JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

pubs.acs.org/JACS

Noncanonical Secondary Structure Stabilizes Mitochondrial tRNA$^{Ser(UCN)}$ by Reducing the Entropic Cost of Tertiary Folding

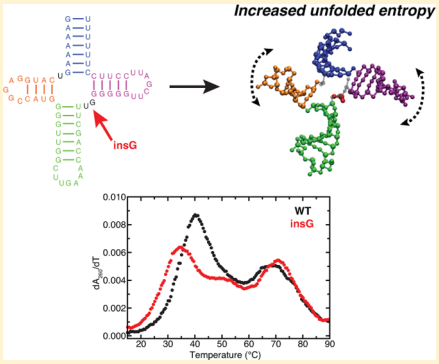
Anthony M. Mustoe,[†] Xin Liu,[‡] Paul J. Lin,[‡] Hashim M. Al-Hashimi,[∥] Carol A. Fierke,[†,‡,§] and Charles L. Brooks, III*[,†,‡]

[†]Department of Biophysics, [‡]Department of Chemistry, and [§]Department of Biological Chemistry, University of Michigan, Ann Arbor, Michigan 48109-1055, United States

[∥]Departments of Biochemistry and Chemistry, Duke University School of Medicine, Durham, North Carolina 27710, United States

Supporting Information

ABSTRACT: Mammalian mitochondrial tRNA$^{Ser(UCN)}$ (mt-tRNA^{Ser}) and pyrrolysine tRNA (tRNA^{Pyl}) fold to near-canonical three-dimensional structures despite having noncanonical secondary structures with shortened interhelical loops that disrupt the conserved tRNA tertiary interaction network. How these noncanonical tRNAs compensate for their loss of tertiary interactions remains unclear. Furthermore, in human mt-tRNA^{Ser}, lengthening the variable loop by the 7472insC mutation reduces mt-tRNA^{Ser} concentration in vivo through poorly understood mechanisms and is strongly associated with diseases such as deafness and epilepsy. Using simulations of the TOPRNA coarse-grained model, we show that increased topological constraints encoded by the unique secondary structure of wild-type mt-tRNA^{Ser} decrease the entropic cost of folding by ∼2.5 kcal/mol compared to canonical tRNA, offsetting its loss of tertiary interactions. Further simulations show that the pathogenic 7472insC mutation disrupts topological constraints and hence destabilizes the mutant mt-tRNA^{Ser} by ∼0.6 kcal/mol relative to wild-type. UV melting experiments confirm that insertion mutations lower mt-tRNA^{Ser} melting temperature by 6−9 °C and increase the folding free energy by 0.8−1.7 kcal/mol in a largely sequence- and salt-independent manner, in quantitative agreement with our simulation predictions. Our results show that topological constraints provide a quantitative framework for describing key aspects of RNA folding behavior and also provide the first evidence of a pathogenic mutation that is due to disruption of topological constraints.



INTRODUCTION

Mammalian mitochondrial tRNAs (mt-tRNAs), along with the mitochondrial ribosome, are responsible for translating 13 proteins encoded by the mitochondrial DNA that are central to oxidative phosphorylation.[1] In order to fulfill this essential biological role, mt-tRNAs must fold to the well-known L-shaped tRNA 3D structure, and mutations to mt-tRNA genes that disrupt folding have been implicated in a growing number of human diseases.[1]

Compared to canonical cytosolic tRNAs (cc-tRNAs), mt-tRNA folding thermodynamics remains poorly understood. Many mt-tRNAs possess "bizarre" secondary structures that deviate significantly from the canonical tRNA cloverleaf, characterized by shortened interhelical linkers and apical loops, lengthened helices, and sometimes lacking entire helical domains.[2,3] Furthermore, mt-tRNAs frequently lack otherwise universally conserved tertiary interactions important for stabilizing the tRNA 3D fold.[2,3] Understanding how mt-RNAs fold to near-canonical 3D structures despite their noncanonical morphologies has the potential to provide broader insights into the design principles of RNA tertiary folding.

Recent studies of canonical tRNA by our groups led us to hypothesize that the noncanonical secondary structures of mt-tRNAs may be central to their tertiary structure stability.[4] Due to the hierarchical nature of the RNA free energy landscape, tertiary structure folding generally proceeds from a state that consists of prefolded secondary structure helices.[5,6] Using simulations of cc-tRNAs, we showed that so-called topological constraints posed by the connectivity and sterics of prefolded secondary structure impose large free energy costs on adopting different 3D conformations, contributing to the specificity, stability, and cooperativity of cc-tRNA tertiary folding.[4] Furthermore, evolutionary conservation of cc-tRNA secondary structure was correlated with conservation of topological constraints. Due to their uniquely shortened interhelical linkers, mt-tRNAs should experience significantly greater topological constraints than cc-tRNA. This additional constraining of the unfolded conformational space may in turn reduce the entropic penalty of 3D folding, potentially explaining how mt-tRNAs fold despite possessing fewer tertiary interactions.

A good system for exploring this hypothesis is human mt-tRNA$^{Ser(UCN)}$ (hereafter referred to as mt-tRNA^{Ser}),[7,8] which has an identical secondary structure and similar sequence as the rare cytosolic pyrolysine tRNA (tRNA^{Pyl}) species.[9,10] In mt-

Received: December 22, 2014
Published: February 23, 2015

DOI: 10.1021/ja5130308
J. Am. Chem. Soc. 2015, 137, 3592−3599

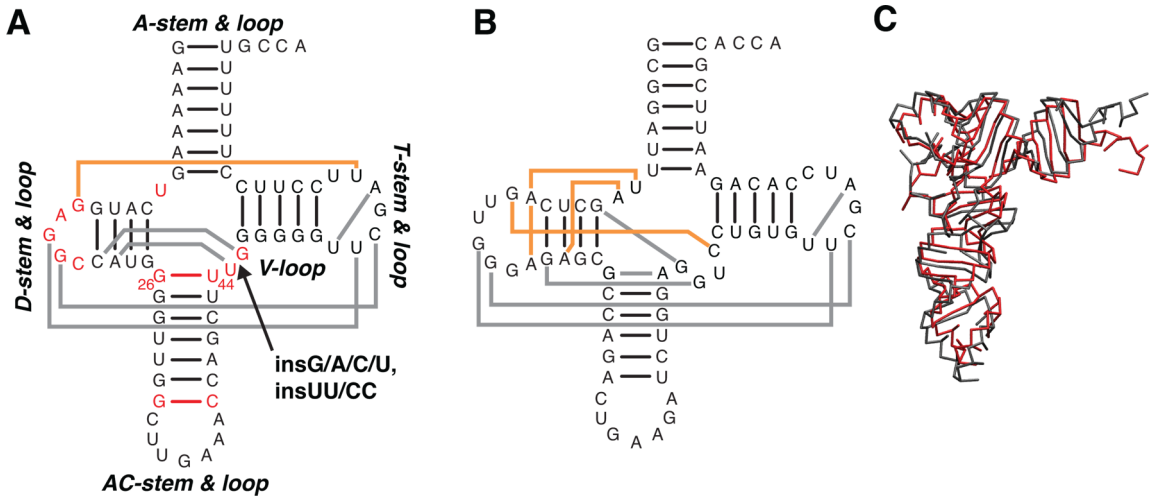


Figure 1. (A) Secondary structure of human mt-tRNA$^{Ser(UCN)}$ and the location of the insertion mutations. Secondary structure features that differ from cc-tRNA are indicated by red. Gray and orange lines indicate tertiary interactions inferred from the tRNA^{Pyl} crystal structure.[13] Tertiary interactions shared by cc-tRNA are shown in gray, and novel interactions are shown in orange. (B) Secondary structure of yeast tRNA^{Phe} (cc-tRNA). Conserved tertiary interactions missing in mt-tRNA^{Ser} are shown in orange. (C) Superposition of TOPRNA representations of the cc-tRNA[21] (gray) and tRNA^{Pyl}[13] (red) crystal structures.

tRNA^{Ser}, the A/D-loop is only one nucleotide (nt) long, compared to 2 nt in all other tRNA species (Figure 1A,B).[11,12] The noncanonical base pair atop the AC-stem is also typically replaced with a canonical base pair, and the V-loop is shortened from 5 to 3 nt.[2,10,11] This unique secondary structure precludes formation of several normally critical tertiary pairs (Figure 1A,B). Nevertheless, while less stable than cc-tRNAs, mammalian mt-tRNA^{Ser} and archaeal tRNA^{Pyl} transcripts fold to near-canonical 3D structures (Figure 1C).[9,13−15]

Intriguingly, a mutation expected to disrupt the topological constraints of mt-tRNA^{Ser} is also pathogenic in humans (Figure 1A). The 7472insC mutation, hereafter referred to as the insG mutation, lengthens the polyG tract that terminates in the V-loop of human mt-tRNA^{Ser} (Figure 1A) and has been strongly implicated as a cause of deafness, progressive encephalomyopathy, and MERFF (myoclonic epilepsy with ragged red fibers).[8,16−18] Consistent with pathogenesis arising from destabilized tertiary structure, the insG mutation does not alter the folded structure but impairs transcript processing and decreases the concentration of mature mt-tRNA^{Ser} by 65%.[20,7,19] One proposal is that the inserted G is accommodated into the T-loop, destabilizing the tertiary pairs formed with the D-loop.[16] However, the T-loop is highly conserved, and disrupting its local structure may be expected to cause a more severe phenotype.[2,12] Alternatively, the inserted G could be accommodated into the V-loop. This would give the V-loop a similar length to that found in cc-tRNA but would increase the topologically available conformational space relative to wild-type (WT) mt-tRNA^{Ser}.

In this study, we explore the role of topological constraints in mt-tRNA^{Ser} folding using simulations of the specialized TOPRNA coarse-grained model.[22] TOPRNA specifically isolates the effects of topological constraints on the RNA free energy landscape by treating base-paired regions as permanent A-form helices, single-stranded nucleotides as freely rotatable chains, and ignores both electrostatic and attractive interactions. Consistent with our hypothesis, our simulations indicate that secondary structure plays a critical role in reducing the entropic cost of WT mt-tRNA^{Ser} folding, and that the pathogenic insG mutation destabilizes mt-tRNA^{Ser} by increasing the size of the accessible unfolded conformational space. These results provide the first example of a topological-constraint-based disease mechanism and suggest that topological constraints are a key design principle of RNA tertiary folding.

RESULTS

Topological Constraints Stabilize WT mt-tRNA$^{Ser(UCN)}$.

To determine whether increased topological constraints from the noncanonical secondary structure of WT mt-tRNA^{Ser} help stabilize tertiary structure, we performed extensive temperature replica exchange simulations of mt-tRNA^{Ser} using an updated version of the TOPRNA[22] coarse-grained model (Supporting Information, Figure S1). We quantified whether mt-tRNA^{Ser} is more topologically constrained than cc-tRNA by counting the number of unique global conformations sampled by the simulations, using Euler angles (α_h, β_h, γ_h), to measure the relative orientation of two helices, where α_h and γ_h describe the helical twists and β_h the interhelical bend (Figure 2A).[23,24] Three sets of Euler angles measured between the A-, D-, and T-stems relative to the AC-stem in turn provide a nine-dimensional description of global tRNA 3D conformation that can be discretized onto a finite grid.[4] Consistent with our expectations, the shorter A/D- and V-loops of mt-tRNA^{Ser} reduce the number of interhelical conformations sampled by mt-tRNA^{Ser} by 2-fold relative to cc-tRNA (2.5% vs 5.7%; Figure 2B). We previously showed that one of the primary mechanisms through which topological constraints restrict cc-tRNA conformation is by coupling the orientation of all four tRNA helices together, giving rise to long-range correlated motions that contribute to folding cooperativity.[4] Consistent with this explanation, we find that helices in mt-tRNA^{Ser} are 10−90% more correlated than in cc-tRNA (Figure S2).

To quantify the extent to which the increased topological restriction of WT mt-tRNA^{Ser} biases the molecule toward native-like conformations, we computed the probability P_{nat} of mt-tRNA^{Ser} and cc-tRNA sampling native-like 3D conformations in our simulations, which we directly ascribe to the

DOI: 10.1021/ja5130308
J. Am. Chem. Soc. 2015, 137, 3592−3599

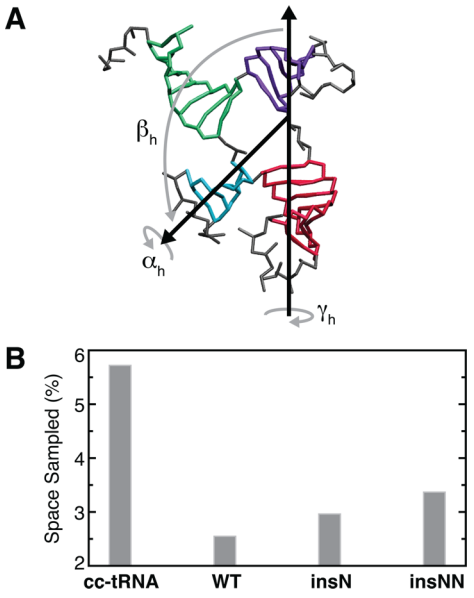


Figure 2. (A) Representative snapshot from a TOPRNA simulation of mt-tRNA^{Ser} illustrating the Euler angle representation of the conformation between the AC-stem (red) and D-stem (blue). Loops, which are treated as freely rotatable chains by TOPRNA, are colored gray. (B) Fraction of theoretically possible interhelical (α_h, β_h, γ_h) conformations sampled by TOPRNA simulations.

topological constraint contribution to the free energy of folding through

$$\Delta G^{topo}_{fold} = -RT\ln\left(\frac{P_{nat}}{1 - P_{nat}}\right) \quad (1)$$

As TOPRNA lacks attractive interactions, our simulations spend the majority of their time sampling extended conformations and P_{nat} is exceedingly small; depending on the metric used to define native-like structure, P_{nat} ranged from 10^{-2} to 10^{-5}. ΔG^{topo}_{fold} is hence large and positive for both molecules, reflecting the entropic penalty of folding that must be offset by favorable tertiary interactions. Strikingly, however, we find that the $\Delta\Delta G^{topo}_{fold}$ (T = 300 K) between cc-tRNA and mt-tRNA^{Ser} is 1.5−2.5 ± 0.1 kcal/mol (Figure 3 and Supporting Information Figure S3; see discussion below), indicating that topological constraints can stabilize mt-tRNA^{Ser} by as much as this amount. Notably, this is comparable to the favorable free energy provided by one to two base pairs,[25]

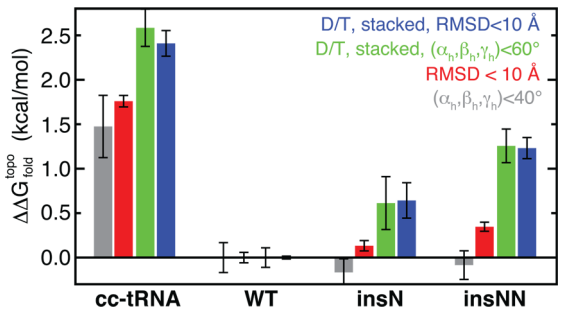


Figure 3. Differences in the topological constraint contribution to the folding free energy $\Delta\Delta G^{topo}_{fold} = \Delta G^{topo}_{fold}(i) - \Delta G^{topo}_{fold}(WT)$. Values and error bars represent the mean and standard deviation of the $\Delta\Delta G$ computed by block averaging over thirds of the simulations.

strongly supporting that topological constraints provide compensation for the decreased number of tertiary interactions in mt-tRNA^{Ser}.

In theory, $\Delta\Delta G^{topo}_{fold}$ between mt-tRNA^{Ser} and cc-tRNA should be independent of the metric used to compute P_{nat}. Specifically, while P_{nat} will vary with the definition of native structure, it should vary proportionally for both tRNAs. Interestingly, however, $\Delta\Delta G^{topo}_{fold}$ varies significantly based on the native metric. To measure global structure, we required tRNA conformations to have <10 Å root-mean-square deviation (rmsd) from their respective crystal structure. Alternatively, we required all interhelical (α_h, β_h, γ_h) angles to be within 40° of their crystal structure values. Both of these global definitions yielded $\Delta\Delta G^{topo}_{fold}$ estimates of 1.5−1.7 kcal/mol (Figure 3). By contrast, when P_{nat} was recomputed also requiring that the tRNAs possess contacts between the D-to-T-loop and both of the native interhelical stacks, $\Delta\Delta G^{topo}_{fold}$ increased to ∼2.5 kcal/mol (Figure 3). This suggests that topological constraints stabilize mt-tRNA^{Ser} by two mechanisms. First, as captured by global metrics alone, the cost of forming near-native interhelical orientations is reduced relative cc-tRNA. Second, the shorter connecting loops of mt-tRNA^{Ser} further restrict the translational freedom of the helices, reducing the entropic cost of forming D- and T-loops and stacking contacts (Figure S3).

The above analysis reports on the free energy cost of each tRNA forming the native structure. Another important aspect of the energy landscape is whether or not the tRNAs can form competing tertiary folds or, in other words, the specificity of the folding landscape. We previously found that topological constraints prevent cc-tRNA from forming non-native tertiary folds,[4] and thus we assessed whether topological constraints also increase mt-tRNA^{Ser} folding specificity. As before,[4] we identified the 500 conformers sampled by TOPRNA that have the best interloop packing (and thus potential to form putative tertiary interactions). Notably, only 0.4% of mt-tRNA^{Ser} compared to ∼12% of cc-tRNA best-packed conformers possess non-native contacts (Figure 4). The degree to which

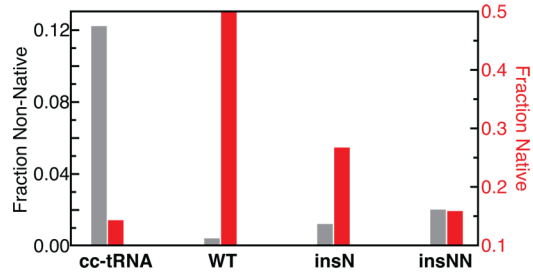


Figure 4. Fraction of the 500 best-packed conformations of each tRNA that have non-native loop−loop contacts (gray) or that have D-to-T-loop contacts and both interhelical stacks (red). The fraction of conformers with both native contacts is weighted by entropy. Note that the third conformational possibility—tRNAs that lack stacking or D-to-T-loop contacts—is not shown.

individual conformers are close in structure to other best-packed conformers also provides a measure of their relative conformational entropy.[26] Compared to cc-tRNA, near-native 3D conformations of mt-tRNA^{Ser} have significantly higher entropies (Figure S4), which is summarized in Figure 4 as the entropy-weighted fraction of best-packed conformers that possess D-to-T-loop contacts and both interhelical stacks. Thus, mt-tRNA^{Ser} topological constraints both lower the cost of

DOI: 10.1021/ja5130308
J. Am. Chem. Soc. 2015, 137, 3592−3599

forming native interactions (ΔG^{topo}_{fold}) and increase the energy gap between native and non-native folds.[27]

The mt-tRNA^{Ser} simulation analyzed above treated the G26−U44 base pair as a permanent feature of secondary structure (Figure 1A). This pair is conserved as either a GU or AU pair in humans[28] and as a canonical pair in the majority of mt-tRNA^{Ser} and tRNA^{Pyl} species. However, A$^+$C, AA, or GA pairs at this position are also observed in other species.[2,10,11] To determine the contribution of intact G26−U44 pairing to our observations, we performed an additional TOPRNA simulation of mt-tRNA^{Ser} where these residues were treated as single strands (mt-tRNA$^{Ser}_{noBP}$). Compared to when G26−U44 is paired, the mt-tRNA$^{Ser}_{noBP}$ simulation samples ∼50% more interhelical conformations, ΔG^{topo}_{fold} is increased by ∼1 kcal/mol, and folding specificity is reduced ∼5% (Figure S5). This decrease in stability provides a feasible explanation for the evolutionary preference for WC/GU pairs between residues 26 and 44; for species without canonical pairing, other sequence changes in the D- and T-loops may serve to strengthen the tertiary interaction network. Significantly, however, ΔG^{topo}_{fold} in mt-tRNA$^{Ser}_{noBP}$ is still ∼1.5 kcal/mol more favorable than in cc-tRNA.

insG Mutation Destabilizes mt-tRNA^{Ser} by Disrupting its Topological Constraints. We next explored whether the insG mutation disrupts the ability of topological constraints to stabilize mt-tRNA^{Ser}, thereby potentially explaining its pathogenicity (Figure 1A). Consistent with our hypothesis, a TOPRNA simulation of mt-tRNA^{Ser} with an insN mutation in the V-loop samples ∼16% more interhelical conformations than the WT simulation (Figure 2B). (We use the insN notation to denote that our simulations are independent of inserted nucleotide identity.) Likewise, the interhelical correlations of the insN mutant are decreased ∼1−20% compared to WT (Figure S2).

Significantly, this reduction in topological constraints is accompanied by a ∼0.6 kcal/mol increase in the ΔG^{topo}_{fold} of the insN mutant compared to WT and a 50% decrease in folding specificity (Figures 3 and 4 and Supporting Information Figures S3 and S4). This is a consequential destabilization of mt-tRNA^{Ser} given its expected instability.[14] Near the melting temperature, a 0.6 kcal/mol increase in folding energy should decrease the folded population by ∼40%, in reasonable qualitative agreement with the 65% decrease in insG mutant concentration observed in cells.[19] An analogous insA mutation in the related tRNA^{Pyl} reduces its translational activity by a similar amount (83%),[29] consistent with the insN mutation being destabilizing in a sequence-independent manner. This destabilization arises from an increased cost of forming tertiary contacts, as near-native global 3D conformations are formed with similar energies (Figure 3). Consistent with the insN mutant's decreased interhelical correlations (Figure S2), ∼0.2 kcal/mol of this increase is due to a loss of cooperativity among interhelical stacking and D-to-T-loop contacts (Figure S3). Thus, in part, due to decoupling of the helices, greater conformational freedom near the native state increases the entropic cost of forming tertiary interactions in the mutant. Simulations of the insN mutant without a G26−U44 base pair also demonstrated similar decreases in topological confinement, $\Delta\Delta G^{topo}_{fold}$ between mutant and WT, and folding specificity (Figure S5).

To further establish the trend between topological constraints and folding, we also performed simulations of an insNN V-loop mutant of mt-tRNA^{Ser} (Figure 1A). As expected, these simulations indicated that the insNN mutant has the effect of further decreasing topological confinement and folding specificity and increasing the free energy cost of folding (Figures 2−4 and Figures S2−S4). Similar results were also observed for an insNN simulation performed without a G26−U44 base pair (Figure S5).

Melting Experiments Reveal Destabilization of Insertion Mutants in Quantitative Agreement with TOPRNA Predictions. Our simulations make several strong predictions regarding mt-tRNA^{Ser} stability: (i) Single and double insertions in the V-loop should destabilize tertiary structure by ∼0.6 and ∼1.2 kcal/mol at 300 K, respectively. (ii) This destabilization should be independent of the inserted nucleotide(s) identity, assuming that, like the insG mutation,[7] the insertions do not cause secondary structure misfolding. (iii) This destabilization should be attributable to an increase in the entropic cost of folding for the mutants. Notably, from this last prediction, we can estimate the expected decrease in melting temperature (T_m) of the mutants. As detailed in Supporting Information, isolated changes in the entropy of the unfolded state will cause, to a first-order approximation, proportional changes in T_m through

$$T_m^{WT}\Delta S^{WT}_{fold,3Dconf} \approx T_m^{mut}\Delta S^{mut}_{fold,3Dconf} \quad (2)$$

Estimating the change in 3D conformational entropy upon folding, $\Delta S_{fold,3Dconf}$, from our simulations, and setting T_m^{WT} = 45 °C, we predict that the T_m of the insN and insNN mutants should be decreased by 8 and 15 °C, respectively (Figure S6).

To test these predictions, we used in vitro transcription to prepare WT, insG, insC, insU, insA, insCC, and insUU human mt-tRNA^{Ser} constructs and characterized their folding stabilities by UV melting experiments. Experiments were done at both near-physiological (2 mM $MgCl_2$; 150 mM NaCl) and "stabilizing" (5 mM $MgCl_2$; 150 mM NaCl) solution conditions. At both conditions, WT tRNA^{Ser} exhibits a pronounced cooperative melting transition at 40 or 45 °C followed by multiple broad transitions at higher temperatures (Figure 5A,B). The increase in melting temperature to 45 °C at higher $MgCl_2$ concentration is consistent with the expected stabilizing effects of $MgCl_2$, and both the shape and T_m of these transitions are in good agreement with prior UV melting studies of bovine mt-tRNA$^{Ser(UCN)}$ transcripts.[14] NMR melting studies of the bovine transcript further demonstrated that the first transition corresponds to tertiary structure melting coupled with some loss of D-stem and T-stem structure.[14] We thus also assign the low T_m = 40 or 45 °C to tertiary structure melting and the higher temperature transitions to unfolding of the remaining secondary structure stems.

The insG, insA, and insU mutants exhibit similar melting curves at both $MgCl_2$ concentrations, yet the T_m of the first transition is decreased by 6−9(±1) °C compared to WT and decreased by 11 ± 1 °C in the insUU mutant (Figure 5A,B). Strikingly, these observed decreases closely match the 8 and 15 °C ΔT_m predictions of our simulations (Figure 5C,D); we consider differences of a few degrees to be within the expected accuracy of our approximated ΔT_m predictions. While we cannot unequivocally assign the first transition in the mutants to tertiary structure melting, several observations beyond the gross similarities of the melting curves support the assumption that the mutants adopt a native-like tertiary fold. First, prior enzymatic probing studies found no structural differences between mature WT and insG mt-tRNA^{Ser} purified from cells.[7]

DOI: 10.1021/ja5130308
J. Am. Chem. Soc. 2015, 137, 3592−3599

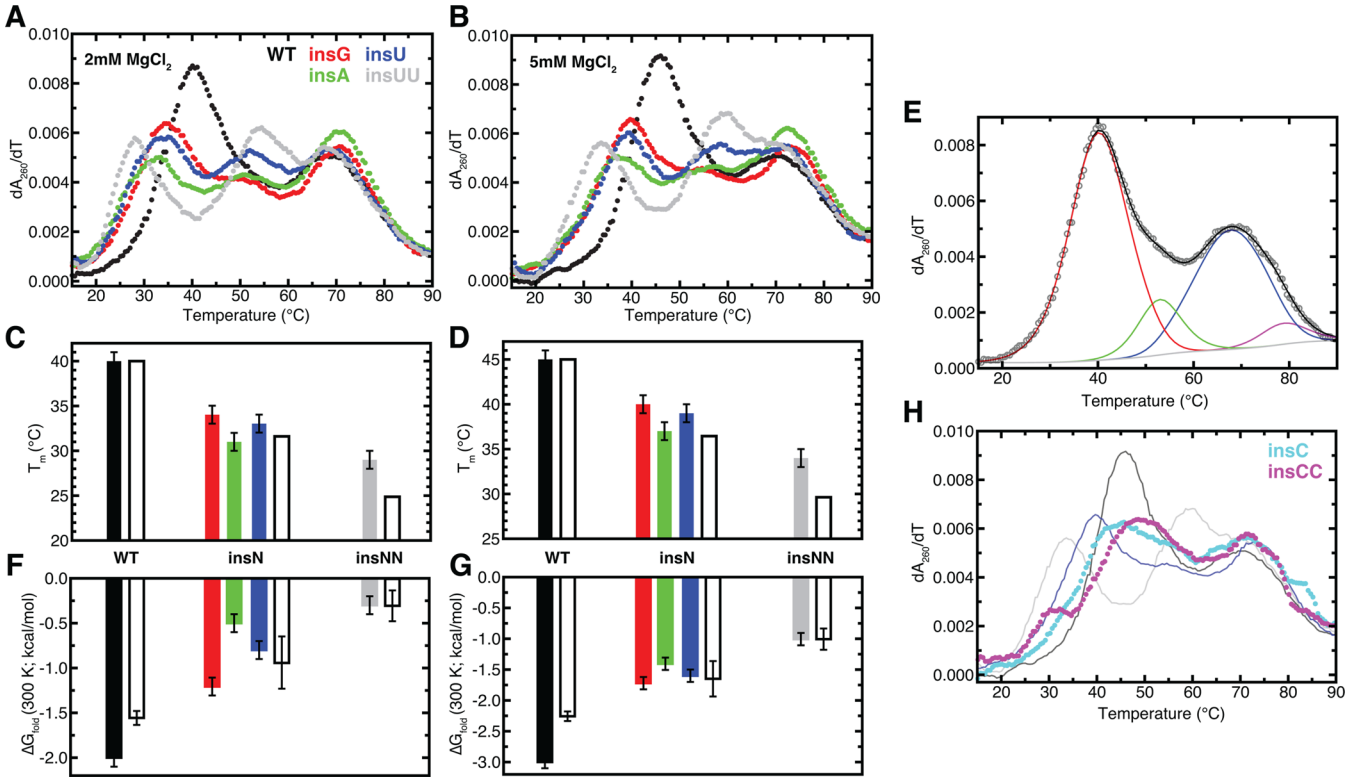


Figure 5. (A,B) Derivative of absorbance at 260 nm of mt-tRNA^{Ser} species in the presence of 2 mM $MgCl_2$ (A,C,F) or 5 mM $MgCl_2$ (B,D,G). All melts were performed in a background of 20 mM sodium cacodylate (pH 7.2) and 150 mM NaCl. Curves of different molecules are colored according to the key in (A). (C,D) Tertiary structure melting temperatures determined from van't Hoff fits to melting curves. Error bars represent estimated 1 °C error (see Materials and Methods). Bars are colored by mutant according to (A). TOPRNA predicted values, in reference to the WT T_m, are shown with open bars and represent the mean of the different ΔT_m estimates in Figure S6. (E) Example van't Hoff fit to the WT transcript at 2 mM $MgCl_2$. The different transitions are shown in color, the baseline in gray, and the overall fit in black. (F,G) ΔG_{fold} determined from van't Hoff fits extrapolated to 300 K. Color scheme is the same as in (C,D). TOPRNA predictions are referenced to the insUU ΔG_{fold}. (H) Melting curves of insC and insCC mutants at 5 mM $MgCl_2$. WT, insG, and insUU melting curves are shown by lines colored according to the key in (A) for reference.

The insG pre-tRNA transcripts are also processed by nuclear RNase P and mitochondrial tRNase Z under solution conditions highly comparable to those used here,[20] indicative of the formation of a native structure. Finally, the observed sensitivity of T_m to Mg^{2+} concentration[30] and the incremental decrease in T_m between the insU and insUU mutants argues against the first transition only representing secondary structure unfolding. Combined, these data support the proposal that changes in topological constraints are responsible for destabilizing the tertiary structure of the mutant transcripts.

Interestingly, the insG/A/U and insUU mutants also melt less cooperatively than the WT transcript (reflected by lower maximum derivatives and greater transition widths in Figure 5A,B), suggesting that the unfolding enthalpies are different. Indeed, van't Hoff analysis confirms that the unfolding enthalpy is typically lower in the mutants (see below; Supporting Information Tables S1 and S2). While this decrease is qualitatively consistent with our prediction that tertiary structure should fold less cooperatively in the mutants (Figure S3), it is more likely that the decrease reflects differential hydration and ion interactions with the folded state and/or differences in base stacking cooperativity. Note that, because of potential entropy−enthalpy compensation effects, this does not necessarily contradict that an increased entropic cost of folding is primarily responsible for decreased stability of the mutants.

We further used a van't Hoff analysis to extract thermodynamic parameters of the first melting transition for the WT, insA, insU, insG, and insUU constructs.[31] Each melting curve was globally fit using a four-state sequential unfolding model (Figure 5E). While fitting generally yielded multiple solutions, all curves shared a "parsimonious" fit that gave similar thermodynamic parameters for each of the four unfolding transitions (Figure 5E and Supporting Information Tables S1 and S2). As alternative fits varied between melting curves or were inconsistent with NMR data collected on the bovine mt-tRNA^{Ser} transcript,[14] we restricted our analysis to the parsimonious fit (Supporting Information). Depending on the mt-tRNA and $MgCl_2$ concentration, the unfolding enthalpy of the first transition varied between 40 ± 5 and 54 ± 3 kcal/mol (Tables S1 and S2), consistent with melting of tertiary structure and some D-stem base pairs.[14] Notably, when used to derive $\Delta G_{fold,vH}$(300 K), the van't Hoff analysis indicates that the insG/A/U mutants are destabilized by 0.8−1.7(±0.1) kcal/mol relative to WT, and the insUU mutant by 1.7−2.0(±0.1) kcal/mol (Figure 5F,G). Again, these values are remarkably close to our TOPRNA predictions of $\Delta\Delta G^{topo}_{fold}$ = 0.6 ± 0.2 and 1.2 ± 0.1 kcal/mol, respectively. The $\Delta\Delta G_{fold,vH} \approx 0.6$ kcal/mol between the insN and insUU mutants exactly matches our prediction. Disruptions of Mg^{2+} interactions or tertiary structure hydrogen bonding caused by the insN mutations, which are not included in our TOPRNA model, can explain the variable and larger than predicted $\Delta\Delta G_{fold,vH}$ between the WT and insN mutants.

DOI: 10.1021/ja5130308
J. Am. Chem. Soc. 2015, 137, 3592−3599

By contrast, the insC and insCC mutants lack a well-defined tertiary structure melting transition (Figure 5H). Both mutants appear to exhibit small shoulders near the predicted T_m values of the insN and insNN mutants, but these shoulders were poorly fit by a van't Hoff analysis (not shown). We propose this different melting behavior is due to secondary structure misfolding, perhaps involving mispairing between the inserted C and the complementary Gs atop the AC-stem (Figure 1A).

DISCUSSION

How mt-tRNA$^{Ser(UCN)}$ and tRNA^{Pyl} stably fold to near-canonical 3D structures despite their noncanonical architectures has long been poorly understood. Our results indicate that these tRNAs employ a novel mechanism where the entropic cost of folding is reduced by a more constrained secondary structure, thereby compensating for a fewer number of tertiary interactions. In particular, our simulations predict that increased topological constraints stabilize the WT mt-tRNA^{Ser} 3D structure by as much as −2.5 kcal/mol and also substantially increase folding specificity compared to cc-tRNA. In the insA, insU, and insG mutants, lengthening the V-loop by a single nucleotide destabilizes mt-tRNA^{Ser} by 0.8−1.7 kcal/mol, leading to a 6−9 °C decrease in melting temperature. A 2 nt insertion destabilizes the insUU mutant by 1.7−2.0 kcal/mol and decreases T_m by 11 °C. The agreement between these experimentally measured values and those predicted by our simulations, as well as the sequence and salt independence of the mutant destabilization, provides strong evidence that topological constraints are indeed responsible for these changes.

Importantly, our results provide a potential explanation for why the insG mutation is pathogenic. Prior studies found that mt-tRNA^{Ser} concentration is decreased 65% in mitochondria when all mt-DNA plasmids contain the insG mutation, which presumably disrupts mitochondrial protein synthesis and hence cellular respiration.[19] This decrease is primarily due to fast degradation of misprocessed and/or unprocessed transcripts.[20,7] However, why the mutant was specifically degraded was unclear. Our data show that WT mt-tRNA^{Ser} transcripts are only marginally stable, such that the destabilized insG mutant is primarily unfolded at physiological temperatures (T_m < 37 °C at near-physiological $MgCl_2$ concentrations). We propose that this shift in folding equilibrium results in increased flux through degradative pathways that compete with maturation (Figure 6). Notably, the 65% decrease of mature insG mt-tRNA^{Ser} in cells[19] is similar to the ∼50% decrease in folding equilibrium expected from a 0.8 kcal/mol higher folding free energy. For the mutant transcripts that enter the maturation pathway, nucleotide modifications likely help stabilize the tertiary structure, explaining the similar cellular half-life of mature insG and WT mt-tRNA^{Ser}.[7] Interestingly, recent studies have shown that unstable cytosolic tRNAs are rapidly degraded by a rapid tRNA decay (RTD) surveillance pathway.[32] We speculate that a similar RTD pathway may exist in mitochondria. Although mt-tRNA^{Ser} lacks consecutive 5′-Gs shown to be important for cytosolic RTD, it possesses the requisite unstable acceptor stem.[32] In support of an RTD mechanism, destabilizing mutations to the mt-tRNA^{Ser} acceptor stem cause similar decreases in mitochondrial concentration and pathogenic phenotypes.[33,34] Finally, we note that the appearance of insG and not insA/U/C mutations in the human population does not necessarily indicate that the former is more tolerable; insG

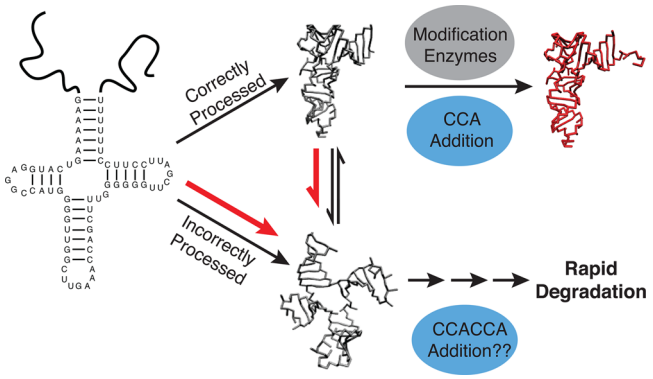


Figure 6. Proposed pathogenic mechanism of the insG mutation. The unstable transcript exists in equilibrium between folded and unfolded tertiary conformations, with only the folded conformation likely efficiently processed into mature tRNA. In competition with maturation is a degradation pathway that degrades unfolded and misprocessed tRNAs. The insG mutation shifts the transcript equilibrium toward unfolded conformations (red arrows), leading to increased misprocessing and degradation.[20] Degradation may be promoted by 3′-CCACCA addition as in cytosolic rapid decay.[32] Once modified, insG mt-tRNA^{Ser} molecules have stabilities similar to those in WT.[7]

mutations are naturally expected at a much higher frequency due to the adjacent G-tract in the T-stem.[16,35]

Inspection of other mt-tRNAs suggests that topological constraints may broadly compensate for noncanonical tertiary interaction networks. Of particular note, in mt-tRNA$^{Ser(AGY)}$, the entire D-stem and loop is replaced with a 3−5 nt single-strand linker between the A- and AC-stems.[2] This short linker, along with its shorter 3 nt V-loop, should place substantially greater constraints on the A-, AC-, and T-stems compared to cc-tRNAs. More generally, the significant majority of mt-tRNAs possess 4 nt V-loops compared to the 5 nt V-loops found in most cc-tRNAs.[2,4] mt-tRNA mutations that disrupt secondary structure are also particularly likely to be pathogenic.[36] However, insertion mutations comparable to 7472insC observed in other mt-tRNAs with more canonical secondary structures are benign polymorphisms,[28] indicating that the pathogenic mechanism we observe here is likely unique to the particular architecture of mt-tRNA$^{Ser(UCN)}$.

We emphasize that the true measure of our simulations lies in their ability to predict overall destabilization trends rather than their exact quantitative agreement with experiment. Clearly, forces ignored in our model, such as electrostatics and sequence-specific attractive interactions, play crucial roles in RNA tertiary folding.[37,38] It is also worth noting that the inherent instability of mt-tRNA^{Ser}—T_m for cc-tRNAs is generally >60 °C[14] compared to 45 °C measured here—makes it clear that greater topological constraints are insufficient to fully offset the significant loss of tertiary interactions in mt-tRNA^{Ser}. A full understanding of RNA tertiary folding requires accounting for all aspects of the RNA free energy landscape, including secondary structure misfolding, which we propose occurs in the insC and insCC mutants. Given these facts, we note that our near-quantitative prediction of the mutant ΔT_m is likely partially due to offsetting errors in our analysis, as we assumed that the conformational distribution sampled by TOPRNA directly matches that of real mt-tRNA^{Ser}. Nevertheless, we stress that our $\Delta\Delta G^{topo}_{fold}$ estimates are assumption free and, due to the additivity of free

DOI: 10.1021/ja5130308
J. Am. Chem. Soc. 2015, 137, 3592−3599

energy, reflect a real component of the RNA free energy landscape.[4,22]

A growing body of literature has suggested that secondary structure, and topological constraints in particular, plays an important role in RNA tertiary folding.[4,22−24,39−43] Significantly, our study is the first to experimentally measure the contribution of such topological constraints to the folding free energy of a complex RNA molecule. Indeed, to our knowledge, our study is the first to quantitatively predict *a priori* the destabilizing effect that a secondary structure mutation will have on RNA tertiary folding. The magnitude of the free energy differences we observe indicates that topological constraints play a comparable role to tertiary interactions in determining RNA tertiary stability. Notably, our study only explored small variations of the optimized tRNA secondary structure scaffold; more significant deviations from this scaffold would likely prohibitively increase the entropic penalty of tertiary folding. Extensive work on the hairpin ribozyme has demonstrated that its evolutionarily selected secondary structure reduces the entropic cost of folding,[44−47] although the origins of this entropic stabilization were unclear. Combined with our results, we propose that entropic stabilization by topological constraints may be a general design principle of RNA folding. Such a mechanism may be a critical bridge toward understanding tertiary structure stability, specificity, and ultimately rationally designing novel RNA folds.

MATERIALS AND METHODS

TOPRNA Model. As described elsewhere,[22] TOPRNA uses three beads to represent the sugar (S), base (B), and phosphate (P) moieties of each nucleotide. GC, AU, and GU paired B-beads are permanently bonded together, and contiguously paired regions are parametrized to adopt A-form helical structure. Single-stranded nucleotides are freely rotatable and do not experience attractive interactions. Paired B-beads experience small attractive interactions to one another, which is meant to simulate intrahelical stacking but also marginally favors interhelical stacking. However, no other beads experience this attractive term. All electrostatic forces are ignored.

Simulation Details. All simulations were performed using the updated force field described in Supporting Information. Initial coordinates of WT, insN, and insNN mt-tRNA^{Ser} were obtained by assembling the molecules from initially linear chains using the **init** option of *toprnaCreate.pl* (brooks.chem.lsa.umich.edu). Langevin dynamics temperature replica exchange simulations were performed in CHARMM[48] using the MMTSB toolset[49] for a total of 10^9 dynamics steps, using a 20 fs time step, 5 ps^{-1} friction coefficient, and 8 temperature windows spanning 300 to 450 K. The first 2×10^6 simulation steps were discarded as equilibration. Exchanges were attempted every 2000 dynamics steps, with exchange rates varying between 42 and 45% for all simulations. All other simulation parameters were set as previously described.[22]

Simulation Analysis. Analysis was restricted to conformations sampled at 300 K. Interhelical Euler angles were measured as previously described;[4,23] for the AC-stem, the three base pairs below G26−U44 were used for alignments to facilitate comparisons to cc-tRNA. The fraction of interhelical conformations sampled was computed by binning on a 60° grid, as described previously.[4] Distances in Euler angle space were computed as the amplitude of the single axis rotation needed to convert a given interhelical (α_h, β_h, γ_h) angle to the value measured in the crystal structure.[23] The D- and T-loops were considered to be in contact if the S-beads from at least one pair of residues were within 14 Å.[4] Interhelical stacking was evaluated using a modified version of the Tyagi and Mathews criteria as previously described.[4,50] The rmsd values were computed using P-beads, excluding the 3′-A/GCCA and V-loops.

The 500 best-packed conformations sampled by each simulation were identified using the energy function $E = n_l\varepsilon_l + n_s\varepsilon_s$, with n_l corresponding to the number of contacts between loop residues, n_s the number of interhelical stacks, and ε_s and ε_l scaling parameters set to −3.5 and −0.6 kcal/mol.[4] Stacking interactions were considered between the A- and D-stems, the D- and AC-stems, and T- and A-stems. The entropy of each best-packed conformation i was computed through

$$S_i = k_B \ln\left[1 + \sum_{j\neq i} \exp\left(\frac{-\mathrm{rmsd}_{i,j}^{\ 3}}{(10\ \text{Å})^3}\right)\right] \quad (3)$$

where the sum is done over all other best-packed conformations and $\mathrm{rmsd}_{i,j}$ is the P-bead rmsd computed excluding the V-loop and 3′-GCCA-loop.[4,26] The entropy-weighted ensemble average was computed as

$$\langle N\rangle = \frac{\sum_i N_i \exp(S_i/k_B)}{\sum_i \exp(S_i/k_B)} \quad (4)$$

UV Melting Experiments. RNA cloning and synthesis procedures are described in Supporting Information. Briefly, human mt-tRNA$^{Ser(UCN)}$ was subcloned into a pUC18 plasmid and prepared by runoff transcription using T7 RNA polymerase. Mutants were obtained by site-directed mutagenesis and similarly prepared by runoff transcription. All DNA templates were verified by sequencing.

Concentrated tRNA was diluted into a lightly buffered water solution containing ∼0.01 mM EDTA and ∼0.1 mM Tris (pH 7.2), denaturated at 95 °C for 2 min, refolded at room temperature for 15 min, followed by addition of folding buffer and incubation at room temperature for 30 min. Final solution conditions consisted of 0.5 μM RNA, 0.01 mM EDTA, 0.1 mM Tris, 150 mM NaCl, 20 mM Na-cacodylate (pH 7.2), and either 2 or 5 mM $MgCl_2$. The solution was placed in 1 cm capped quartz cuvettes and was topped with 50 μL of mineral oil to prevent evaporation. Melts were performed at a rate of 0.5 °C/min from 15 to 90 °C, monitoring absorbance at 260 nm on a Cary 100Bio UV/vis spectrophotometer. Derivatives were obtained by subtracting the buffer baseline, normalizing to the absorbance at 15 °C, and computing the slope of the linear least-squares fit to all data points within a ±3 °C temperature window.[31] Experimental replicates showed good agreement (not shown).

Global Melt Fit (GMF) provided by D.E. Draper (Johns Hopkins) was used to extract melting enthalpies and temperatures from the WT, insA, insU, insG, and insUU dAbs(260 nm)/dT melting data utilizing custom python wrapper scripts to interface with the GMF fitting routine.[31] Scripts and documentation are available for download at brooks.chem.lsa.umich.edu. Curves were globally fit assuming three, four, or five sequential transitions, with ΔH and ΔAbs of all transitions required to be >0 for the fit to be considered successful. Three- and four-transition models successfully fit all curves, but five-transition models generally failed to converge. Four-transition models were selected because they yielded a total ΔH of unfolding consistent with expectations from nearest neighbor stacking parameters (200−260 kcal/mol), whereas three-transition models yielded lower estimates (<190 kcal/mol). As detailed in Supporting Information, our analysis is based on the "parsimonious" four-state fit of each individual melting.

$\Delta G_{\mathrm{fold,vH}}$ was computed from the fitted T_m and ΔH_{vH} of the first transition according to

$$\Delta G_{\mathrm{fold,vH}}(300\ \mathrm{K}) = \Delta H_{\mathrm{vH}}(1 - 300\ \mathrm{K}/T_m) \quad (5)$$

Errors of fitted parameters were computed for each melting curve by bootstrap analysis.[31] However, in some cases, the bootstrap estimated errors were smaller than the differences observed between the fits of experimental duplicates. Hence, we report the error as the larger of the two values.

ASSOCIATED CONTENT

Supporting Information

Description of the TOPRNA2 force field; derivation of ΔT_m predictions; mt-tRNA^{Ser} synthesis; supplemental discussion of van't Hoff analysis; Figures S1−S6 and Tables S1 and S2. This

DOI: 10.1021/ja5130308
J. Am. Chem. Soc. 2015, 137, 3592−3599

material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

brookscl@umich.edu

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank D.E. Draper for providing the GMF source code; Yu Chen for help preparing samples; and members of the Al-Hashimi, Brooks, and Fierke laboratories for helpful discussions. A.M.M. is supported by an NSF graduate research fellowship. This work was supported by the National Institutes of Health (P01 GM0066275 to H.M.A., R01 GM55387 to C.A.F., R21 GM09615 to H.M.A. and C.L.B., and R01 GM37554 to C.L.B.).

REFERENCES

(1) Suzuki, T.; Nagao, A. *Annu. Rev. Genet.* **2011**, *45*, 299.
(2) Helm, M.; Brule, H.; Friede, D.; Giege, R.; Putz, D.; Florentz, C. *RNA* **2000**, *6*, 1356.
(3) Giege, R.; Juhling, F.; Putz, J.; Stadler, P.; Sauter, C.; Florentz, C. *Wiley Interdiscip Rev. RNA* **2012**, *3*, 37.
(4) Mustoe, A. M.; Brooks, C. L., III; Al-Hashimi, H. M. *Nucleic Acids Res.* **2014**, *42*, 11792.
(5) Brion, P.; Westhof, E. *Annu. Rev. Biophys. Biomol. Struct.* **1997**, *26*, 113.
(6) Tinoco, I., Jr.; Bustamante, C. *J. Mol. Biol.* **1999**, *293*, 271.
(7) Toompuu, M.; Yasukawa, T.; Suzuki, T.; Hakkinen, T.; Spelbrink, J. N.; Watanabe, K.; Jacobs, H. T. *J. Biol. Chem.* **2002**, *277*, 22240.
(8) Yokogawa, T.; Watanabe, Y.; Kumazawa, Y.; Ueda, T.; Hirao, I.; Miura, K.; Watanabe, K. *Nucleic Acids Res.* **1991**, *19*, 6101.
(9) Theobald-Dietrich, A.; Frugier, M.; Giege, R.; Rudinger-Thirion, J. *Nucleic Acids Res.* **2004**, *32*, 1091.
(10) Gaston, M. A.; Jiang, R.; Krzycki, J. A. *Curr. Opin. Microbiol.* **2011**, *14*, 342.
(11) Juhling, F.; Morl, M.; Hartmann, R. K.; Sprinzl, M.; Stadler, P. F.; Putz, J. *Nucleic Acids Res.* **2009**, *37*, D159.
(12) Marck, C.; Grosjean, H. *RNA* **2002**, *8*, 1189.
(13) Nozawa, K.; O'Donoghue, P.; Gundllapalli, S.; Araiso, Y.; Ishitani, R.; Umehara, T.; Soll, D.; Nureki, O. *Nature* **2009**, *457*, 1163.
(14) Hayashi, I.; Kawai, G.; Watanabe, K. *J. Mol. Biol.* **1998**, *284*, 57.
(15) Watanabe, Y.; Kawai, G.; Yokogawa, T.; Hayashi, N.; Kumazawa, Y.; Ueda, T.; Nishikawa, K.; Hirao, I.; Miura, K.; Watanabe, K. *Nucleic Acids Res.* **1994**, *22*, 5378.
(16) Tiranti, V.; Chariot, P.; Carella, F.; Toscano, A.; Soliveri, P.; Girlanda, P.; Carrara, F.; Fratta, G. M.; Reid, F. M.; Mariotti, C.; Zeviani, M. *Hum. Mol. Genet.* **1995**, *4*, 1421.
(17) Verhoeven, K.; Ensink, R. J.; Tiranti, V.; Huygen, P. L.; Johnson, D. F.; Schatteman, I.; Van Laer, L.; Verstreken, M.; Van de Heyning, P.; Fischel-Ghodsian, N.; Zeviani, M.; Cremers, C. W.; Willems, P. J.; Van Camp, G. *Eur. J. Hum. Genet.* **1999**, *7*, 45.
(18) Jaksch, M.; Klopstock, T.; Kurlemann, G.; Dorner, M.; Hofmann, S.; Kleinle, S.; Hegemann, S.; Weissert, M.; Muller-Hocker, J.; Pongratz, D.; Gerbitz, K. D. *Ann. Neurol.* **1998**, *44*, 635.
(19) Toompuu, M.; Tiranti, V.; Zeviani, M.; Jacobs, H. T. *Hum. Mol. Genet.* **1999**, *8*, 2275.
(20) Toompuu, M.; Levinger, L. L.; Nadal, A.; Gomez, J.; Jacobs, H. T. *Biochem. Biophys. Res. Commun.* **2004**, *322*, 803.
(21) Sussman, J. L.; Holbrook, S. R.; Warrant, R. W.; Church, G. M.; Kim, S. H. *J. Mol. Biol.* **1978**, *123*, 607.
(22) Mustoe, A. M.; Al-Hashimi, H. M.; Brooks, C. L., III. *J. Phys. Chem. B* **2014**, *118*, 2615.
(23) Bailor, M. H.; Mustoe, A. M.; Brooks, C. L., III; Al-Hashimi, H. M. *Nat. Protoc.* **2011**, *6*, 1536.
(24) Bailor, M. H.; Sun, X. Y.; Al-Hashimi, H. M. *Science* **2010**, *327*, 202.
(25) Turner, D. H.; Sugimoto, N.; Freier, S. M. *Annu. Rev. Biophys. Biophys. Chem.* **1988**, *17*, 167.
(26) Xiang, Z.; Soto, C. S.; Honig, B. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 7432.
(27) Onuchic, J. N.; Luthey-Schulten, Z.; Wolynes, P. G. *Annu. Rev. Phys. Chem.* **1997**, *48*, 545.
(28) Putz, J.; Dupuis, B.; Sissler, M.; Florentz, C. *RNA* **2007**, *13*, 1184.
(29) Ambrogelly, A.; Gundllapalli, S.; Herring, S.; Polycarpo, C.; Frauer, C.; Soll, D. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 3141.
(30) Stein, A.; Crothers, D. M. *Biochemistry* **1976**, *15*, 160.
(31) Draper, D. E.; Bukhman, Y. V.; Gluick, T. C. Thermal Methods for the Analysis of RNA Folding Pathways. In *Current Protocols in Nucleic Acid Chemistry*; Wiley: New York, 2000, Chapter 11, Unit 11.3.1.
(32) Wilusz, J. E.; Whipple, J. M.; Phizicky, E. M.; Sharp, P. A. *Science* **2011**, *334*, 817.
(33) Müllers, M.; Maniura-Weber, K.; Kiseljakovic, E.; Bust, M.; Hayrapetyan, A.; Jaksch, M.; Helm, M.; Wiesner, R. J.; von Kleist-Retzow, J. C. *Nucleic Acids Res.* **2005**, *33*, 5647.
(34) Li, X.; Fischel-Ghodsian, N.; Schwartz, F.; Yan, Q.; Friedman, R. A.; Guan, M. X. *Nucleic Acids Res.* **2004**, *32*, 867.
(35) Garcia-Diaz, M.; Kunkel, T. A. *Trends Biochem. Sci.* **2006**, *31*, 206.
(36) McFarland, R.; Elson, J. L.; Taylor, R. W.; Howell, N.; Turnbull, D. M. *Trends Genet.* **2004**, *20*, 591.
(37) Butcher, S. E.; Pyle, A. M. *Acc. Chem. Res.* **2011**, *44*, 1302.
(38) Lipfert, J.; Doniach, S.; Das, R.; Herschlag, D. *Annu. Rev. Biochem.* **2014**, *83*, 813.
(39) Chu, V. B.; Lipfert, J.; Bai, Y.; Pande, V. S.; Doniach, S.; Herschlag, D. *RNA* **2009**, *15*, 2195.
(40) Laing, C.; Schlick, T. *J. Mol. Biol.* **2009**, *390*, 547.
(41) Lescoute, A.; Westhof, E. *RNA* **2006**, *12*, 83.
(42) Sim, A. Y. L.; Levitt, M. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 3590.
(43) Hajdin, C. E.; Ding, F.; Dokholyan, N. V.; Weeks, K. M. *RNA* **2010**, *16*, 1340.
(44) Walter, N. G.; Burke, J. M.; Millar, D. P. *Nat. Struct. Biol.* **1999**, *6*, 544.
(45) Zhao, Z. Y.; Wilson, T. J.; Maxwell, K.; Lilley, D. M. *RNA* **2000**, *6*, 1833.
(46) Tan, E.; Wilson, T. J.; Nahas, M. K.; Clegg, R. M.; Lilley, D. M.; Ha, T. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 9308.
(47) Klostermeier, D.; Millar, D. P. *Biochemistry* **2000**, *39*, 12970.
(48) Brooks, B. R.; Brooks, C. L., III; Mackerell, A. D., Jr.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. *J. Comput. Chem.* **2009**, *30*, 1545.
(49) Feig, M.; Karanicolas, J.; Brooks, C. L., III. *J. Mol. Graphics Modell.* **2004**, *22*, 377.
(50) Tyagi, R.; Mathews, D. H. *RNA* **2007**, *13*, 939.

DOI: 10.1021/ja5130308
J. Am. Chem. Soc. 2015, 137, 3592−3599