THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

Sequence-Dependent Base-Stacking Stabilities Guide tRNA Folding Energy Landscapes

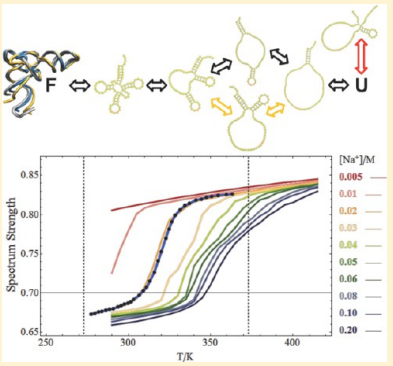
Rongzhong Li,[†] Heming W. Ge,[†] and Samuel S. Cho*[,†,‡]

[†]Department of Physics, Wake Forest University, Winston-Salem, North Carolina 27106, United States

[‡]Department of Computer Science, Wake Forest University, Winston-Salem, North Carolina 27106, United States

Supporting Information

ABSTRACT: The folding of bacterial tRNAs with disparate sequences has been observed to proceed in distinct folding mechanisms despite their structural similarity. To explore the folding landscapes of tRNA, we performed ion concentration-dependent coarse-grained TIS model MD simulations of several *E. coli* tRNAs to compare their thermodynamic melting profiles to the classical absorbance spectra of Crothers and co-workers. To independently validate our findings, we also performed atomistic empirical force field MD simulations of tRNAs, and we compared the base-to-base distances from coarse-grained and atomistic MD simulations to empirical base-stacking free energies. We then projected the free energies to the secondary structural elements of tRNA, and we observe distinct, parallel folding mechanisms whose differences can be inferred on the basis of their sequence-dependent base-stacking stabilities. In some cases, a premature, nonproductive folding intermediate corresponding to the Ψ hairpin loop must backtrack to the unfolded state before proceeding to the folded state. This observation suggests a possible explanation for the fast and slow phases observed in tRNA folding kinetics.



INTRODUCTION

Transfer RNA (tRNA) molecules comprise the most well-studied class of RNA molecules because they are small and of universal importance in the translation of contiguous codons in mRNA into sequences of amino acids that result in a specific protein.[1−4] Each tRNA molecule's anticodon forms Watson−Crick or wobble base pairs with its corresponding mRNA codon and carries the matching amino acid to the ribosome for protein synthesis.[5] Every organism has at least 20 different tRNA molecules, at least one for each amino acid. Despite their sequence diversity, their structures share a conserved "cloverleaf" secondary structure that consists of the D, Anticodon (A), and TΨC (Ψ) hairpin loops and an acceptor (Acc) stem (Figure 1A). With few exceptions, tRNAs adopt an "L"-shaped tertiary structure, even though they have low sequence similarity (Figure 1).[6] Hence, their conformation must evolve under at least two opposing functional constraints: (1) a general, conserved 3D shape that requires it to fit into ribosome binding sites and (2) specific recognition by their cognate aminoacyl-tRNA synthetase and many other potential partners.[7] For specific recognition, the structure alone is unlikely be the primary determinant, which is often the case in protein folding,[8,9] although there are notable exceptions.[10−12]

Crothers and co-workers first established the general features of tRNA folding in their seminal experiments.[13,14] In their thermal denaturation studies, the melting profiles of the various *E. coli* tRNA molecules differed. While their folding was highly dependent on temperature and ionic strength, multiphasic melting curves were observed under different salt concentrations, suggesting intermediate structure(s) in the folding process.[13] Native structure-based[15] and ab initio coarse-grained MD simulations[16] also predict intermediate states in the folding of yeast $tRNA^{Phe}$. Temperature-dependent unfolding occurs through distinct sequence-dependent ensembles, and the melting order of the secondary and tertiary structures was not conserved.[17]

More recent experiments showed that Na^+- and Mg^{2+}-induced kinetic folding of $tRNA^{Phe}$ involves at least two transitions that include secondary structure formation. Furthermore, the exact order of folding depends on hairpin stability in the presence of salts, and the kinetics can involve fast and slow phases, suggesting parallel folding mechanisms.[18−21] Very recently, Perona and co-workers introduced a new method to track tRNA folding kinetics using the activities of its cognate amino-acyl tRNA synthetase as a probe. They also observe fast and slow phases derived from a biexponential fit[22] that may also indicate parallel folding mechanisms.

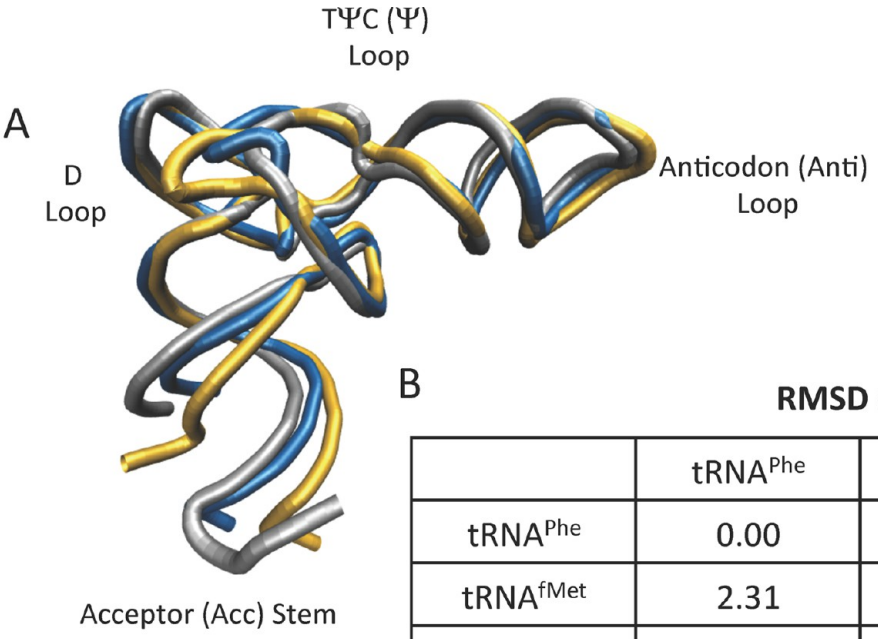
In our present study, we performed ion concentration-dependent coarse-grained MD simulations of several *E. coli* tRNAs, including those whose structures are not presently known, to generate thermodynamic melting profiles that are directly compared to the classical absorbance spectra of Crothers and co-workers. We also performed CHARMM and AMBER empirical force field atomistic MD simulations of tRNA, and we

Special Issue: Peter G. Wolynes Festschrift

Received: March 1, 2013
Revised: July 8, 2013

   dx.doi.org/10.1021/jp402114p | J. Phys. Chem. B XXXX, XXX, XXX−XXX

A



B

RMSD (Å)

	tRNA^{Phe}	tRNA^{fMet}	tRNA^{Cys}
tRNA^{Phe}	0.00	2.31	2.56
tRNA^{fMet}	2.31	0.00	2.69
tRNA^{Cys}	2.56	2.69	0.00

C

tRNA^{Phe} GCCCGGAUAGCUCAGUCGGUAGAGCAGGGGAUUGAAAAUCCCCGUGUCCUUGGUUCGAUUCCGAGUCCGGGCACCA
((((((((((( )))) ((((( ))))) ((((( ))))))))))))

tRNA^{fMet} CGCGGGGUGGAGCAGCCUGGUAGCUCGUCGGGCUCAUAACCCGAAGAUCGUCGGUUCAAAUCCGGCCCCCGCAACCA
((((((((((( ))))((((((( ))))) ((((( ))))))))))))

tRNA^{Tyr} GGUGGGGUUCCCGAGCUGGCAAGGGAGCAGACUGUAAAUCUGCCGUUCGAAGGUUCGAAUCCUUCCCCCACCACCA
(((((((((( ))) ((((( ))))) ((((( ))))))))))))

tRNA^{Cys} GGCGCGUUAACAAAGCGGUUAUGUAGCGGAUUGCAAAUCCGUCUAGUCCGGUUCGACUCCGGAACGCGCCUCCA
(((((((((( ))) ((((( ))))) ((((( ))))))))))))

D

Sequence Similarity (%)

	tRNA^{Phe}	tRNA^{fMet}	tRNA^{Tyr}	tRNA^{Cys}
tRNA^{Phe}	100.0	58.1	56.9	67.2
tRNA^{fMet}	58.1	100.0	58.7	73.1
tRNA^{Tyr}	56.9	58.7	100.00	58.3
tRNA^{Cys}	67.2	73.1	58.3	100.00

Figure 1. Structural and sequence characterization of various tRNA molecules. (A) The structures of yeast tRNA^{Phe} (PDB: 1EHZ; blue), bacterial tRNA^{fMet} (PDB: 2FMT; yellow), and bacterial tRNA^{Cys} (PDB: 1U0B; gray) are superimposed. (B) The differences between the three structures are measured by the root-mean-square deviation (RMSD). (C) The primary sequences (top row) and the corresponding secondary structure base-pairing (bottom row) of bacterial tRNA used in the study. (D) The percent similarity between the tRNA sequences as measured by LALIGN.

compared them to our coarse-grained MD simulations and empirical base-stacking free energies. With coarse-grained MD simulations, we observe sequence-dependent tRNA folding mechanisms and our simulations suggest a possible contributor of the fast and slow phases observed in folding.

METHODS

TIS Model Native Structure Based Coarse-Grained Hamiltonian. To perform folding MD simulations of various tRNA molecules and determine their folding free energy landscapes, we used a coarse-grained model called the three-

dx.doi.org/10.1021/jp402114p | J. Phys. Chem. B XXXX, XXX, XXX−XXX

interaction-site (TIS) model for RNA. The detailed energy function and parameters for the TIS model can be found elsewhere,[23−26] but we briefly discuss it here. In this simplified representation of RNA, each nucleotide in the model consists of three beads for the nitrogenous base, ribose sugar, and negatively charged phosphate moieties. The TIS model consists of short-range interactions for the bond (H_{bond}), angle (H_{angle}), and torsional potentials ($H_{torsion}$). In addition to the connectivity and rotational degrees of freedom, it also consists of long-range interactions for stabilizing the native structure (H_{NC}), the electrostatic interactions for the charged phosphates (H_{elec}), and aromatic base-stacking interactions ($H_{stacking}$) based on the well-known Turner rules.[27]

$$H_{total} = H_{short} + H_{long}$$

$$H_{short} = H_{bond} + H_{angle} + H_{torsion}$$

$$H_{long} = H_{NC} + H_{elec} + H_{stacking}$$

The electrostatic interactions between two phosphate beads are represented by the Debye−Hückel potential:

$$V_{P_iP_j} = \frac{z_{P_i} z_{P_j} e^2}{4\pi\epsilon_0\epsilon_r r} e^{-r/l_D}$$

The Debye length, l_D, is tuned to reflect the changes in ion concentration. In an earlier study, the melting temperature obtained by the TIS model was found to be in excellent agreement with previously experimentally measured absorbance spectra.[24]

TIS Model Simulations of tRNA without Structures. In the TIS model, the long-range interactions consist of the native structure-based contact interactions, the sequence length-dependent electrostatic interactions that depend on the number of phosphates present, and sequence-dependent base-stacking interactions. Although Crothers and co-workers studied *E. coli* $tRNA^{Phe}$, $tRNA^{fMet}$, and $tRNA^{Tyr}$, only the structure of $tRNA^{fMet}$ from *E. coli* is known (PDB: 2FMT), and structures of $tRNA^{Phe}$ and $tRNA^{Tyr}$ from *E. coli* do not exist, although yeast $tRNA^{Phe}$ has been resolved (PDB: 1EHZ). A computational challenge is to integrate the existing plethora of sequence information with limited structural data to gain biophysical data.[28] Toward that end, we generated a TIS model Hamiltonian based on the structure of yeast $tRNA^{Phe}$ and the sequence of *E. coli* $tRNA^{Phe}$ based on the observation that the tertiary structures and the sequence length of tRNAs that Crothers and co-workers studied are very similar (Figure 1A,B) and their sequence similarity is low (Figure 1C,D). Similarly, we generated a TIS model Hamiltonian of *E. coli* $tRNA^{Tyr}$ based on the structure of yeast $tRNA^{Phe}$ and the sequence of *E. coli* $tRNA^{Tyr}$ because of sequence similarities. Also, we generated the TIS model Hamiltonian of *E. coli* $tRNA^{Cys}$ (PDB: 1U0B) even though to our knowledge an absorbance melting profile has not been performed but its structure is already known.

We performed Langevin MD simulations with a low friction coefficient for more effective sampling of the free energy landscape.[29] For each of the four tRNAs we studied, we performed three sets of independent MD simulations over a range of ion concentrations $[Na^+]$ = 0.005−0.200 M (uncalibrated), with temperatures ranging from 250 to 375 K. As previously noted, the temperature was offset by 35 K.[23]

To measure the conformational change of the molecule for comparison to the experimentally observed melting process, we adopted the fraction of native contacts Q as an order parameter.[30] The experimental spectrum strength cannot be explicitly derived using physical quantities, so we fitted the experimental spectrum strength to the ensemble-averaged Q using $1/(Q+\alpha)$, where α is a fitting parameter for the asymptotic cooperative behavior. The melting temperatures derived from our simulations are not fitted, and they can be directly compared with experiments.

Atomistic CHARMM and AMBER Empirical Force Field MD Simulations. We performed atomistic empirical force field MD simulations of *E. coli* $tRNA^{fMet}$, $tRNA^{Phe}$, $tRNA^{Tyr}$, and $tRNA^{Cys}$ using the NAMD software suite with the CHARMM27 and AMBER S99b0 force fields. For the tRNAs whose sequences are known but their structures are unresolved, we can again take advantage of remarkable similarity of tRNA structures (Figure 1A,B) by threading their sequences to the yeast $tRNA^{Phe}$, whose structure has been resolved in the absence of a binding partner. We altered the sequence of the atomistic yeast $tRNA^{Phe}$ structure into the *E. coli* $tRNA^{Phe}$ and $tRNA^{Tyr}$ by replacing its bases with other bases with minimal changes based on ideal structures using the MutateNA.pl program from the MMTSB toolbox.[31]

For each tRNA, we generated five sets of 30 ns trajectories (150 ns total). To set up these simulations, each tRNA molecule was centered in a cubic solvent box based on the TIP3P water model such that no water molecule was within 1.6 Å of a tRNA heavy atom. The dimensions of the solvent box were such that at least a 10 Å solvent buffer exists in each of the *x*-, *y*-, and *z*-dimensions. Within this box, we added sodium and chloride ions so that the total ion concentration was 0.005, 0.005, 0.02, and 0.02 M for $tRNA^{fMet}$, $tRNA^{Phe}$, $tRNA^{Tyr}$, and $tRNA^{Cys}$, respectively. These concentrations correspond to the experimentally measured absorbance spectra of Crothers and co-workers at low ion concentrations. Our simulations were performed using a standard protocol using periodic boundary conditions and treated the electrostatics with the particle mesh Ewald method.

For each trajectory of the tRNAs, the system was equilibrated by first performing 2000 steps of energy minimization and then raising the simulated temperatures for 50 ns until the target temperature was reached. For $tRNA^{fMet}$, $tRNA^{Phe}$, $tRNA^{Tyr}$, and $tRNA^{Cys}$, the target temperatures were 313, 313, 323, and 323 K, respectively, which corresponded to the melting temperature observed by Crothers and co-workers.[14] Then, we performed a production run MD simulation at the target temperature for 30 ns for each of the five trajectories for analysis.

RESULTS AND DISCUSSION

To determine how tRNA molecules with similar structures can give rise to different melting profiles as observed by Crothers and co-workers, we performed TIS model coarse-grained MD simulations of *E. coli* $tRNA^{fMet}$, $tRNA^{Phe}$, $tRNA^{Tyr}$, and $tRNA^{Cys}$, which all have very similar structures and sequence lengths (Figure 1A,B) but different sequence compositions (Figure 1C,D). The long-range interactions of the TIS model Hamiltonian consist of native structure contact interactions, sequence length-dependent electrostatic interactions for the phosphate backbone, and sequence-dependent base-stacking interactions.

However, there are only a few tRNA molecules whose X-ray crystallographic or NMR structures have been determined in isolation. We compared the yeast $tRNA^{Phe}$, which was recently determined and refined using modern techniques, with *E. coli* $tRNA^{fMet}$ and $tRNA^{Cys}$, whose structures are known when bound to their cognate aminoacyl-tRNA synthetase (AARS). Their

dx.doi.org/10.1021/jp402114p | J. Phys. Chem. B XXXX, XXX, XXX−XXX

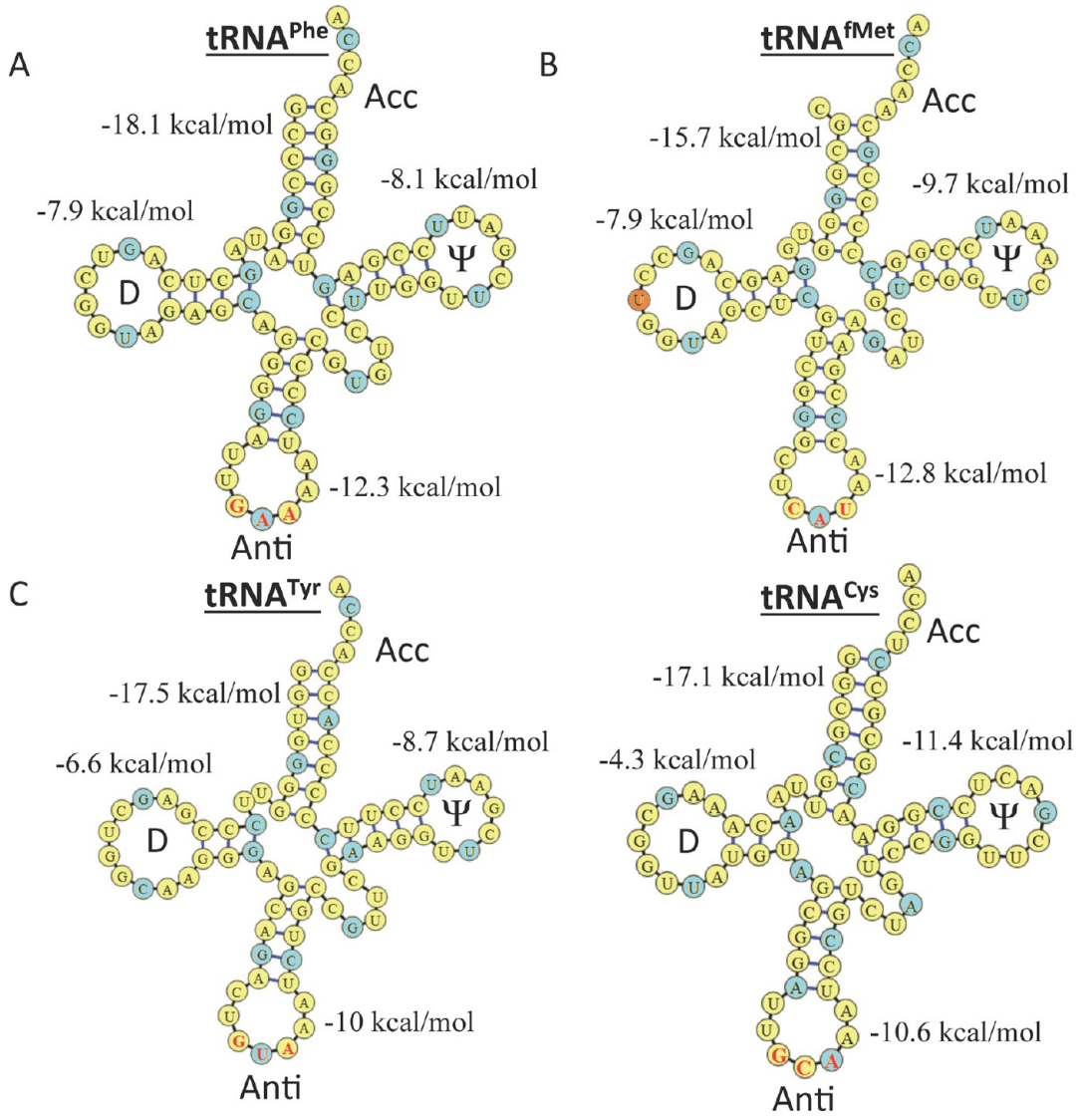


Figure 2. Secondary structure and base-stacking energies of bacterial tRNA molecules. For the bacterial (A) $tRNA^{Phe}$, (B) $tRNA^{fMet}$, (C) $tRNA^{Tyr}$, and (D) $tRNA^{Cys}$, the secondary structures are shown with the canonical Watson−Crick base pairs shown with a black line and non-canonical base pairs shown with a red line. Also, the base-stacking free energies are shown for the D, Anti, and ψ hairpin loops, as well as the acceptor stem. Every fifth nucleotide is colored blue according to traditional tRNA numbering except for $tRNA^{fMet}$, which has an "extra" nucleotide according to that numbering system.

backbone RMSDs are less than 3.0 Å between their structures (Figure 1A,B). We next compared the sequences of *E. coli* $tRNA^{fMet}$, $tRNA^{Phe}$, $tRNA^{Tyr}$, and $tRNA^{Cys}$, and the sequence percent similarity ranged from 56.9 to 73.1% between them (Figure 1C,D). Their secondary structural configurations of base-pair interactions, however, all consist of the familiar cloverleaf configuration of base pairs (Figure 2).

To construct the TIS model Hamiltonians for the various tRNA molecules we studied, we threaded the sequence-dependent base-stacking interactions of *E. coli* $tRNA^{Phe}$ and $tRNA^{Tyr}$ to the TIS model Hamiltonian native structure and electrostatic terms of yeast $tRNA^{Phe}$. Specifically, we determined the $H_{stacking}$ based on the sequence, while all other terms were based on the yeast $tRNA^{Phe}$. For $tRNA^{fMet}$ and $tRNA^{Cys}$, sequence-dependent base-stacking interaction threading was unnecessary because their full structures (bound to AARS) were known.

Empirical Base-Stacking Interaction Trends Are Comparable to Base−Base Distances from Coarse-Grained and Atomistic MD Simulations. To complement the experiments and coarse-grained TIS model simulations, we performed five sets of 30 ns empirical force field MD simulations of the *E. coli* $tRNA^{fMet}$, $tRNA^{Phe}$, $tRNA^{Tyr}$, and $tRNA^{Cys}$ using the CHARMM27[32] and AMBER S99b0[33] force fields with the NAMD simulation package.[34] While this time scale is not sufficient to observe folding events, we can make direct comparisons to the coarse-grained MD simulations by analyzing equivalent trajectories. We therefore computed the average distances between the base pair centers of mass for the CHARMM and AMBER MD simulations that we compared to the base-to-base bead distances from the TIS model simulations that were also 100 ns in length (Figure 3). We also compared the base-to-base distances to the base-stacking stabilization free energies from Turner's rules,[27] which define the TIS model Hamiltonian base-stacking terms (Figure 3).

Coarse-Grained MD Simulations Quantitatively Reproduce Classical tRNA Thermal Melting Profiles at Low Ion Concentrations. For each of the tRNAs we studied, we

dx.doi.org/10.1021/jp402114p | J. Phys. Chem. B XXXX, XXX, XXX−XXX

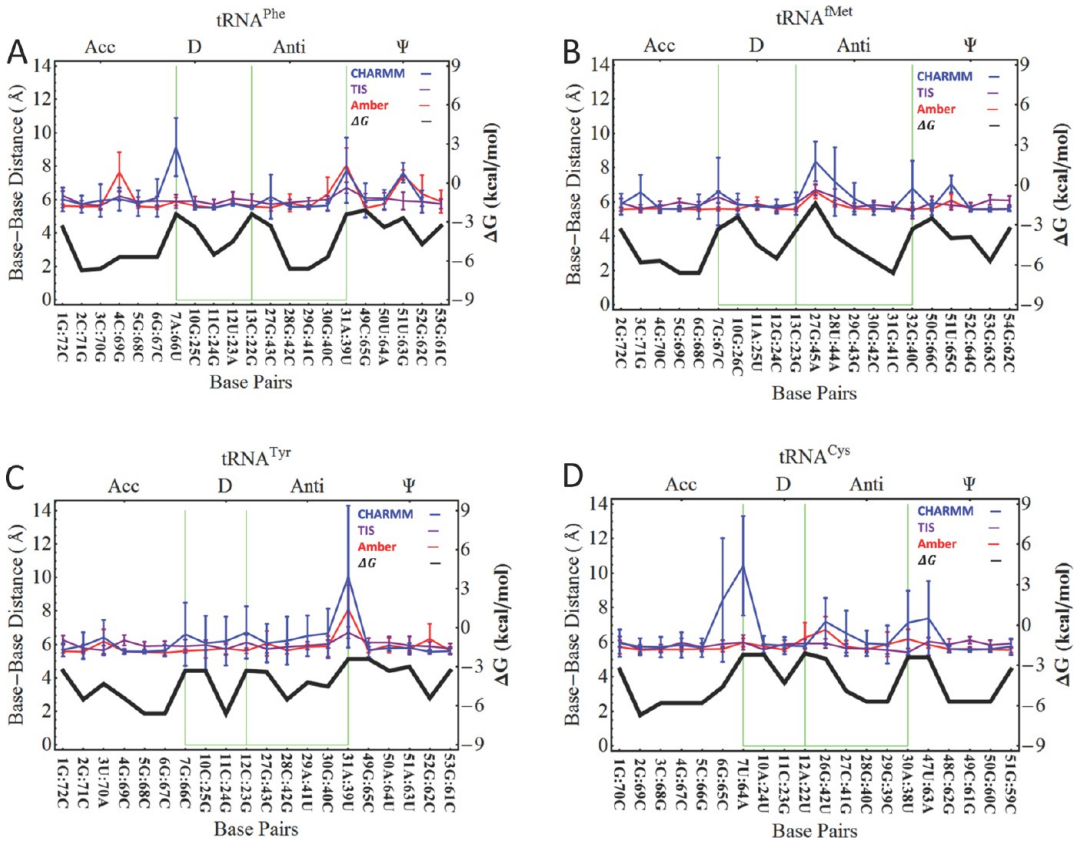


Figure 3. Ensemble-averaged base-to-base distances of base pairs in (A) $tRNA^{Phe}$, (B) $tRNA^{fMet}$, (C) $tRNA^{Tyr}$, and (D) $tRNA^{Cys}$ from CHARMM, TIS model, and AMBER MD simulations compared with base-stacking free energies from MFold. The standard deviations are shown as error bars.

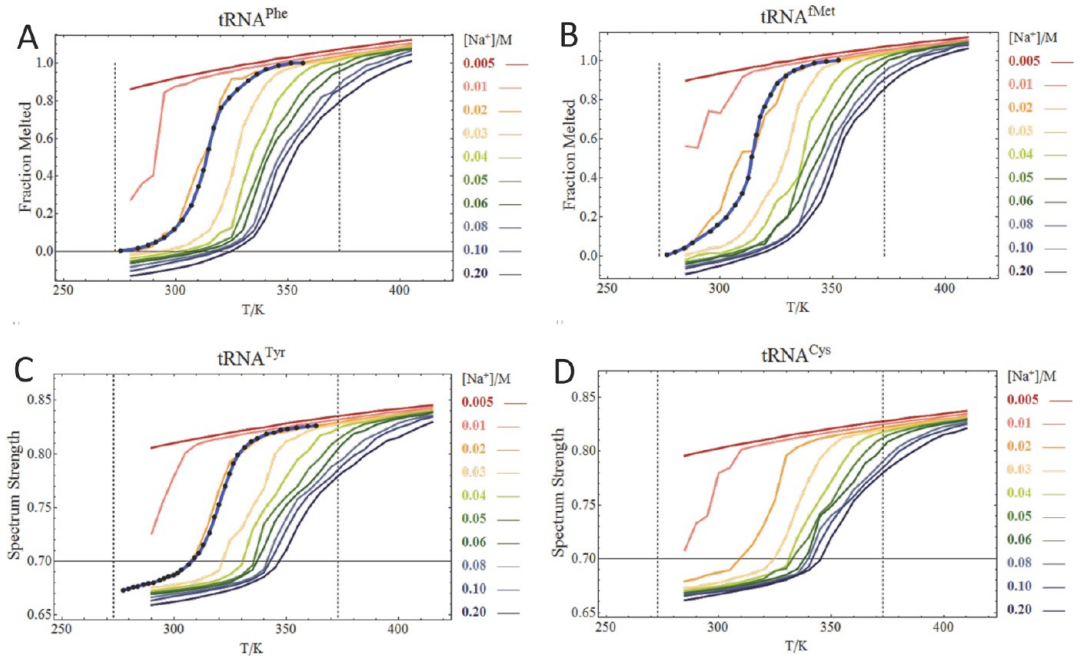


Figure 4. Ion concentration-dependent melting profiles from simulations compared with absorbance spectra measured by Crothers and co-workers. For the bacterial (A) $tRNA^{Phe}$, (B) $tRNA^{fMet}$, (C) $tRNA^{Tyr}$, and (D) $tRNA^{Cys}$, the fraction of native contacts is probed as a function of the temperature over a broad ion concentration range ($[Na^+]$ = 0.005−0.20 M). The dotted black vertical lines indicate the freezing and boiling points of water. For comparison, the absorbance spectra experimentally measured by Crothers and co-workers are shown for $tRNA^{Phe}$, $tRNA^{fMet}$, and $tRNA^{Tyr}$ in a dark blue dotted line.

performed multiple independent TIS model coarse-grained MD simulations over a broad ion concentration and temperature range. For thermodynamic analyses, we observed multiple folding and unfolding events in the MD simulations. For

dx.doi.org/10.1021/jp402114p | J. Phys. Chem. B XXXX, XXX, XXX−XXX

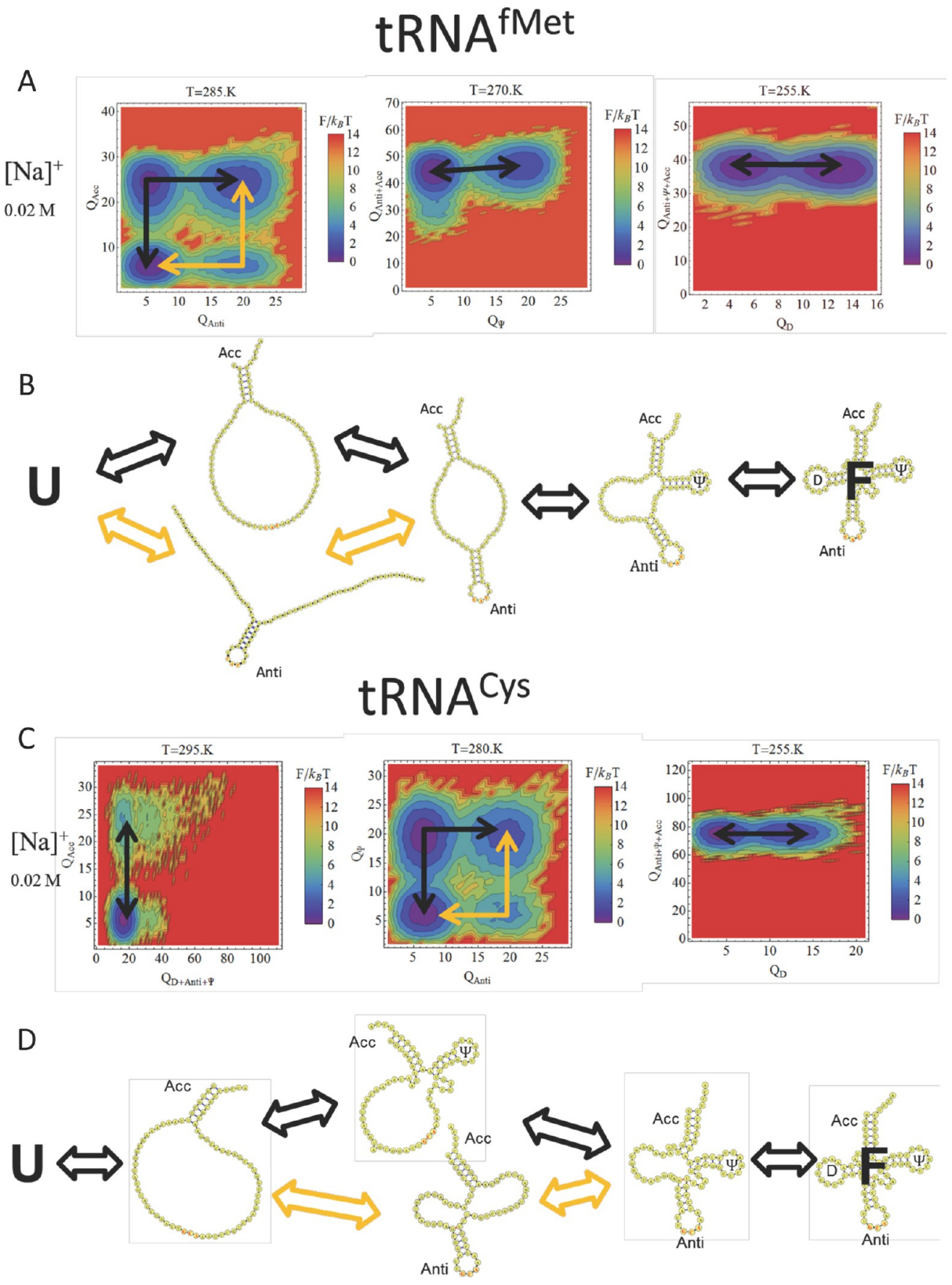


Figure 5. Free energy folding landscapes of $tRNA^{fMet}$ and $tRNA^{Cys}$. (A, C) The free energy is projected with respect to Q of the different hairpin loops and acceptor stem of $tRNA^{fMet}$ and $tRNA^{Cys}$ at different melting temperatures. (B, D) Schematics of the corresponding folding mechanisms inferred from the free energy profiles.

$tRNA^{fMet}$, $tRNA^{Phe}$, and $tRNA^{Tyr}$, we calculated the ensemble-averaged fraction of native contacts, $\langle Q \rangle$, as a probe of the folding and directly compared them to the absorbance spectra from the classical tRNA melting profiles of Crothers and co-workers. We observe similar melting profiles and melting temperatures at $[Na^+] = 0.020$ M, the conditions of the experiments at which the

dx.doi.org/10.1021/jp402114p | J. Phys. Chem. B XXXX, XXX, XXX−XXX

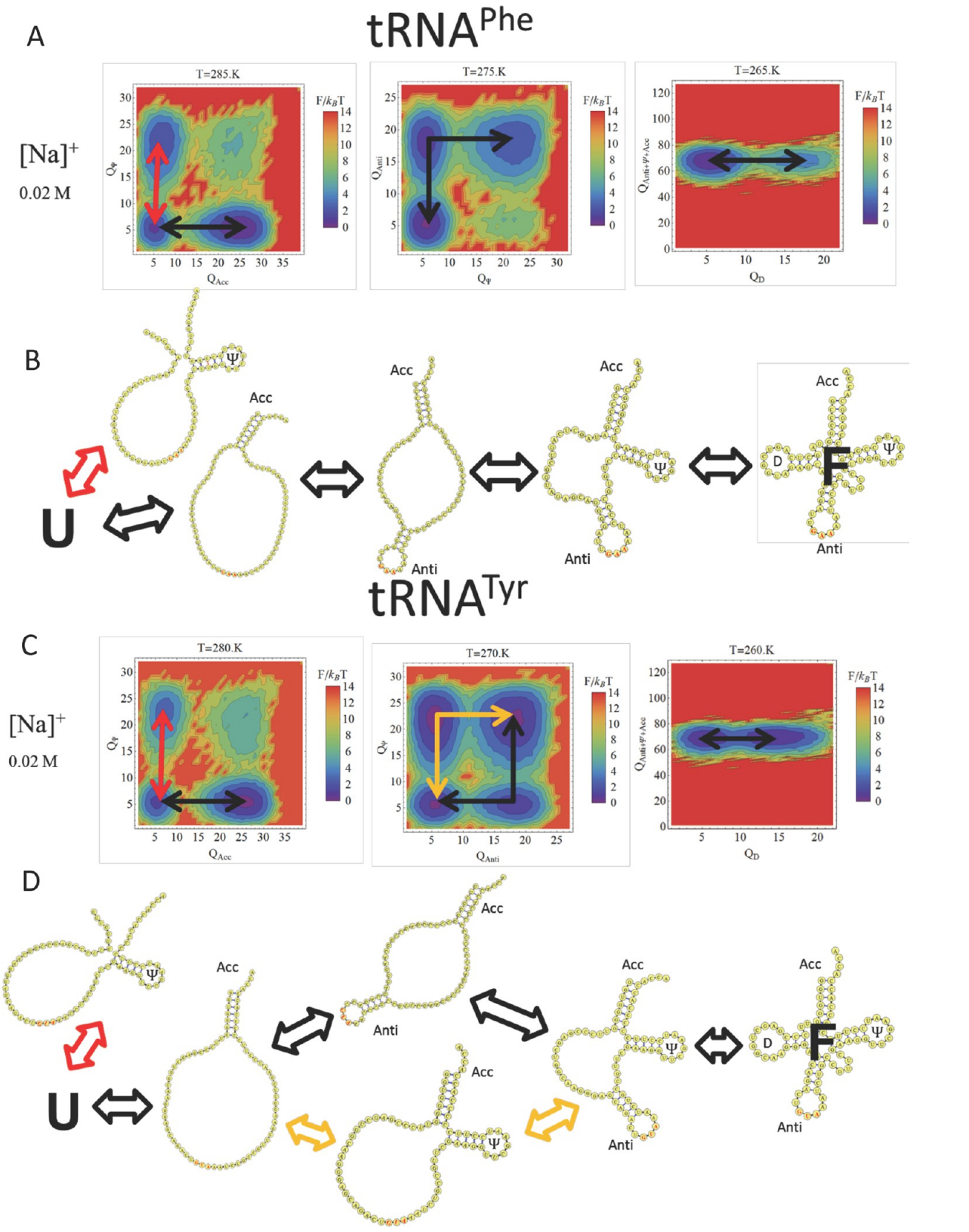


Figure 6. Free energy folding landscapes of $tRNA^{Phe}$ and $tRNA^{Tyr}$ with backtracking. (A, C) The free energy is projected with respect to Q of the different hairpin loops and acceptor stem of $tRNA^{Phe}$ and $tRNA^{Tyr}$ at different melting temperatures. (B, D) Schematics of the corresponding folding mechanisms inferred from the free energy profiles. The backtracking mechanisms are highlighted with red arrows.

secondary structure is expected to dominate (Figure 4A–C). The difference between the melting temperatures from experiments and our simulations is about 5 K. For $[Na^+] < 0.020$ M, the tRNAs remain significantly unfolded, even below the freezing point of water. This ion concentration regime seems to represent a significant limitation of the Debye–Hückel electrostatic

dx.doi.org/10.1021/jp402114p | J. Phys. Chem. B XXXX, XXX, XXX–XXX

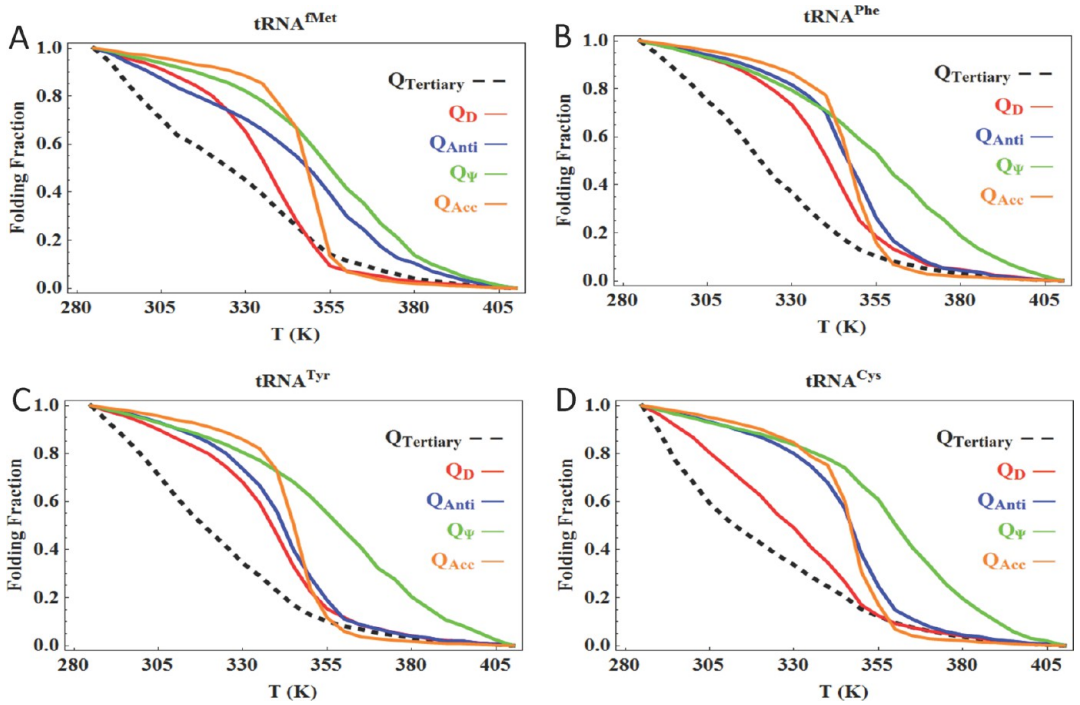


Figure 7. Secondary and tertiary structure formation at high ion concentration. For the bacterial (A) $tRNA^{Phe}$, (B) $tRNA^{fMet}$, (C) $tRNA^{Tyr}$, and (D) $tRNA^{Cys}$, the fraction of native contacts of each secondary structural element and tertiary structure is probed as a function of the temperature at $[Na^+]$ = 0.200 M.

potential. We performed the same analysis for $tRNA^{Cys}$ (Figure 4D), although the melting profiles have not yet been experimentally measured. We predict that the melting temperature of $tRNA^{Cys}$ is significantly higher than $tRNA^{fMet}$, which can be readily tested experimentally.

tRNA Folding Free Energy Landscapes. To obtain a microscopic view of the tRNA folding mechanisms, we projected the free energy profiles with respect to the tRNA secondary structural elements at their respective melting temperatures using the fraction of native contacts (Q) as a reaction coordinate, which has been demonstrated to accurately probe folding mechanisms.[30]

We calculated the folding free energy landscapes for $tRNA^{fMet}$ and $tRNA^{Cys}$. These two systems were chosen because, although the D hairpin loop base-stacking free energies are comparably low, the Anti hairpin loop is more stable than the Ψ hairpin loops with a $\Delta\Delta G \sim 3$ kcal/mol between the two loops in the $tRNA^{fMet}$, but their $\Delta\Delta G \sim 0.8$ kcal/mol is in the $tRNA^{Cys}$. The Acceptor stem is more stable in the $tRNA^{Cys}$ than in the $tRNA^{fMet}$ (Figure 2B,D).

In the free energy profiles projected to each of the constituent secondary structural elements, the $tRNA^{fMet}$ folding involves competitive and parallel folding of the Acc stem and Anti hairpin loop at the highest temperature (285 K). Regardless of the folding mechanism at the higher temperature, the Ψ hairpin loop folds at an intermediate temperature (270 K), and the D hairpin loop folds at the lowest temperature, completing the folding mechanism (Figure 5A,B). Note that the base-stacking stability of the Acc stem ($\Delta G = -15.7$ kcal/mol) and Anti hairpins ($\Delta G = -12.8$ kcal/mol), which have a difference of 2.9 kcal/mol (Figure 2B), fold at the same temperature, and the Acc stem folding mechanism involves lower free energy barriers. For the $tRNA^{Cys}$, the Acc stem (base-stacking stability of −17.1 kcal/mol) folding occurs at the highest temperature (295 K), and there exists a parallel folding of the Anti and Ψ hairpin loops with very similar base-stacking stabilities ($\Delta G = -10.6$ and −11.4 kcal/mol, respectively) at the intermediate temperature (280 K), and the least stable ($\Delta G = -4.3$ kcal/mol) D hairpin loop again folds at the lowest temperature to complete the folding mechanism (Figure 5C,D), again in agreement with the order of the stabilities of the constituent secondary structural elements.

Backtracking Mechanism Partitions Fast vs Slow Phase Folding. The natural next step is to calculate free energy profiles with respect to each of the individual components of the tRNA secondary structure elements to determine the folding (or melting) mechanism of *E. coli* $tRNA^{Phe}$ and $tRNA^{Tyr}$. For both cases, we also observe that different folding mechanisms are dependent on their base-stacking stabilities (Figure 6).

However, the Ψ hairpin loop of $tRNA^{Phe}$ and $tRNA^{Tyr}$ folds at the highest melting temperature, while just the Acc stem would be expected to fold on the basis of their relative base-stacking stabilities. In both cases, the premature folding of the Ψ hairpin loop was not productive, and it must first be unfolded for the folding mechanism to proceed (Figure 6). The dissolution of prematurely formed native contacts has been ascribed to topological frustration in protein folding[9,35−37] for funneled energy landscapes,[38−41] and similar mechanisms are seen here. Although the 2D free energy profiles cannot conclusively preclude that backtracking was occurring in our simulations, the 3D free energy profiles were necessary to confirm that the folding of the Ψ hairpin loop did not proceed further to the folded state (Figure S2, Supporting Information). These observations in our simulations are not readily predicted on the basis of the base-stacking stabilities of the individual secondary structural elements.

Tertiary Structure Forms after Secondary Structure Is Complete at High Ion Concentration. We next compared the secondary and tertiary structure formation. In RNA folding, it

dx.doi.org/10.1021/jp402114p | J. Phys. Chem. B XXXX, XXX, XXX−XXX

is generally observed that secondary structure forms before tertiary structure due to the energetic strong hydrogen bonding and base-stacking interactions that stabilize the secondary structure. The stabilities of the individual helices can largely determine the order of the folding of assembly of pseudoknots and riboswitches.[24−26,42] In addition, Brooks, Al-Hashimi, and co-workers have recently shown that the connectivity and sterics that define the topological constraints of interhelical orientations across bulges in RNA can limit the possible tertiary structure it can adopt.[43] Other recent examples have shown that, due to the low stability gap separating the native state from other non-native states,[18,44] secondary structural rearrangements can occur to accommodate the tertiary structure.[45−47]

In the work by Crothers and co-workers, their experiments were done at low ion concentrations ($[Na^+]$ = 0.020 M) where tertiary structure would not be expected to form due to strong charge−charge repulsions between the phosphate groups. They also performed their experiments at high ion concentrations ($[Na^+]$ = 0.174−0.200 M). Their melting profiles at high ion concentrations are very distinct from the ones we observe in our simulations. In their experimentally measured melting profiles, they observe at least two transitions that correspond to the secondary and tertiary structures. Our single transition melting profiles would not be expected to reproduce the experimentally measured melting profiles even with our fitting procedure, and it is clearly a limitation of our coarse-grained model.

Instead, we computed the ensemble-averaged fraction of native tertiary contacts, $\langle Q_{Tertiary} \rangle$, that we define to be any native contact that is not a secondary structure contact (see Figure S1 in the Supporting Information for exact contacts), to compare with the ensemble-averaged fraction of the native secondary contacts. In each case, at high ion concentration ($[Na^+]$ = 0.200 M), we observe that the native tertiary contact formation is complete only after the native secondary contacts are largely complete (Figure 7). At low ion concentrations, as expected, the tertiary native contacts can occur with secondary contacts (Figure S3, Supporting Information).

CONCLUSIONS

The folding mechanism of tRNA molecules can be different, even for cases where their structures and length are very similar. In our present study, we performed a combination of coarse-grained TIS model and atomistic empirical force field MD simulations to dissect the folding mechanisms of various *E. coli* tRNA whose absorbance spectra have been measured experimentally by Crothers and co-workers. Our simulations accurately reproduced their melting profiles when we matched the ion concentration in our simulations at low ion concentrations where secondary structure dominates. At higher ion concentrations, the tertiary structure formation occurs in our simulations after secondary structure is complete, which need not be the case, as recent experimental and computational studies have shown.[45−47]

When we calculated the free energy landscapes of the different tRNAs, we observed parallel and competitive folding mechanisms that we isolate to the sequence-dependent base-stacking stabilities, which determine the melting temperatures of the individual secondary structural elements. In some cases, we observed a premature, nonproductive intermediate corresponding to the formation of the Ψ hairpin loop, which must unfold before proceeding to the folded state. This process may ultimately slow down the folding kinetics, and this may be one possible source of the observation of biphasic fast and slow folding kinetics for tRNAs.

ASSOCIATED CONTENT

Supporting Information

Native contact maps of $tRNA^{Phe}$, $tRNA^{fMet}$, $tRNA^{Tyr}$, and $tRNA^{Cys}$. Three-dimensional free energy folding landscapes of $tRNA^{Phe}$ and $tRNA^{Tyr}$ projected to the fraction of native contacts, Q, with respect to the Acc stem and Anti and Ψ hairpin loops to identify backtracking mechanisms. Melting profiles of $tRNA^{Phe}$, $tRNA^{fMet}$, $tRNA^{Tyr}$, and $tRNA^{Cys}$ with respect to each of the secondary structural elements and the tertiary structure. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: choss@wfu.edu. Phone: 336-758-3922.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

S.S.C. is grateful to Peter Wolynes for many fond memories as a graduate student, and he is honored by the invitation to contribute to this special issue. He also thanks Eric Westhof for preliminary advice and encouragement on the simulation methodology we took. The National Science Foundation (CBET-1232724) financially supported this work. We acknowledge computational support from the Wake Forest University DEAC Cluster, where the MD simulations were performed. H.G. and S.S.C. acknowledge financial support from the Wake Forest University Center for Molecular and Cellular Communication.

REFERENCES

(1) Crick, F. H. *Symp. Soc. Exp. Biol.* **1958**, *12*, 138−163.
(2) Robertus, J. D.; Ladner, J. E.; Finch, J. T.; Rhodes, D.; Brown, R. S.; Clark, B. F.; Klug, A. *Nature* **1974**, *250*, 546−551.
(3) Shi, H.; Moore, P. B. *RNA* **2000**, *6*, 1091−1105.
(4) Alexander, R. W.; Eargle, J.; Luthey-Schulten, Z. *FEBS Lett.* **2010**, *584*, 376−386.
(5) Crick, F. H. *J. Mol. Biol.* **1966**, *19*, 548−555.
(6) Westhof, E.; Auffinger, P. Transfer RNA Structure. In *eLS*; John Wiley & Sons, Ltd: 2012, DOI: 10.1002/9780470015902.a0000527.pub2.
(7) Phizicky, E. M.; Hopper, A. K. *Genes Dev.* **2010**, *24*, 1832−1860.
(8) Leopold, P. E.; Montal, M.; Onuchic, J. N. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 8721−8725.
(9) Clementi, C.; Nymeyer, H.; Onuchic, J. N. *J. Mol. Biol.* **2000**, *298*, 937−953.
(10) Karanicolas, J.; Brooks, C. L. *Protein Sci.* **2002**, *11*, 2351−2361.
(11) Scott, K. A.; Batey, S.; Hooton, K. A.; Clarke, J. *J. Mol. Biol.* **2004**, *344*, 195−205.
(12) Cho, S. S.; Levy, Y.; Wolynes, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 434−439.
(13) Cole, P. E.; Yang, S. K.; Crothers, D. M. *Biochemistry* **1972**, *11*, 4358−4368.
(14) Yang, S. K.; Crothers, D. M. *Biochemistry* **1972**, *11*, 4375−4381.
(15) Sorin, E. J.; Nakatani, B. J.; Rhee, Y. M.; Jayachandran, G.; Vishal, V.; Pande, V. S. *J. Mol. Biol.* **2004**, *337*, 789−797.
(16) Ding, F.; Sharma, S.; Chalasani, P.; Demidov, V. V.; Broude, N. E.; Dokholyan, N. V. *RNA* **2008**, *14*, 1164−1173.
(17) Hilbers, C. W.; Shulman, R. G.; Kim, S. H. *Biochem. Biophys. Res. Commun.* **1973**, *55*, 953−960.
(18) Treiber, D. K.; Williamson, J. R. *Curr. Opin. Struct. Biol.* **1999**, *9*, 339−345.
(19) Shelton, V. M.; Sosnick, T. R.; Pan, T. *Biochemistry* **2001**, *40*, 3629−3638.

(20) Serebrov, V.; Clarke, R. J.; Gross, H. J.; Kisselev, L. *Biochemistry* **2001**, *40*, 6688−6698.
(21) Chen, S. J. *Annu. Rev. Biophys.* **2008**, *37*, 197−214.
(22) Bhaskaran, H.; Rodriguez-Hernandez, A.; Perona, J. J. *RNA* **2012**, *18*, 569−580.
(23) Hyeon, C.; Thirumalai, D. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6789−6794.
(24) Cho, S. S.; Pincus, D. L.; Thirumalai, D. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 17349−17354.
(25) Biyun, S.; Cho, S. S.; Thirumalai, D. *J. Am. Chem. Soc.* **2011**, *133*, 20634−20643.
(26) Hyeon, C.; Thirumalai, D. *Nat. Commun.* **2011**, *2*, 487.
(27) Mathews, D. H.; Sabina, J.; Zuker, M.; Turner, D. H. *J. Mol. Biol.* **1999**, *288*, 911−940.
(28) Cruz, J. A.; Westhof, E. *Cell* **2009**, *136*, 604−609.
(29) Veitshans, T.; Klimov, D.; Thirumalai, D. *Folding Des.* **1997**, *2*, 1−22.
(30) Cho, S. S.; Levy, Y.; Wolynes, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 586−591.
(31) Feig, M.; Karanicolas, J.; Brooks, C. L. *J. Mol. Graphics Modell.* **2004**, *22*, 377−395.
(32) MacKerell, A. D.; Banavali, N.; Foloppe, N. *Biopolymers* **2000**, *56*, 257−265.
(33) Perez, A.; Marchan, I.; Svozil, D.; Sponer, J.; Cheatham, T. E.; Laughton, C. A.; Orozco, M. *Biophys. J.* **2007**, *92*, 3817−3829.
(34) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781−1802.
(35) Gosavi, S.; Chavez, L. L.; Jennings, P. A.; Onuchic, J. N. *J. Mol. Biol.* **2006**, *357*, 986−996.
(36) Hills, R. D., Jr.; Brooks, C. L., III. *J. Mol. Biol.* **2008**, *382*, 485−495.
(37) Hills, R. D., Jr.; Kathuria, S. V.; Wallace, L. A.; Day, I. J.; Brooks, C. L., III; Matthews, C. R. *J. Mol. Biol.* **2010**, *398*, 332−350.
(38) Bryngelson, J. D.; Onuchic, J. N.; Socci, N. D.; Wolynes, P. G. *Proteins* **1995**, *21*, 167−195.
(39) Wolynes, P. G.; Onuchic, J. N.; Thirumalai, D. *Science* **1995**, *267*, 1619−1620.
(40) Onuchic, J. N.; Wolynes, P. G. *Curr. Opin. Struct. Biol.* **2004**, *14*, 70−75.
(41) Brooks, C. L., III; Onuchic, J. N.; Wales, D. J. *Science* **2001**, *293*, 612−613.
(42) Cao, S.; Chen, S. J. *Nucleic Acids Res.* **2006**, *34*, 2634−2652.
(43) Bailor, M. H.; Mustoe, A. M.; Brooks, C. L., III; Al-Hashimi, H. M. *Curr. Opin. Struct. Biol.* **2011**, *21*, 296−305.
(44) Thirumalai, D.; Hyeon, C. *Biochemistry* **2005**, *44*, 4957−4970.
(45) Wu, M.; Tinoco, I. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 11555−11560.
(46) Tinoco, I.; Bustamante, C. *J. Mol. Biol.* **1999**, *293*, 271−281.
(47) Koculi, E.; Cho, S. S.; Desai, R.; Thirumalai, D.; Woodson, S. A. *Nucleic Acids Res.* **2012**, DOI: 10.1093/nar/gks468.