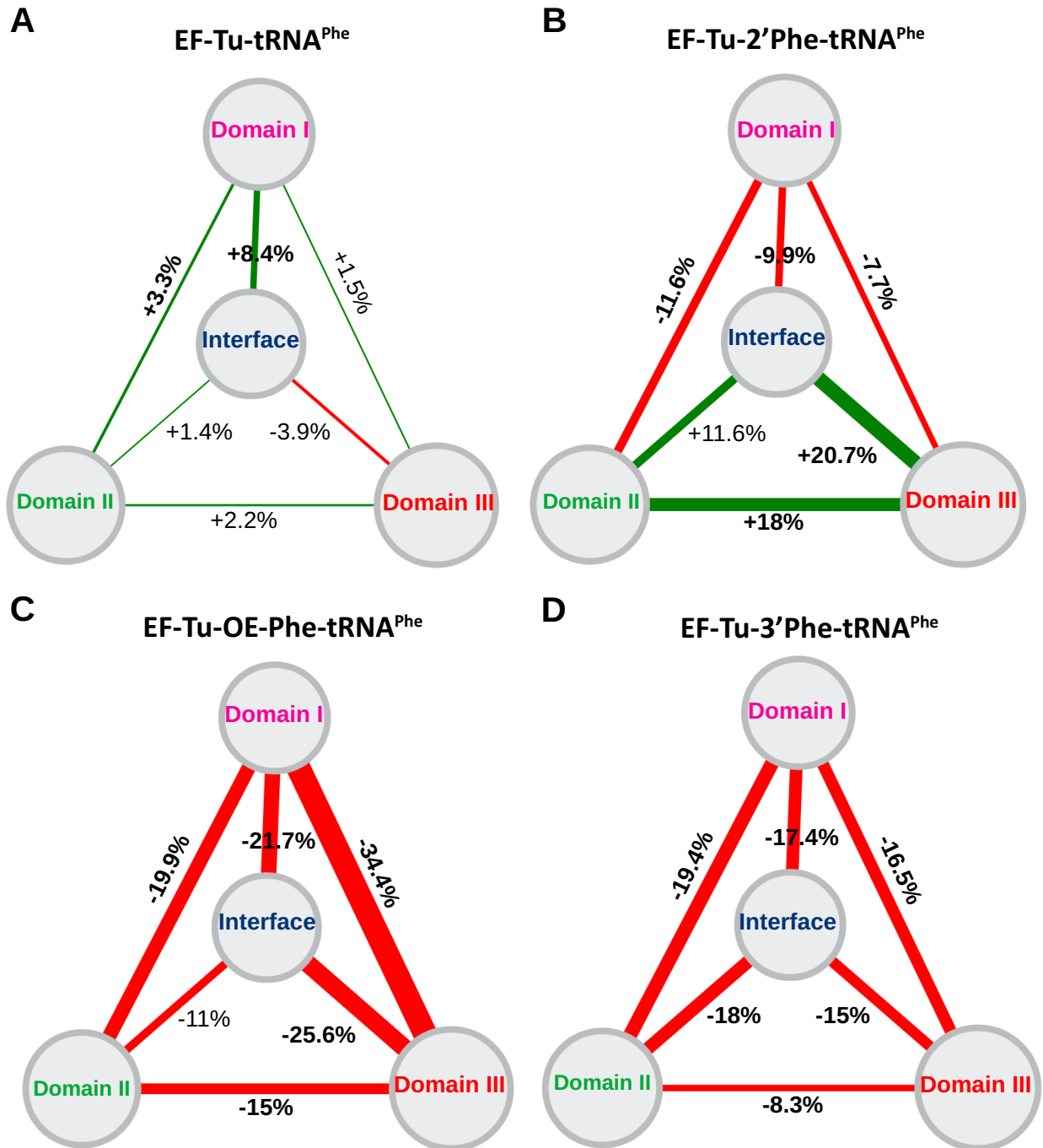


Figure 13: Relative change in MI between residue subsets identified by spectral clustering across tRNA charging isomer EF-Tu complexes (see also Figure 12). Clusters mapped to their majority domains with the cluster containing the switch regions denoted as interface. **A** tRNA^{Phe}·EF-Tu. **B** 2'Phe-tRNA^{Phe}·EF-Tu. **C** OE-Phe-tRNA^{Phe}·EF-Tu. **D** 3'Phe-tRNA^{Phe}·EF-Tu. Values above the noise level (see Table 8) for each channel are highlighted.

7.6.2 Mutual information from MSA

MI calculated from MD simulations was compared to the MI calculated from covarying positions in multiples sequence alignments (MSA) of tRNA and EF-Tu. For tRNA an MSA was generated with sequences from the manually curated Sprinzl et al. tRNAdb, guaranteeing an overlap of canonical positions of tRNA as well as accurate annotation of modified bases. All elongator tRNA species from all organisms were included to identify the most general trends in positional covariation possible. Some features of the MI maps correspond between the MD MI and those from sequence alignment (Figure 14).

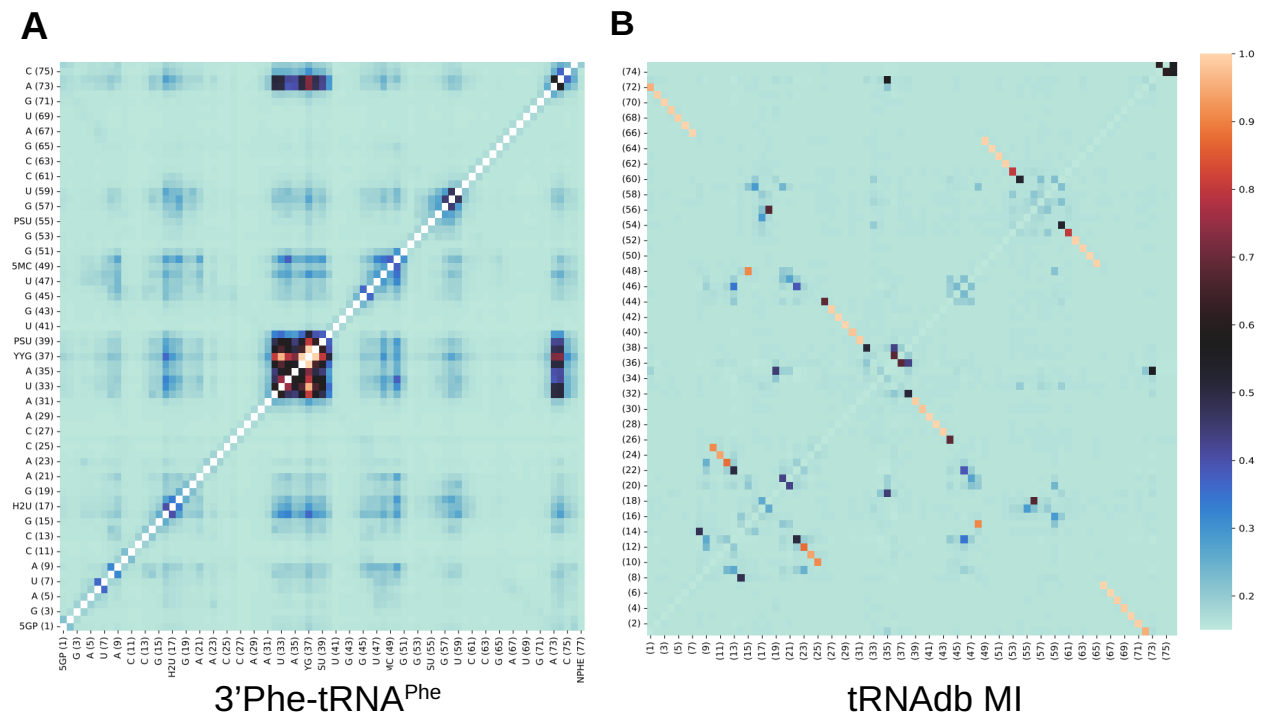


Figure 14: Comparison of pairwise MI maps generated from MD simulations and tRNA sequence alignment. **A** MI map from 3'Phe-tRNA^{Phe} MD trajectory. **B** Pairwise three dimensional MI between nucleotide positions calculated from all annotated tRNA sequences in the Sprinzl et al. tRNAdb.

The highest MI between positions calculated from the tRNA MSA generally corresponds to contacts (Supplement 8). The coupling between the anticodon and the acceptor stem observed from MD simulations is visible as well as the coupling between both T-stems and certain D-stem residues. Two different MSA's for EF-Tu were generated using the deepMSA2 server and mmseqs2 with the *T. aquaticus* EF-Tu sequence (Figure 15). Visualization of the pairwise MI map and clustering analogously to the EF-Tu MD trajectories, reveals an almost one to one mapping of MI clusters to the three EF-Tu domains as well as a fourth cluster containing the two switch regions (Figure 15B) when clustering the MI map calculated from the mmseqs2 generated MSA.

Only one alpha helix between domains I and II is clustered together with domain II. Results based on the deepMSA2 alignment have a less strict distinction between domains in the clustering with some overlap between domains I and II as well as II and III (Figure 15A).

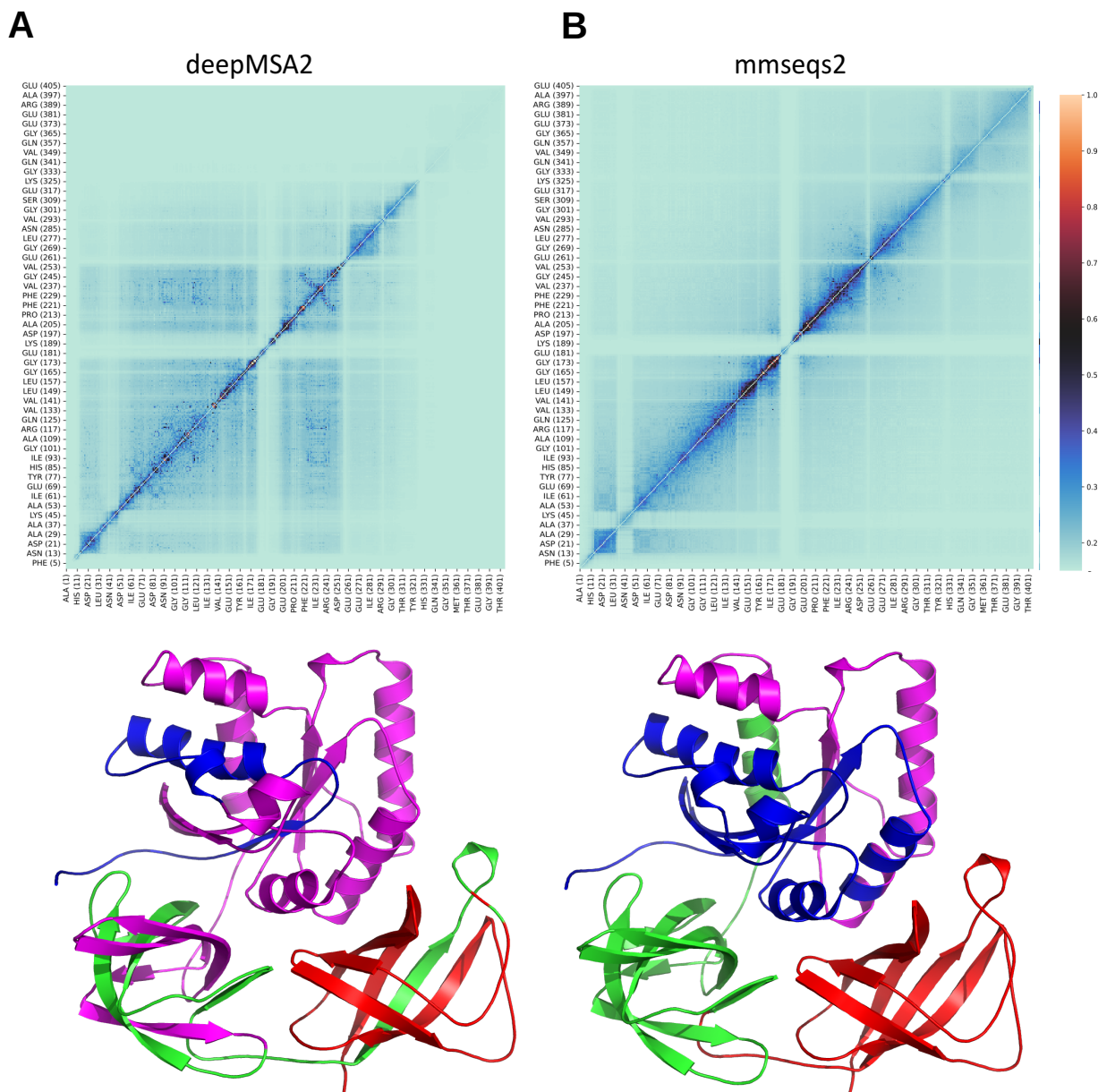


Figure 15: Pairwise MI maps calculated from MSA mapped onto the *T. aquaticus* EF-Tu sequence. **A** Pairwise MI map for residue positions annotated from *T. aquaticus* sequence and result of spectral clustering of pairwise MI. MSA generated by deepMSA2. **B** Pairwise MI map for residue positions annotated from *T. aquaticus* sequence and result of spectral clustering of pairwise MI. MSA generated by mmseqs2.

8. Discussion

The impact of aminoacylation isomerism on several conformational and dynamic aspects of tRNA^{Phe} and tRNA^{Phe}·EF-Tu complexes has been investigated in this work. Several aspects have been found to vary between isomers and some general qualities of tRNA dynamics and conformation (to the authors knowledge) not explicitly described before have been observed.

8.1 Impact of attachment isomerism on tRNA and EF-Tu

The impact of this 2'-/3'-isomerism on the dynamics and structure of tRNA have previously been investigated via aminoacylated RNA trinucleotides as a tRNA model by Polyansky et al. With the present results, some further steps in characterizing the impact of charging isomerism not only on full scale tRNA but its essential binding partner in early translation, EF-Tu have been taken. The aminoacylation site as well as the binding of EF-Tu are shown here to have a strong impact on solvent exposure of the aminoacyl ester carbonyl group as a proxy for hydrolytic stability of the aminoacyl bond. The orthoester (OE) isomer displays the highest carbonyl solvent exposure, followed by the 2'-isomer and the 3'-aminoacylated tRNA^{Phe} trailing behind. The binding of EF-Tu buries the carbonyl group in all but 3'-isomer aminoacyl-tRNA^{Phe}, where the opposite is the case (Figure 3, Table 4). The establishment of a distribution of a relatively high carbonyl exposure for 3'-Phe-tRNA^{Phe} after binding EF-Tu may indicate that the 3'-Phe-tRNA^{Phe}·EF-Tu forms a well-defined complex with a distribution around “intended” exposure of the aminoacyl bond for facilitating further steps in translation. Notably the highest absolute value for carbonyl exposure was observed for the free 3'-isomer tRNA^{Phe} trajectory. As Phe tRNA species are charged at the 2'-hydroxyl of A76, the 3'-isomer might be present only after binding EF-Tu or an isomer specifically recognized for binding by EF-Tu, but too hydrolytically unstable in free tRNA. The stability of the acyl linkage has been tested for many 3'-attached amino acids (including Phe) and has been shown to depend on the species of the attached amino acid³¹. The only variable that was able to protect the weaker linkages from hydrolysis was EF-Tu. The Phe linkage stability was found to be similar to that of lysine and tyrosine and as such slightly lower than the median, in agreement with relative stabilities reported by Hentzen et al.¹²⁴. These findings suggest that an isomer with higher exposure of the aminoacyl linkage to solvent than the 3'-isomer such as the OE-attachment would be likely even more prone to hydrolysis in the absence of EF-Tu. This could mean that the interconversion state is either very short lived in free tRNA, or may help to facilitate EF-Tu binding. The underlying geometric causes leading to this disparity in ester bond exposure are likely the distance between the 5'- and 3'- ends of the tRNA in general and the propensity of G1 – A76 hydrogen bonding and base ring stacking specifically. The 3'-OH of A76 looks to be disproportionally prone to forming hydrogen bonds to the G1 phosphate group with this imbalance being present in uncharged tRNA^{Phe} where both hydroxyl groups are available and further magnified in the presence of the amino acid.

The relative fraction of identified hydrogen bonds 3'-OH/2'-OH grows from about 5 to almost 10 when Phe is attached (Table 5). Additionally, the presence of hydrogen bonds between the G1 phosphate and the amino group was detected twice as often in the 2'-attached free tRNA^{Phe} trajectory compared to its 3'-counterpart (Table 5). Together both these interactions could be used to explain differences in 3'- and 5'-strand separation between charging isomers. Indeed when no A67 hydroxyl groups are available for hydrogen bonding to the G1 phosphate as is the case of the OE-isomer, the distance distribution between both strands is shifted towards higher distances even in the presence of a small amount of frames presenting with amino group-G1 hydrogen bonds (Figure 6A, Table 7). The stacking of base rings between G1 and A76 does not seem to be significantly impacted by the amino-acid attachment site between 2'- and 3'-OH with a difference of only 2%. However, the absence of any steric hindrance by the amino acid in uncharged tRNA^{Phe} increases the presence of this interaction to almost 70% of MD simulation frames (Table 6). The distance between the tRNA^{Phe} acceptor stem 3'- and 5'-strands is increased in all EF-Tu bound tRNA^{Phe} molecules (Figure 6B, Table 5) indicating the need for some degree of strand separation before binding EF-Tu. The OE- and 3'-attached tRNA^{Phe} might have a distinct advantage when binding EF-Tu as their strands are generally more separated due to lacking interactions between the acceptor stem ends. This is supported by an approximately 2250 fold lower binding affinity of uncharged yeast tRNA^{Phe} to *E. coli* EF-Tu·GTP¹²⁵ as the non-acylated tRNA^{Phe} system displayed the highest amount of strand interactions of all MD simulations. Early structural data suggested a possible OE configuration present in EF-Tu bound tRNA²⁷. One may speculate that the spontaneous inter-conversion between the attachment isomers is at some point exploited to facilitate easier accommodation of the tRNA into the EF-Tu binding site. This might not be necessary as the 3'-isomer that has previously been shown to be preferred by EF-Tu^{18,22} also shows a decreased amount of A76-G1 hydrogen bonding and higher acceptor stem strand separation compared to the 2'-isomer and the uncharged tRNA^{Phe} (Table 5, Table 7, Figure 6). As the binding of the 3'- and 5'-ends of the tRNA are key components in its interaction with EF-Tu, the accessibility of these termini might be an important factor in EF-Tu's ability to discriminate between charging isomers. Indeed the tRNA 5'-end while stabilized in all MD trajectories where EF-Tu is present, is uniquely stable in 3'-aminoacylated tRNA^{Phe} (Figure 2C).

EF-Tu is considered to bind poorly to tRNA^{Phe} lacking an attached amino acid¹²⁵, and both the tRNA body and the attached amino acid have been shown to contribute to the binding affinity in a compensatory way to achieve uniform binding affinity^{54,57,126}. More recent work shows that the largest part of the tRNA body contribution to the EF-Tu binding affinity can be attributed to three base-pairs in the tRNA T-stem (49-65, 50-64 and 51-63)^{63,127}. Additionally, computational work identified significant contributions to binding energies from G52, G52, and G53 in Cys-tRNA^{Cys}·EF-Tu¹²⁸. In the present work, only the attachment site of the amino acid is modified across different MD systems and as such it could be assumed that all T-stem contributions to the tRNA-EF-Tu interface

are unaffected. This might not be the case as the proposed point of contact between EF-Tu and the 51-63 base pair, Glu390 is located in a region of EF-Tu with differing RMSF across aminoacylation isomers (Figure 8). Specifically, residue 391 fluctuates less in 3'-Phe-tRNA^{Phe}·EF-Tu as compared to all other MD trajectories. This could indicate the existence of additional effects on the EF-Tu-tRNA binding interface that can not entirely be described by local sequence variation. The contributions of the tRNA T-stem interface at EF-Tu Glu391 to the binding affinity might be subject to tuning through changes in conformation brought about by the attachment isomerism. Computational work identified the universally conserved CCA acceptor triplet as well as the highly conserved G1 base as major contributors to tRNA-EF-Tu binding¹²⁸. The interactions between G1 and A76 described here might explain in part the relevance of these bases for the tRNA-EF-Tu affinity. Especially the impact on the separation of strands in tRNA could be relevant for EF-Tu binding as both tRNA ends are bound separately by EF-Tu³⁵.

MD conformations of tRNA^{Phe} were investigated by comparison to cryo EM structures during different steps of decoding from Loveland et al⁴⁵ (Supplement 9). Distribution of backbone RMSD during 1 μ s MD for the complete tRNA^{Phe} molecule was closest to the structure termed 'initial binding' for all MD trajectories. 3'-Phe-tRNA^{Phe}·EF-Tu had the highest fraction of frames within 4.5 Å backbone RMSD (39.9%). The 2'-isomer was basically at the same level of similarity to this reference at 39.2%. For the OE-Phe-tRNA^{Phe} isomer in complex with EF-Tu, this fraction dropped to 6.2% and similarly 7.7% for the uncharged-tRNA^{Phe}·EF-Tu complex. Fractions for unbound tRNA^{Phe} ranged from 9-20%. In the case of OE- and 3'-isomer, the number of frames satisfying the RMSD cutoff was similar at ~14% while 2'-Phe-tRNA^{Phe} MD trajectory contained only 9% of frames satisfying the cutoff. Similarity to structures in subsequent stages of translation, such as codon/anticodon sampling and tRNA accommodation, was much lower in terms of backbone RMSD across all MD trajectories potentially reflecting the influence of the ribosome on tRNA conformation. Together these results indicate that the EF-Tu tRNA MD simulations adopt conformations close to experimental data to a substantial amount. Additionally, the presence and attachment site of the amino acid influence the global tRNA^{Phe} backbone conformation, as seen by the differences in RMSD between charging isomer MD simulations and experimental reference structure. For example, the OE-aminoacylated tRNA^{Phe} bound to EF-Tu satisfies the RMSD similarity cutoff in just 6% of frames compared to the 3'-aminoacylated system (39.9%).

EF-Tu has been shown to bind both 2'- and 3'-attachment isomers but the non native attachment site has an about 50 fold decreased affinity^{18,28,29}. EF-Tu therefore effectively acts as a tRNA attachment isomerase and transforms the mixture of 2'- and 3'-isomers present after aminoacylation into a uniform set of 3'-aa-tRNA·EF-Tu species²². The earlier mentioned conformational aspects of tRNA charging isomers might help to explain the preferential binding of 3'-aminoacylated tRNA^{Phe}, while the possibility of binding the other isomer makes observations of EF-Tu bound to 2'-Phe-tRNA^{Phe} relevant to elucidating the EF-Tu isomerase activity. The differences in mutual information (MI)

between amino acid subgroups from clustering with limited prior knowledge might therefore point to residues critical in communicating the precise attachment site of the amino acid in EF-Tu to facilitate the conversion. The detected clusters can be mapped to established domains and functional elements in EF-Tu (Figure 1A, Figure 12) and a clear reduction in correlated conformations between these domains is visible in both 3'- and OE-aminoacylated tRNA^{Phe} bound to EF-Tu (Figure 13C, Figure 13D). On the other hand, the MI between the clusters containing the tRNA binding interface as well as the cluster containing the switch II region is increased in 2'-Phe-tRNA^{Phe}-EF-Tu complex MD simulations (Figure 13B). This indicates that EF-Tu is able to integrate the information about the attachment site of the bound tRNA, adapt its internal communication and this might be how the isomerase activity is triggered. The increased independence of domain movements visible from the decrease in MI between the identified amino acid subgroups, might be necessary to facilitate the relative domain movements occurring before GTP hydrolysis. EF-Tu MD simulations analysis showed a separation of domain I from domains II and III, extending the linker and making the inter domain rotation necessary for tRNA release possible¹²⁹. In contrast, EF-Tu shows very little change in MI between the clusters in complex with uncharged tRNA^{Phe} compared to single EF-Tu (Figure 13A). This further strengthens the role of the amino acid attachment in modulating the intramolecular response to tRNA binding.

Calculation of configurational entropy suggests an entropic benefit to the 3'-attachment of the amino acid compared to both other isomers (Figure 7C). This could be a contributing factor to the imbalance between attachment isomers observed in NMR experiments to be in favor of the 3'-isomer²⁶. The complex of OE-Phe-tRNA^{Phe}-EF-Tu increases the contribution of configurational entropy to free energy $T\Delta S$ for both of its components by about 23 kJ/mol (Figure 7, Figure 9) at 310K. The binding energy of aminoacylated tRNA to EF-Tu has been calculated at around 40 kJ/mol and is almost uniform among all elongator tRNA species¹²⁶. Misacylated tRNA varies widely in its binding free energy due to the contribution of the amino acid and it is not clear if this difference might permit misacylated tRNA to bind in the OE configuration due to the entropic benefit even if such a configuration was stable enough to participate in binding EF-Tu before converting to either 2'- or 3'-isomer. The difference in the binding free energy between tRNA^{Gln} misacylated with alanine was calculated as 9.6 kJ/mol which could be theoretically compensated¹²⁶. This data was obtained through misacylation of unmodified and mutated tRNA (to allow for misacylation) and the configurational entropy calculated in this work only gives a correct upper bound but may well be overestimated due to incomplete sampling.

The clustering of MI values taken as a measure of similarity between residues corresponds well to the sequence- and structure-based categorization of EF-Tu domains and motifs (Figure 12). This increases the confidence in the mutual information values calculated from MD as a measure of biologically relevant coupling between residues.

This is further strengthened by the good agreement of clustering between MI from MD and MI from MSA in both tRNA and EF-Tu (Figure 14, Figure 15).

The presence of high MI between both sequence position and degrees of freedom between the anticodon and the acceptor triplet in all tRNA^{Phe} MD trajectories is a noteworthy observation (Figure 10, Figure 11, Figure 14). The MI between sequence positions is usually hard to decompose into covariation arising from sterics, contacts, phylogenetic history, interactions between charges, or long range communication/allosteric coupling. However, the acceptor stem and anticodon loop are far apart in tRNA, so any sort of steric or electrostatic influences can likely be dismissed. It seems that this connection might be significant enough to be encoded in MI from both sequence data as well as atomic movements from MD simulation. The direct coupling between the most distant parts of the molecule points to a high degree of intermolecular communication in tRNA. It is further noteworthy that the acceptor stem directly attaches to the amino acid. This unifies the link between anticodon and amino acid through the genetic code with a link in the dynamical behavior of the adapter molecule. This link is also present when clustering the MI from the 3'-Phe-tRNA^{Phe}·EF-Tu MD trajectory (Supplement 6). This directly links the codon sensing part of the ternary complex to the acceptor stem influencing crucial EF-Tu residues as described in this work.

The profile of RMSF in tRNA^{Phe} is consistent across isomers and presence or absence of EF-Tu in the anticodon displaying high fluctuation (Figure 2C). Finally, the RMSF profile showing an increase in fluctuation around the anticodon loop and residue H2U17 of the D-loop, is consistent with cryo-EM experimental findings of distortion in these regions when bound to GTPase activated EF-Tu⁶⁻⁸. Taken together with the need for an intact tRNA present to trigger the EF-Tu GTP hydrolysis⁹, the present results might add further evidence supporting a non-static, active role for tRNA in translation possibly necessitating tRNA conformational flexibility, especially in the anticodon loop, for EF-Tu GTPase activity and proper accommodation of cognate species during decoding.

8.2 Limitations and outlook

The three strongest findings from the present work are: (1) the attachment position and presence of phenylalanine on tRNA^{Phe}, although being a small modification, strongly impacts conformation of the tRNA^{Phe} acceptor stem; (2) these differences robustly impact MI between functional domains of EF-Tu in ternary complex validated by the establishment of a random noise thresholds through replicate MD trajectories and (3) tRNA dynamics, as captured by both RMSF and MI profiles, points to long-range correlated conformations of the anticodon and the acceptor stem regardless of amino acid attachment site. This finding is corroborated by sequence data. Limiting the confidence in these findings are the lack of replicates for most systems due to high computational requirements and time constraints. Findings regarding the fundamental

geometric aspects of the attachment isomers are large enough in magnitude to be considered relevant even in the absence of replicate MD trajectories. However, especially the calculation of configurational entropy and MI is critically dependent on both sampling (as has been exploited for the systemic noise reduction in this work) and more total simulated time as well as smaller time steps for trajectory output would definitely increase the accuracy for these calculations. A number of possible implications concerning both conformational aspects (such as carbonyl solvent accessible surface area and 3', 5'-strand separation) and the results on the dynamics of the elements of the ternary complex (stabilization of the 5'-end of 3'Phe-tRNA^{Phe} in complex with EF-Tu compared to other isomer ternary complexes) remain speculative until experimental validation. Others such as the clustering of MI between residues into communities of high internal MI are validated by their agreement with MI maps from MSA, combining a well studied and fundamental quantity from information theory, with data generated from experiment and processed by mature bioinformatic formalisms. The workflow for analyzing MI between residues from MD simulations presented here might be an interesting addition to investigations of dynamic and allosteric properties of proteins and other bio molecules, especially if considerations of sampling requirements are taken into account in MD simulation design e.g in terms of trajectory output. To further investigate the tRNA·EF-Tu·GTP ternary complex, the existing approaches from network theory for investigating vertex centrality or network path analysis could be employed or adapted to suit the investigation of functional couplings in biomolecules represented as weighted graphs. Some modifications in consolidating degrees of freedom into biologically relevant units might be worth considering to eliminate/minimize bias introduced by differing numbers of atoms per unit, especially in the presence of outliers such as GTP or strongly modified tRNA bases in the investigated systems. This is partially remedied already, because such residues contribute disproportionately to systemic noise and the MI bias from residue size is dampened through the noise reduction via artificial reduction in sampling as implemented in ARTEMIS.

9. References

1. Le Quesne, J. P., Spriggs, K. A., Bushell, M. & Willis, A. E. Dysregulation of protein synthesis and disease. *J. Pathol.* **220**, 140–151 (2010).
2. Alberts, B. *et al.* From RNA to Protein. in *Molecular Biology of the Cell*. 4th edition (Garland Science, 2002).
3. Uemura, S. *et al.* Real-time tRNA transit on single translating ribosomes at codon resolution. *Nature* **464**, 1012–1017 (2010).
4. Agirrezabala, X. & Valle, M. Structural Insights into tRNA Dynamics on the Ribosome. *Int. J. Mol. Sci.* **16**, 9866–9895 (2015).
5. Rodnina, M. V., Fricke, R. & Wintermeyer, W. Transient Conformational States of Aminoacyl-tRNA during Ribosome Binding Catalyzed by Elongation Factor Tu. *Biochemistry* **33**, 12267–12275 (1994).
6. Valle, M. *et al.* Cryo-EM reveals an active role for aminoacyl-tRNA in the accommodation process. *EMBO J.* **21**, 3557–3567 (2002).
7. Valle, M. *et al.* Incorporation of aminoacyl-tRNA into the ribosome as seen by cryo-electron microscopy. *Nat. Struct. Biol.* **10**, 899–906 (2003).
8. Li, W. *et al.* Recognition of aminoacyl-tRNA: a common molecular mechanism revealed by cryo-EM. *EMBO J.* **27**, 3322–3331 (2008).
9. Piepenburg, O. *et al.* Intact Aminoacyl-tRNA Is Required To Trigger GTP Hydrolysis by Elongation Factor Tu on the Ribosome. *Biochemistry* **39**, 1734–1738 (2000).
10. Fei, J., Richard, A. C., Bronson, J. E. & Gonzalez, R. L. Transfer RNA-mediated regulation of ribosome dynamics during protein synthesis. *Nat. Struct. Mol. Biol.* **18**, 1043–1051 (2011).
11. Blanchard, S. C., Kim, H. D., Gonzalez, R. L., Puglisi, J. D. & Chu, S. tRNA dynamics on the ribosome during translation. *Proc. Natl. Acad. Sci.* **101**, 12893–12898 (2004).
12. Allen, G. S., Zavialov, A., Gursky, R., Ehrenberg, M. & Frank, J. The Cryo-EM Structure of a Translation Initiation Complex from *Escherichia coli*. *Cell* **121**, 703–712 (2005).
13. Myasnikov, A. G. *et al.* Conformational transition of initiation factor 2 from the GTP- to GDP-bound state visualized on the ribosome. *Nat. Struct. Mol. Biol.* **12**, 1145–1149 (2005).
14. Julián, P. *et al.* The Cryo-EM Structure of a Complete 30S Translation Initiation Complex from *Escherichia coli*. *PLOS Biol.* **9**, e1001095 (2011).
15. Hopfield, J. J. Kinetic Proofreading: A New Mechanism for Reducing Errors in Biosynthetic Processes Requiring High Specificity. *Proc. Natl. Acad. Sci.* **71**, 4135–4139 (1974).
16. Ogle, J. M., Murphy, F. V., Tarry, M. J. & Ramakrishnan, V. Selection of tRNA by the Ribosome Requires a Transition from an Open to a Closed Form. *Cell* **111**, 721–732 (2002).
17. Rubio Gomez, M. A. & Ibba, M. Aminoacyl-tRNA synthetases. *RNA* **26**, 910–936 (2020).
18. Sprinzl, M. & Cramer, F. Site of aminoacylation of tRNAs from *Escherichia coli* with respect to the 2'- or 3'-hydroxyl group of the terminal adenosine. *Proc. Natl. Acad. Sci.* **72**, 3049–3053 (1975).

19. Ruff, M. *et al.* Class II Aminoacyl Transfer Rna Synthetases: Crystal Structure of Yeast Aspartyl-trna Synthetase Complexed with tRNA^{Asp}. *Science* **252**, 1682–1689 (1991).
20. Safro, M., Moor, N. & Lavrik, O. Phenylalanyl-tRNA Synthetases. in *Madame Curie Bioscience Database [Internet]* (Landes Bioscience, 2013).
21. Sprinzl, M. Elongation factor Tu: a regulatory GTPase with an integrated effector. *Trends Biochem. Sci.* **19**, 245–250 (1994).
22. Sprinzl, M. Chemistry of aminoacylation and peptide bond formation on the 3'-terminus of tRNA. *J. Biosci.* **31**, 489–496 (2006).
23. Taiji, M., Yokoyama, S. & Miyazawa, T. Slow transacylation of peptidyladenosine allows analysis of the 2'/3'-isomer specificity of peptidyltransferase. *Biochemistry* **24**, 5776–5780 (1985).
24. Weinger, J. S. & Strobel, S. A. Participation of the tRNA A76 Hydroxyl Groups throughout Translation. *Biochemistry* **45**, 5939–5948 (2006).
25. Nordin, B. E. & Schimmel, P. Plasticity of Recognition of the 3'-End of Mischarged tRNA by Class I Aminoacyl-tRNA Synthetases *. *J. Biol. Chem.* **277**, 20510–20517 (2002).
26. Taiji, M., Yokoyama, S. & Miyazawa, T. Transacylation rates of aminoacyladeniosine moiety at the 3'-terminus of aminoacyl-transfer ribonucleic acid. *Biochemistry* **22**, 3220–3225 (1983).
27. Förster, C., Limmer, S., Zeidler, W. & Sprinzl, M. Effector region of the translation elongation factor EF-Tu.GTP complex stabilizes an orthoester acid intermediate structure of aminoacyl-tRNA in a ternary complex. *Proc. Natl. Acad. Sci.* **91**, 4254–4257 (1994).
28. Hecht, S. M., Tan, K. H., Chinault, A. C. & Arcari, P. Isomeric aminoacyl-tRNAs are both bound by elongation factor Tu. *Proc. Natl. Acad. Sci.* **74**, 437–441 (1977).
29. Alford, B. L., Pezzuto, J. M., Tan, K. H. & Hecht, S. M. Both positional isomers of aminoacyl-tRNA's are bound by elongation factor Tu. *J. Biol. Chem.* **254**, 6894–6903 (1979).
30. Pleiss, J. A. & Uhlenbeck, O. C. Identification of thermodynamically relevant interactions between EF-Tu and backbone elements of tRNA. *J. Mol. Biol.* **308**, 895–905 (2001).
31. Peacock, J. R. *et al.* Amino acid-dependent stability of the acyl linkage in aminoacyl-tRNA. *RNA* **20**, 758–764 (2014).
32. Polyansky, A. A., Kreuter, M., Sutherland, J. D. & Zagrovic, B. Direct interplay between stereochemistry and conformational preferences in aminoacylated oligoribonucleotides. *Nucleic Acids Res.* **47**, 11077–11089 (2019).
33. Kavaliauskas, D., Nissen, P. & Knudsen, C. R. The Busiest of All Ribosomal Assistants: Elongation Factor Tu. *Biochemistry* **51**, 2642–2651 (2012).
34. Gromadski, K. B., Wieden, H.-J. & Rodnina, M. V. Kinetic Mechanism of Elongation Factor Ts-Catalyzed Nucleotide Exchange in Elongation Factor Tu. *Biochemistry* **41**, 162–169 (2002).
35. Nissen, P. *et al.* Crystal Structure of the Ternary Complex of Phe-tRNA^{Phe}, EF-Tu, and a GTP Analog. *Science* **270**, 1464–1472 (1995).
36. Berchtold, H. *et al.* Crystal structure of active elongation factor Tu reveals major domain rearrangements. *Nature* **365**, 126–132 (1993).

37. Polekhina, G. *et al.* Helix unwinding in the effector region of elongation factor EF-Tu-GDP. *Structure* **4**, 1141–1151 (1996).
38. Ogle, J. M. & Ramakrishnan, V. STRUCTURAL INSIGHTS INTO TRANSLATIONAL FIDELITY. *Annu. Rev. Biochem.* **74**, 129–177 (2005).
39. Schmeing, T. M. *et al.* The Crystal Structure of the Ribosome Bound to EF-Tu and Aminoacyl-tRNA. *Science* **326**, 688–694 (2009).
40. Kothe, U., Wieden, H.-J., Mohr, D. & Rodnina, M. V. Interaction of Helix D of Elongation Factor Tu with Helices 4 and 5 of Protein L7/12 on the Ribosome. *J. Mol. Biol.* **336**, 1011–1021 (2004).
41. Diaconu, M. *et al.* Structural Basis for the Function of the Ribosomal L7/12 Stalk in Factor Binding and GTPase Activation. *Cell* **121**, 991–1004 (2005).
42. Vorstenbosch, E., Pape, T., Rodnina, M. V., Kraal, B. & Wintermeyer, W. The G222D mutation in elongation factor Tu inhibits the codon-induced conformational changes leading to GTPase activation on the ribosome. *EMBO J.* **15**, 6766–6774 (1996).
43. Voorhees, R. M., Schmeing, T. M., Kelley, A. C. & Ramakrishnan, V. The Mechanism for Activation of GTP Hydrolysis on the Ribosome. *Science* **330**, 835–838 (2010).
44. Loveland, A. B., Demo, G., Grigorieff, N. & Korostelev, A. A. Ensemble cryo-EM elucidates the mechanism of translation fidelity. *Nature* **546**, 113–117 (2017).
45. Loveland, A. B., Demo, G. & Korostelev, A. A. Cryo-EM of elongating ribosome with EF-Tu•GTP elucidates tRNA proofreading. *Nature* **584**, 640–645 (2020).
46. Liu, W. *et al.* EF-Tu dynamics during pre-translocation complex formation: EF-Tu-GDP exits the ribosome via two different pathways. *Nucleic Acids Res.* **43**, 9519–9528 (2015).
47. Noel, J. K. & Whitford, P. C. How EF-Tu can contribute to efficient proofreading of aa-tRNA by the ribosome. *Nat. Commun.* **7**, 13314 (2016).
48. Warias, M., Grubmüller, H. & Bock, L. V. tRNA Dissociation from EF-Tu after GTP Hydrolysis: Primary Steps and Antibiotic Inhibition. *Biophys. J.* **118**, 151–161 (2020).
49. Jared M. Schrader, Margaret E. Saks, Olke C. Uhlenbeck. The specific interaction between aminoacyl-tRNAs and elongation factor Tu. in *Ribosomes: structure, function, and dynamics* (eds. Rodnina, M. V., Wintermeyer, W. & Green, R.) (Springer, Wien ; New York, 2011).
50. Kjeldgaard, M., Nyborg, J. & Clark, B. F. The GTP binding motif: variations on a theme. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* **10**, 1347–1368 (1996).
51. Parmeggiani, A. *et al.* Properties of a genetically engineered G domain of elongation factor Tu. *Proc. Natl. Acad. Sci.* **84**, 3141–3145 (1987).
52. Kothe, U. & Rodnina, M. V. Delayed Release of Inorganic Phosphate from Elongation Factor Tu Following GTP Hydrolysis on the Ribosome. *Biochemistry* **45**, 12767–12774 (2006).
53. Schümmer, T., Gromadski, K. B. & Rodnina, M. V. Mechanism of EF-Ts-Catalyzed Guanine Nucleotide Exchange in EF-Tu: Contribution of Interactions Mediated by Helix B of EF-Tu. *Biochemistry* **46**, 4977–4984 (2007).
54. Louie, A., Ribeiro, N. S., Reid, B. R. & Jurnak, F. Relative affinities of all Escherichia coli aminoacyl-tRNAs for elongation factor Tu-GTP. *J. Biol. Chem.* **259**, 5010–5016 (1984).

55. Louie, A. & Jurnak, F. Kinetic studies of Escherichia coli elongation factor Tu-guanosine 5'-triphosphate-aminoacyl-tRNA complexes. *Biochemistry* **24**, 6433–6439 (1985).
56. Ott, G. *et al.* Ternary complexes of Escherichia coli aminoacyl-tRNAs with the elongation factor Tu and GTP: Thermodynamic and structural studies. *Biochim. Biophys. Acta BBA - Gene Struct. Expr.* **1050**, 222–225 (1990).
57. Asahara, H. & Uhlenbeck, O. C. The tRNA Specificity of *Thermus thermophilus* EF-Tu. *Proc. Natl. Acad. Sci.* **99**, 3499–3504 (2002).
58. Jakubowski, H. Negative correlation between the abundance of Escherichia coli aminoacyl-tRNA families and their affinities for elongation factor Tu-GTP. *J. Theor. Biol.* **133**, 363–370 (1988).
59. Asahara, H. & Uhlenbeck, O. C. Predicting the Binding Affinities of Misacylated tRNAs for *Thermus thermophilus* EF-Tu-GTP. *Biochemistry* **44**, 11254–11261 (2005).
60. Dale, T., Sanderson, L. E. & Uhlenbeck, O. C. The Affinity of Elongation Factor Tu for an Aminoacyl-tRNA Is Modulated by the Esterified Amino Acid. *Biochemistry* **43**, 6159–6166 (2004).
61. Sanderson, L. E. & Uhlenbeck, O. C. Directed Mutagenesis Identifies Amino Acid Residues Involved in Elongation Factor Tu Binding to yeast Phe-tRNAPhe. *J. Mol. Biol.* **368**, 119–130 (2007).
62. Sanderson, L. E. & Uhlenbeck, O. C. Exploring the Specificity of Bacterial Elongation Factor Tu for Different tRNAs. *Biochemistry* **46**, 6194–6200 (2007).
63. Schrader, J. M., Chapman, S. J. & Uhlenbeck, O. C. Understanding the Sequence Specificity of tRNA Binding to Elongation Factor Tu using tRNA Mutagenesis. *J. Mol. Biol.* **386**, 1255–1264 (2009).
64. Fischer, W., Doi, T., Ikehara, M., Ohtsuka, E. & Sprinzl, M. Interaction of methionine-specific tRNAs from *Escherichia coli* with immobilized elongation factor Tu. *FEBS Lett.* **192**, 151–154 (1985).
65. Schulman, L. H., Pelka, H. & Sundari, R. M. Structural requirements for recognition of Escherichia coli initiator and non-initiator transfer ribonucleic acids by bacterial T factor. *J. Biol. Chem.* **249**, 7102–7110 (1974).
66. Seong, B. L. & RajBhandary, U. L. Mutants of Escherichia coli formylmethionine tRNA: a single base change enables initiator tRNA to act as an elongator in vitro. *Proc. Natl. Acad. Sci.* **84**, 8859–8863 (1987).
67. Karplus, M. & McCammon, J. A. Molecular dynamics simulations of biomolecules. *Nat. Struct. Biol.* **9**, 646–652 (2002).
68. Lindorff-Larsen, K. *et al.* Systematic validation of protein force fields against experimental data. *PLoS One* **7**, e32131 (2012).
69. Salomon-Ferrer, R., Götz, A. W., Poole, D., Le Grand, S. & Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *J. Chem. Theory Comput.* **9**, 3878–3888 (2013).
70. Cantara, W. A. *et al.* The RNA modification database, RNAMDB: 2011 update. *Nucleic Acids Res.* **39**, D195–D201 (2011).
71. Zhang, X., Walker, R. C., Phizicky, E. M. & Mathews, D. H. Influence of Sequence and Covalent Modifications on Yeast tRNA Dynamics. *J. Chem. Theory Comput.* **10**, 3473–3483 (2014).

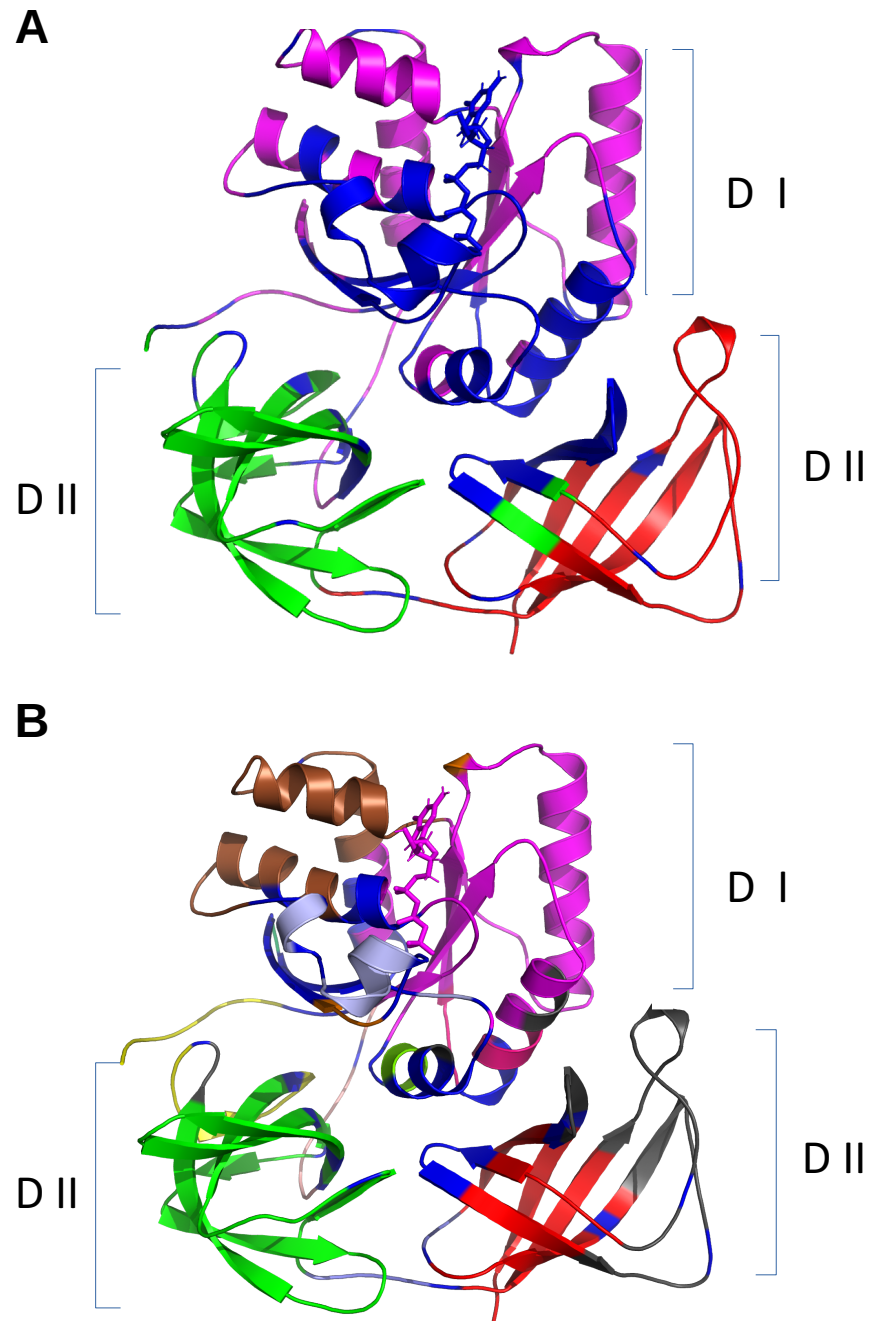
72. Björk, G. R. *et al.* Transfer RNA modification: influence on translational frameshifting and metabolism. *FEBS Lett.* **452**, 47–51 (1999).
73. Persson, B. C. Modification of tRNA as a regulatory device. *Mol. Microbiol.* **8**, 1011–1016 (1993).
74. Urbonavičius, J., Qian, Q., Durand, J. M. B., Hagervall, T. G. & Björk, G. R. Improvement of reading frame maintenance is a common function for several tRNA modifications. *EMBO J.* **20**, 4863–4873 (2001).
75. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **25**, 1157–1174 (2004).
76. Bryce, R. AMBER parameter database.
77. Girodat, D., Blanchard, S. C., Wieden, H.-J. & Sanbonmatsu, K. Y. Elongation Factor-Tu switch I is a gate for aminoacyl-tRNA selection. *J. Mol. Biol.* **432**, 3064–3077 (2020).
78. Colell, E. A., Iserte, J. A., Simonetti, F. L. & Marino-Buslje, C. MISTIC2: comprehensive server to study coevolution in protein families. *Nucleic Acids Res.* **46**, W323–W328 (2018).
79. Ackerman, S. H., Tillier, E. R. & Gatti, D. L. Accurate simulation and detection of coevolution signals in multiple sequence alignments. *PLoS One* **7**, e47108 (2012).
80. McClendon, C. L., Friedland, G., Mobley, D. L., Amirkhani, H. & Jacobson, M. P. Quantifying Correlations Between Allosteric Sites in Thermodynamic Ensembles. *J. Chem. Theory Comput.* **5**, 2486–2502 (2009).
81. Louvaris, D., Rahman, N. & Nasif, A. O. An Approach of Mutual Information-based Characterization of Allosteric Macromolecules. in *2022 IEEE 22nd International Conference on Nanotechnology (NANO)* 79–82 (IEEE, Palma de Mallorca, Spain, 2022). doi:10.1109/NANO54668.2022.9928758.
82. Sogunmez, N. & Akten, E. D. Distinctive communication networks in inactive states of β_2 - adrenergic receptor: Mutual information and entropy transfer analysis. *Proteins* **88**, 1458–1471 (2020).
83. Lin, M. M. Timing Correlations in Proteins Predict Functional Modules and Dynamic Allostery. *J. Am. Chem. Soc.* **138**, 5036–5043 (2016).
84. LeVine, M. V. & Weinstein, H. AIM for Allostery: Using the Ising Model to Understand Information Processing and Transmission in Allosteric Biomolecular Systems. *Entropy Basel Switz.* **17**, 2895–2918 (2015).
85. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
86. Clark, G. W., Ackerman, S. H., Tillier, E. R. & Gatti, D. L. Multidimensional mutual information methods for the analysis of covariation in multiple sequence alignments. *BMC Bioinformatics* **15**, 157 (2014).
87. Bradley, M. J., Chivers, P. T. & Baker, N. A. Molecular dynamics simulation of the Escherichia coli NikR protein: equilibrium conformational fluctuations reveal interdomain allosteric communication pathways. *J. Mol. Biol.* **378**, 1155–1173 (2008).

88. Chennubhotla, C., Yang, Z. & Bahar, I. Coupling between global dynamics and signal transduction pathways: a mechanism of allostery for chaperonin GroEL. *Mol. Biosyst.* **4**, 287–292 (2008).
89. Zheng, W., Brooks, B. R. & Thirumalai, D. Low-frequency normal modes that describe allosteric transitions in biological nanomachines are robust to sequence variations. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 7664–7669 (2006).
90. Mallaev, R. <https://github.com/nalsur-veallam/ARTEMIS>. (2024).
91. Tzeng, S.-R. & Kalodimos, C. G. Protein activity regulation by conformational entropy. *Nature* **488**, 236–240 (2012).
92. Freire, E. Do enthalpy and entropy distinguish first in class from best in class? *Drug Discov. Today* **13**, 869–874 (2008).
93. Frederick, K. K., Marlow, M. S., Valentine, K. G. & Wand, A. J. Conformational entropy in molecular recognition by proteins: A first example from the calmodulin system. *Nature* **448**, 325–329 (2007).
94. Marlow, M. S., Dogan, J., Frederick, K. K., Valentine, K. G. & Wand, A. J. The role of conformational entropy in molecular recognition by calmodulin. *Nat. Chem. Biol.* **6**, 352–358 (2010).
95. Hnizdo, V. & Gilson, M. K. Thermodynamic and Differential Entropy under a Change of Variables. *Entropy Basel Switz.* **12**, 578–590 (2010).
96. Killian, B. J., Kravitz, J. Y. & Gilson, M. K. Extraction of Configurational Entropy from Molecular Simulations via an Expansion Approximation. *J. Chem. Phys.* **127**, 024107 (2007).
97. King, B. M., Silver, N. W. & Tidor, B. Efficient calculation of molecular configurational entropies using an information theoretic approximation. *J. Phys. Chem. B* **116**, 2891–2904 (2012).
98. King, B. M. & Tidor, B. MIST: Maximum Information Spanning Trees for dimension reduction of biological data sets. *Bioinforma. Oxf. Engl.* **25**, 1165–1172 (2009).
99. Elements of Information Theory, 2nd Edition | Wiley. *Wiley.com*
<https://www.wiley.com/en-it/Elements+of+Information+Theory%2C+2nd+Edition-p-9780471241959>.
100. Fleck, M., Polyansky, A. A. & Zagrovic, B. PARENT: A Parallel Software Suite for the Calculation of Configurational Entropy in Biomolecular Systems. *J. Chem. Theory Comput.* **12**, 2055–2065 (2016).
101. Prim, R. C. Shortest Connection Networks And Some Generalizations. *Bell Syst. Tech. J.* **36**, 1389–1401 (1957).
102. Abraham, M. J. *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25 (2015).
103. Abraham, M. *et al.* GROMACS 2023.2 Manual. (2023) doi:10.5281/zenodo.8134388.
104. Shi, H. & Moore, P. B. The crystal structure of yeast phenylalanine tRNA at 1.93 Å resolution: A classic structure revisited. *RNA* **6**, 1091–1105 (2000).
105. Kjeldgaard, M., Nissen, P., Thirup, S. & Nyborg, J. The crystal structure of elongation factor EF-Tu from *Thermus aquaticus* in the GTP conformation. *Structure* **1**, 35–50 (1993).
106. Lindorff-Larsen, K. *et al.* Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins Struct. Funct. Bioinforma.* **78**, 1950–1958 (2010).

107. Jorgensen, W. L. Quantum and statistical mechanical studies of liquids. 10. Transferable intermolecular potential functions for water, alcohols, and ethers. Application to liquid water. *J. Am. Chem. Soc.* **103**, 335–340 (1981).
108. Case, D. A. *et al.* The Amber biomolecular simulation programs. *J. Comput. Chem.* **26**, 1668–1688 (2005).
109. Woods, R. J. & Chappelle, R. Restrained electrostatic potential atomic partial charges for condensed-phase simulations of carbohydrates. *J. Mol. Struct. THEOCHEM* **527**, 149–156 (2000).
110. Frisch, M. J. *et al.* Gaussian 16 Revision C.01. (2016).
111. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).
112. Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190 (1981).
113. Evans, D. J. & Holian, B. L. The Nose–Hoover thermostat. *J. Chem. Phys.* **83**, 4069–4074 (1985).
114. Hess, B., Bekker, H., Berendsen, H. J. C. & Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **18**, 1463–1472 (1997).
115. Eisenhaber, F., Lijnzaad, P., Argos, P., Sander, C. & Scharf, M. The double cubic lattice method: Efficient approaches to numerical integration of surface area and volume and to dot surface contouring of molecular assemblies. *J. Comput. Chem.* **16**, 273–284 (1995).
116. Turner, P. XMGRACE, Version 5.1. 19. *Cent. Coast. Land-Margin Res. Or. Grad. Inst. Sci. Technol. Beaverton OR* **2**, (2005).
117. Schrödinger, LLC. The PyMOL Molecular Graphics System, Version 1.8. (2015).
118. Belkin, M. & Niyogi, P. Laplacian Eigenmaps for dimensionality reduction and data representation. *Neural Comput.* **15**, 1373–1396 (2003).
119. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
120. Traag, V. A., Waltman, L. & van Eck, N. J. From Louvain to Leiden: guaranteeing well-connected communities. *Sci. Rep.* **9**, 5233 (2019).
121. Jühling, F. *et al.* tRNADB 2009: compilation of tRNA sequences and tRNA genes. *Nucleic Acids Res.* **37**, D159–162 (2009).
122. Zheng, W. *et al.* Improving deep learning protein monomer and complex structure prediction using DeepMSA2 with huge metagenomics data. *Nat. Methods* **21**, 279–289 (2024).
123. Steinegger, M. & Söding, J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat. Biotechnol.* **35**, 1026–1028 (2017).
124. Hentzen, D., Mandel, P. & Garell, J. P. Relation between aminoacyl-tRNA stability and the fixed amino acid. *Biochim. Biophys. Acta* **281**, 228–232 (1972).
125. Janiak, F. *et al.* Fluorescence characterization of the interaction of various transfer RNA species with elongation factor Tu.GTP: evidence for a new functional role for elongation factor Tu in protein biosynthesis. *Biochemistry* **29**, 4268–4277 (1990).

-
126. LaRiviere, F. J., Wolfson, A. D. & Uhlenbeck, O. C. Uniform Binding of Aminoacyl-tRNAs to Elongation Factor Tu by Thermodynamic Compensation. *Science* **294**, 165–168 (2001).
 127. Sanderson, L. E. & Uhlenbeck, O. C. The 51-63 base pair of tRNA confers specificity for binding by EF-Tu. *RNA N. Y. N* **13**, 835–840 (2007).
 128. Eargle, J., Black, A. A., Sethi, A., Trabuco, L. G. & Luthey-Schulten, Z. Dynamics of Recognition between tRNA and elongation factor Tu. *J. Mol. Biol.* **377**, 1382–1405 (2008).
 129. Lai, J., Ghaemi, Z. & Luthey-Schulten, Z. The Conformational Change in Elongation Factor Tu Involves Separation of Its Domains. *Biochemistry* **56**, 5972–5979 (2017).

10. Supplementary material



*Supplement 1: Comparison between Spectral and Leiden clustering of MI for EF-Tu high temporal resolution (1 million frames) trajectory. **A** Spectral clustering, 4 k-means clusters. **B** Leiden clustering, 15 clusters.*

Supplement 2: Implementation of Leiden clustering (see methods) for MI maps in Python

```
import sys
import json
import numpy as np
import igraph as ig
import leidenalg as la

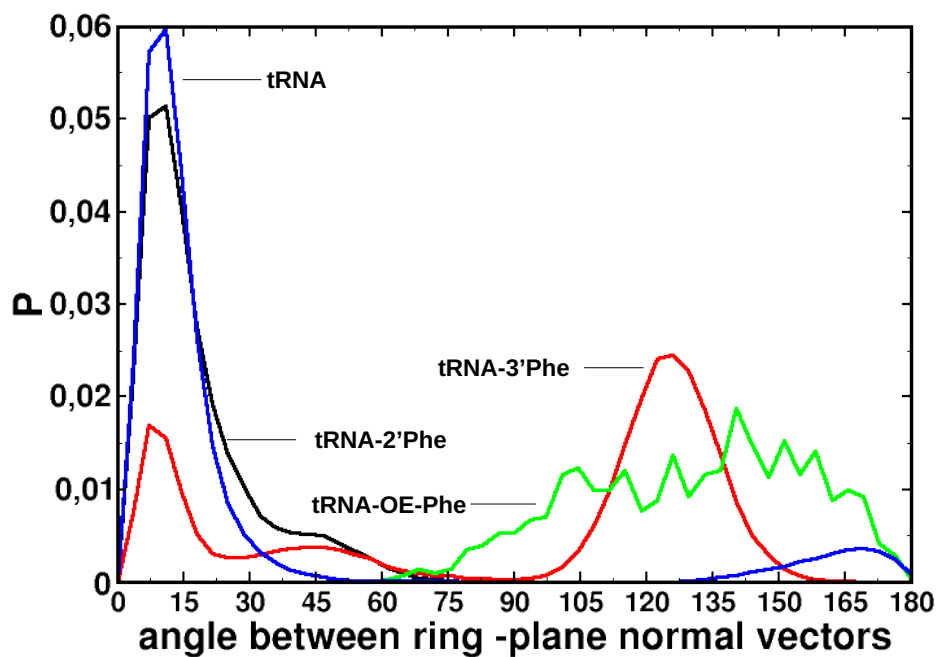
# parse input MI map
inpt = sys.argv[1]
with open(inpt) as json_file:
    data = json.load(json_file)
mmap = np.array(data['map'])

# bound MI values and treat self loop inconsistencies (see formula 7)
i = 0
normalized_arr = []
while i in range(len(mmap)):
    row=[]
    j= 0
    while j in range(len(mmap)):
        row.append(mmap[i][j]/max(mmap[i][i], mmap[j][j]))
        j+=1
    normalized_arr.append(row)
    i+=1
normalized_arr = np.array(normalized_arr)

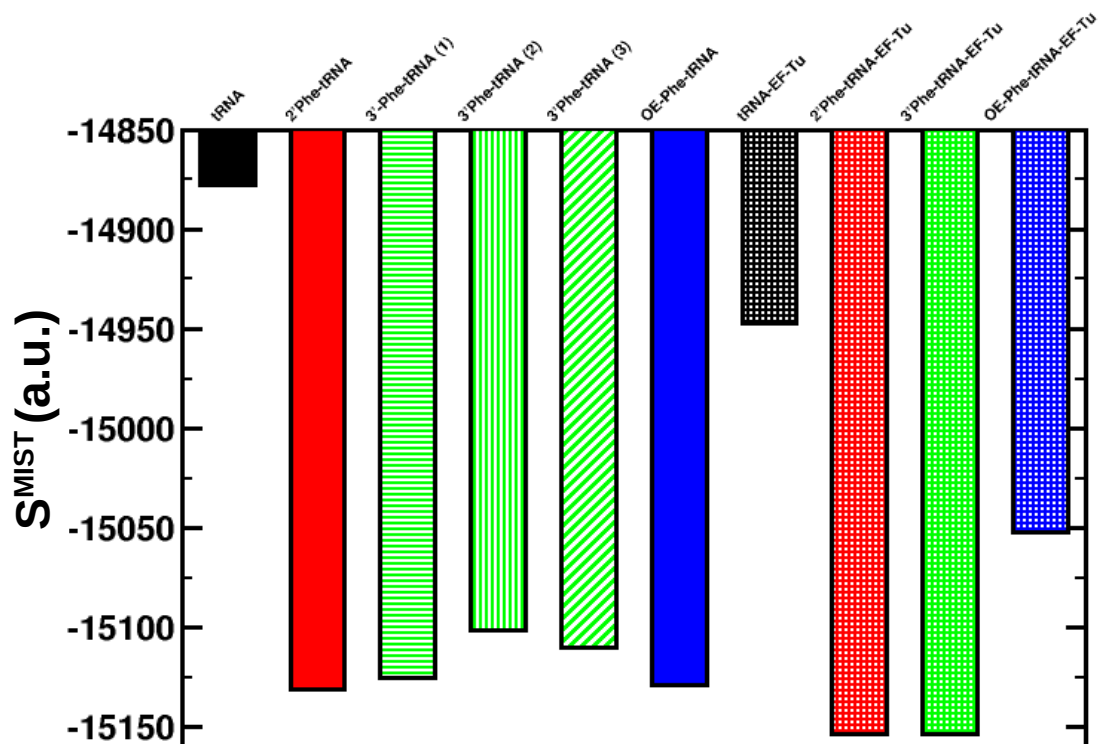
# build undirected weighted graph from similarity matrix
def build_igraph(similarityMatrix):
    edges = (np.unravel_index(np.arange(similarityMatrix.shape[0] *
similarityMatrix.shape[1]), similarityMatrix.shape))
    edges = list(zip(*edges))
    weights = similarityMatrix.ravel()
    g = ig.Graph(directed=False)
    g.add_vertices(similarityMatrix.shape[0])
    g.add_edges(edges)
    g.es['weight'] = weights
    return g

# partition graph with leiden algorithm and strict modularity
def cluster_by_sim_matrix(graph):
    weights = np.array(graph.es["weight"]).astype(np.float64)
    partition_type = la.ModularityVertexPartition
    partition_kwargs = {}
    partition_kwargs["weights"] = weights
    part = la.find_partition(graph, partition_type, **partition_kwargs)
    optimiser = la.Optimiser()
    diff = optimiser.optimise_partition(part, n_iterations=1000)
    return part

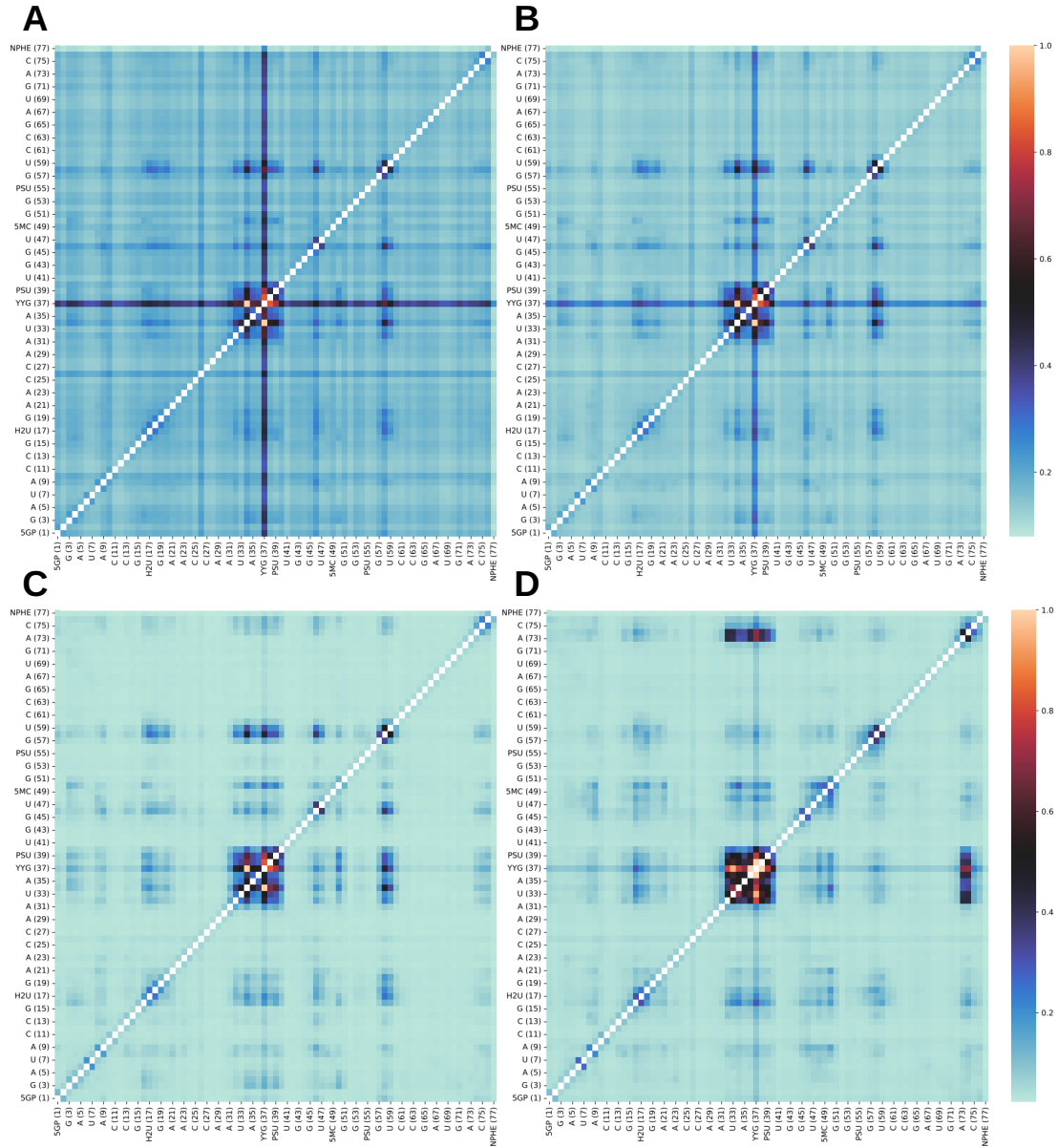
# retrieve partitions
g = build_igraph(normalized_arr)
partition = cluster_by_sim_matrix(g)
labels = np.array(partition.membership)
```



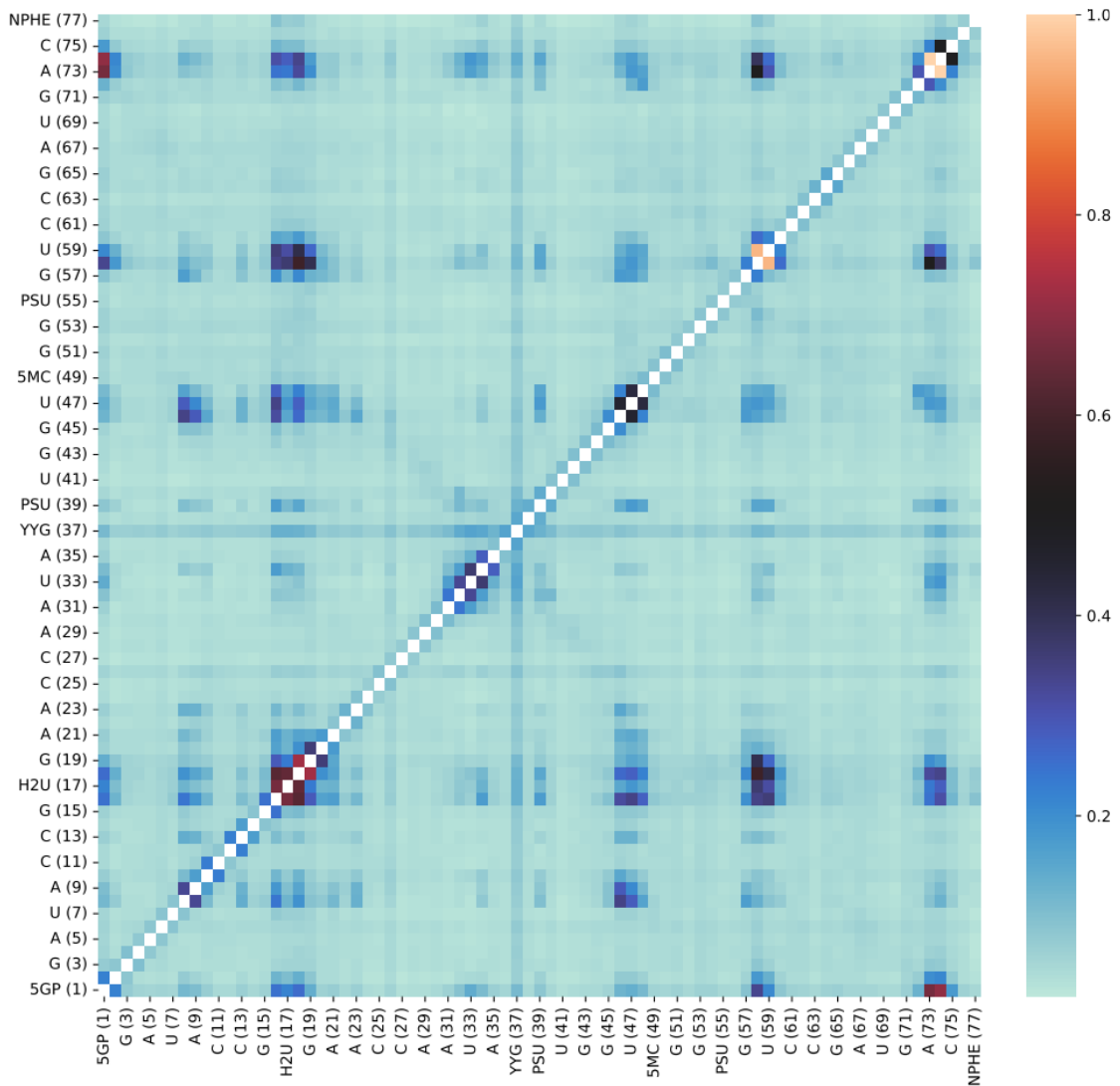
Supplement 3: Distribution of angles between the base ring-plane normal vectors of bases G1 and A76 over 1 μ s MD simulations.



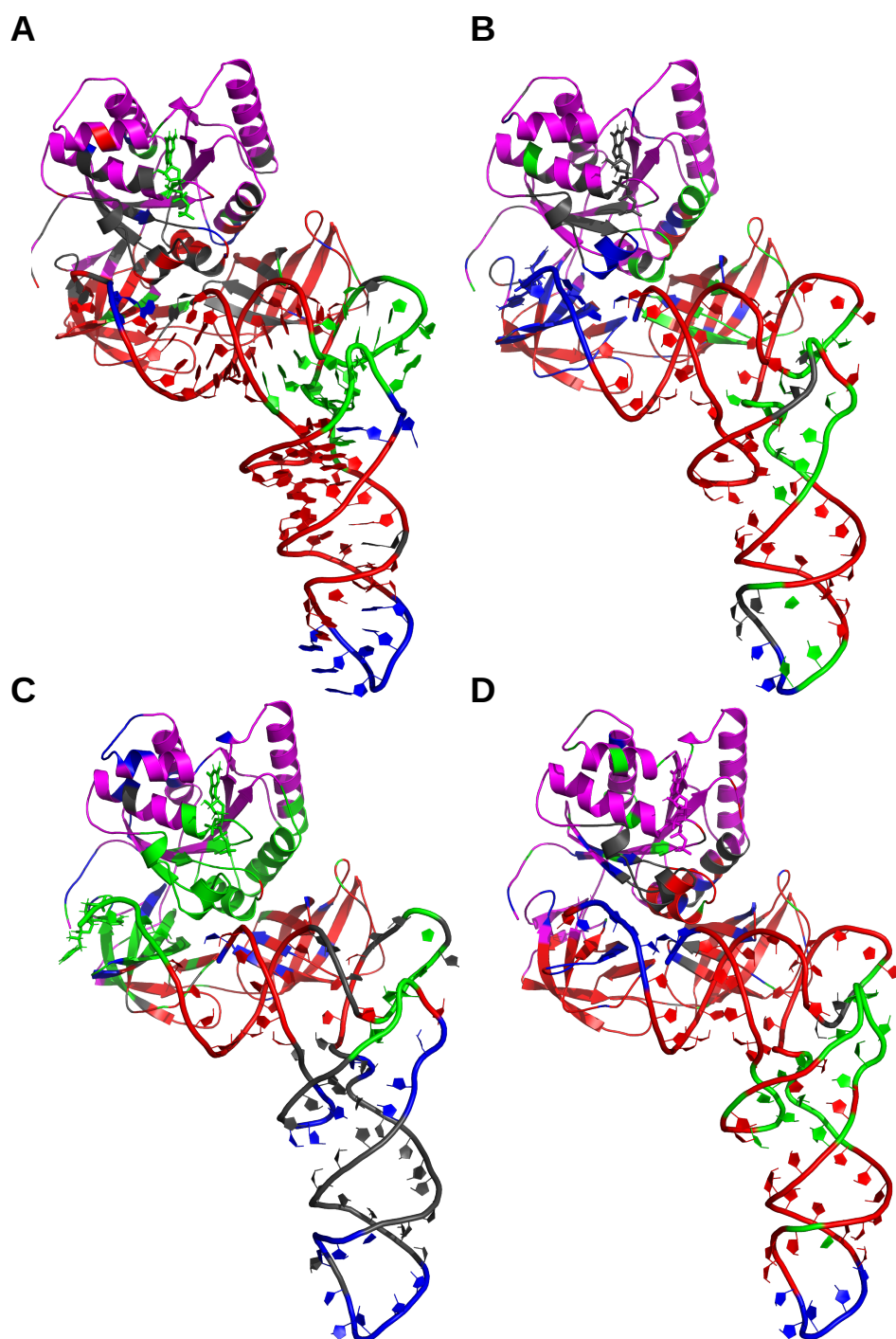
Supplement 4: S^{MIST} of tRNA molecules for all MD trajectories



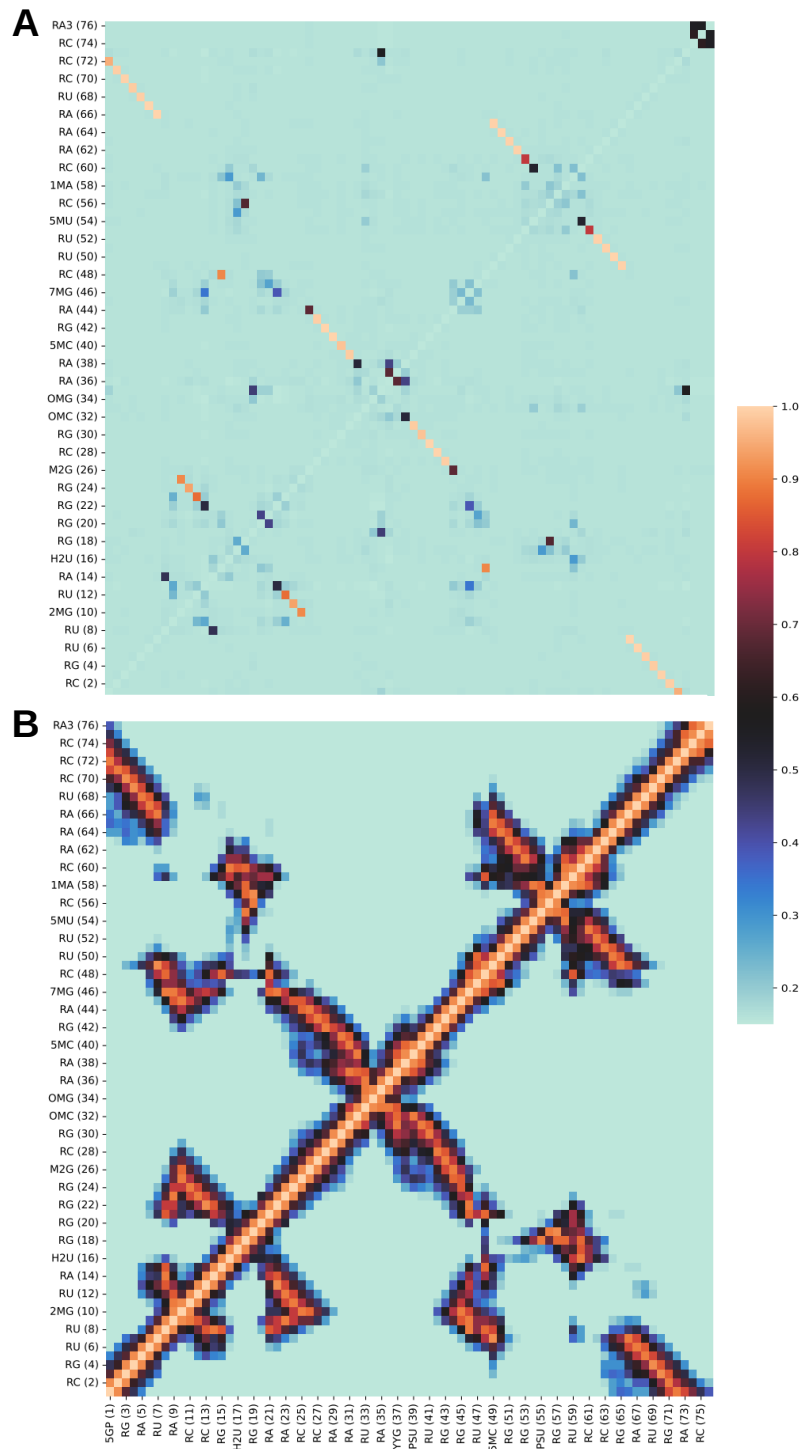
Supplement 5: Visualization of systematic noise in MI maps as a result of sampling sufficiency. A 3'Phe-tRNA^{Phe} trajectory timestep 10ps (1×10^5 frames) artificially decreased sampling. **B** 3'Phe-tRNA^{Phe} trajectory timestep 5ps (2×10^5 frames) native. **C** 3'Phe-tRNA^{Phe} trajectory timestep 5ps (1×10^5 frames) after removal of systematic noise. **D** 3'Phe-tRNA^{Phe} trajectory timestep 1ps (1×10^6 frames) native replica.



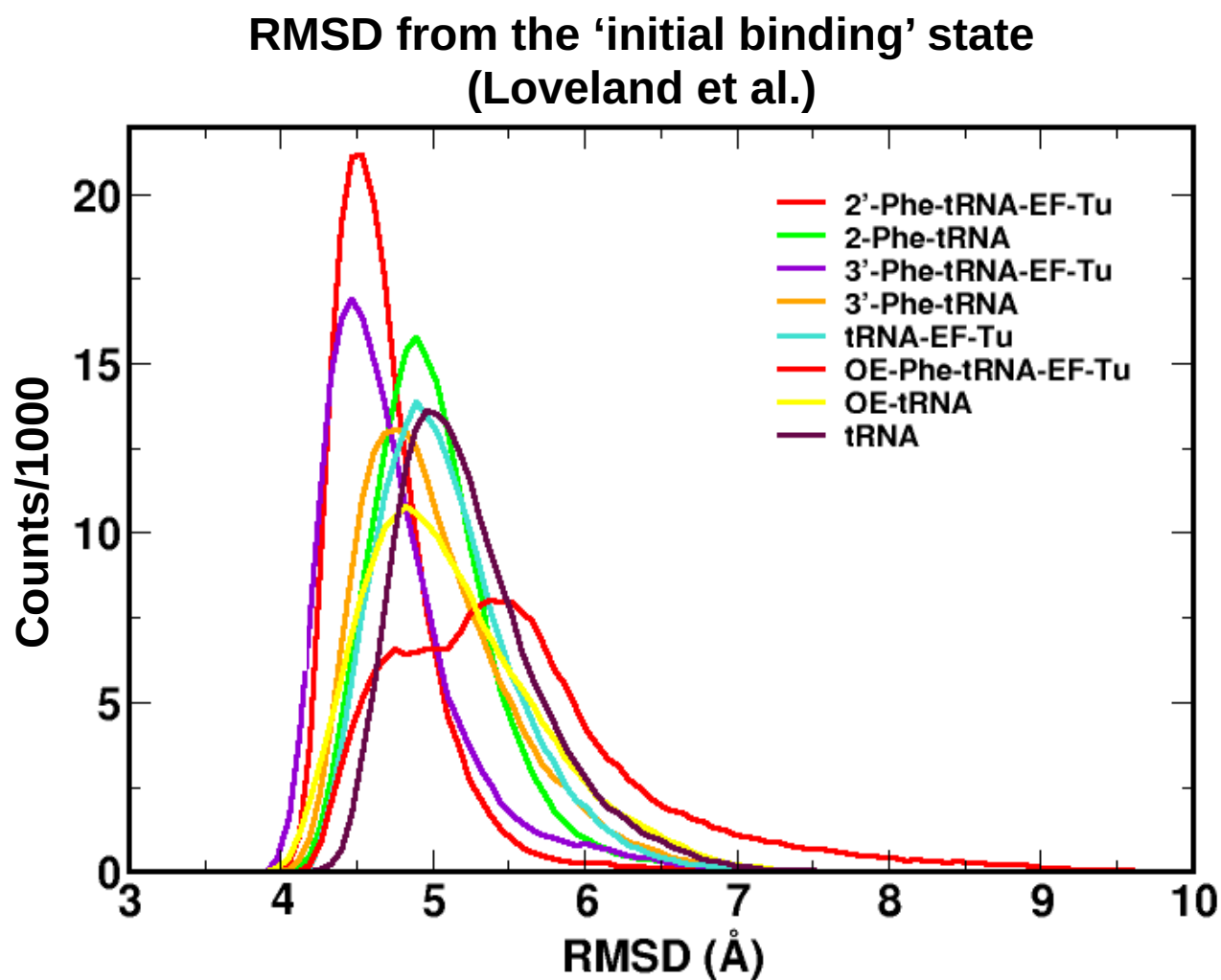
Supplement 6: MI between bases of 3'Phe-tRNA^{Phe} (bound to EF-Tu) rescaled by maximum value.



Supplement 7: Spectral clustering of MI maps for the $tRNA^{Phe} \cdot EF-Tu \cdot GTP$ ternary complex MD trajectories into 5 clusters. **A** $tRNA^{Phe} \cdot EF-Tu$. **B** $2'Phe-tRNA^{Phe} \cdot EF-Tu$. **C** $OE-Phe-tRNA^{Phe} \cdot EF-Tu$. **D** $3'Phe-tRNA^{Phe} \cdot EF-Tu$.



Supplement 8: Comparison of pairwise distance map for 3'Phe-tRNA^{Phe} (PDB code: 1EHZ) to MI between positions calculated from MSA



Supplement 9: Distribution of Backbone atom RMSD of tRNA MD conformers compared to the initial binding state structure from ensemble cryo-EM⁴³.