J|A|C|S COMMUNICATIONS

Published on Web 11/17/2009

Urea Destabilizes RNA by Forming Stacking Interactions and Multiple Hydrogen Bonds with Nucleic Acid Bases

U. Deva Priyakumar,[†] Changbong Hyeon,[‡] D. Thirumalai,*[,§] and Alexander D. MacKerell, Jr.*[,†]

Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, Baltimore, Maryland 21201, Department of Chemistry, Chung-Ang University, Seoul 156-756, Republic of Korea, and Biophysics Program, Institute for Physical Science and Technology, University of Maryland, College Park, Maryland 20742

Received July 13, 2009; E-mail: thirum@umd.edu; amackere@rx.umaryland.edu

Urea has long been used to probe the stability and folding kinetics of proteins.[1] In contrast only recently it was shown that the RNA molecules that have a high propensity to misfold can be resolved using moderate amounts of urea.[2] Urea titrations can also be used to probe the interactions that stabilize the folded states of RNA.[2c] Although the mechanism by which urea denatures proteins is now fairly well understood,[3] the nature of interactions by which urea destabilizes RNA is not known. To provide a microscopic basis for the action of urea on RNA we have carried out extensive all atom molecular dynamics (MD) simulations on two RNA constructs using two urea force fields. Destabilization of RNA is due to disruption of base-pair interactions by direct hydrogen bonding of urea with the bases. The simulations also reveal a novel mechanism in which urea molecules engage in stacking interactions with the purine bases.[4]

Analyses of 20 ns trajectories generated using MD simulations with a urea force field that was created as a part of the present work (see Supporting Information (SI) for simulation details, SI Figures 1 and 2 and Tables 1−3 for urea parameter development and for assessing the validity of the force field) of the 22-nucleotide RNA hairpin P5GA[5] (Figure 1A) in various urea concentrations ([C]s) reveal that at high [C] the solvent-exposed stem regions lead to disruption of base pairing. The fraction of intact hydrogen bonds associated with the bases in the stem decreases from ∼0.71 in the absence of urea to 0.46 in 8 M urea. The loss of the Watson−Crick (WC) hydrogen bonds is accompanied by opening of the base pairs, which is reflected in the distribution of the hydrogen bond donor−acceptor distances (R_{HB}) in the hairpin stem (Figure 1B). The base-paired state is indicated by a sharp peak at $R_{HB} = 3$ Å, whose height decreases as [C] increases to 6 M. The probability of sampling R_{HB} distances that are greater than 10 Å (Figure 1B) increases greatly in high [C], which results in a rotation of the bases of the helix leading to N1−N3 distances of ∼16 Å.[6] Examination of opening at the individual base pair level reveals considerable heterogeneity[7] with the largest fluctuations occurring at the GA and GU mismatches. We also show that urea-induced disruption of the base opening due to the loss of WC hydrogen bonds is nonspecific in the sense that urea does not preferentially interact with a specific base pair. These findings suggest that denaturation of RNA is due to favorable nonspecific interactions with amide-like surfaces of the nucleic acids. The average base−base interaction energies (GC, AU, AG, and GU) decrease substantially at high [C] (SI Table 4). When averaged over all base pair interactions in the stem, the interactions become less favorable by ∼2.7 kcal/mol at 6 M relative to [C] = 0 (SI Table 4). The average interaction energies

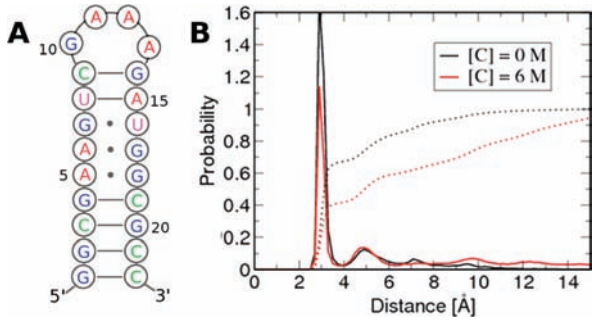

Figure 1. Effect of urea on the P5GA hairpin. (A) Secondary structure map of P5GA. (B) Probability distribution of the N1−N3, N1−N1, and N1−O2 interatomic distances of the GC(AU), GA, and GU base pairs, respectively, in the stem. The dotted line is the integrated probability over the distances.

for certain base pairs (for example A6G17 and U8A15) are substantially less at high [C] relative to their values in water (see SI Table 4).

In contrast, the backbone conformational properties in the presence of urea are unperturbed, which is reflected in the distribution functions characterizing the phophodiester linkages. The angle distributions for α, β, γ, χ, and δ do not depend significantly on the urea concentration (Figure 2 and SI Figure 3). There are minor changes in the distributions of ε and ζ change as [C] increases to 6 or 8 M (Figure 2). The small peak in the ζ distribution at 30° (Figure 2) corresponds to the opening of the bases, which is in accord with previous studies that probed the base-flipping dynamics.[8] The [C]-dependent distributions of the ζ angle for each nucleotide show that the peak at 30° is also sampled by the GAAA tetra-loop (Figure 1A). Taken together the results in Figure 2 and SI Figure 3 show that urea does not induce structural changes in the RNA backbone. The minor perturbations in the distributions of the χ, ε, and ζ angles are merely a consequence of the opening of the bases.

To provide a molecular picture of urea−RNA interactions we introduce the dehydration ratio $\lambda_{DR} = \Delta N_W/N_U$, where ΔN_W is the difference in the number of water molecules in the first solvation shell of RNA as [C] increases from 0 and N_U is the number of urea molecules in the first solvation shell at [C] (SI Table 5). A value of $\lambda_{DR} > 1$ implies that more than one water molecule is exchanged for each urea. The values of λ_{DR} change from 2.54 at 1 M urea to ∼0.85 at [C] = 8 M. The decrease in λ_{DR} at higher [C] is because the number of water molecules ceases to decrease while the number of urea molecules in the first solvation shell increases. The value of λ_{DR} around the phosphodiester backbone is approximately independent of [C] (SI Table 5), which further indicates that the primary disruption of RNA structure is due to interactions of urea with the bases.

[†] University of Maryland, Baltimore.
[‡] Chung-Ang University.
[§] University of Maryland, College Park.

10.1021/ja905795v

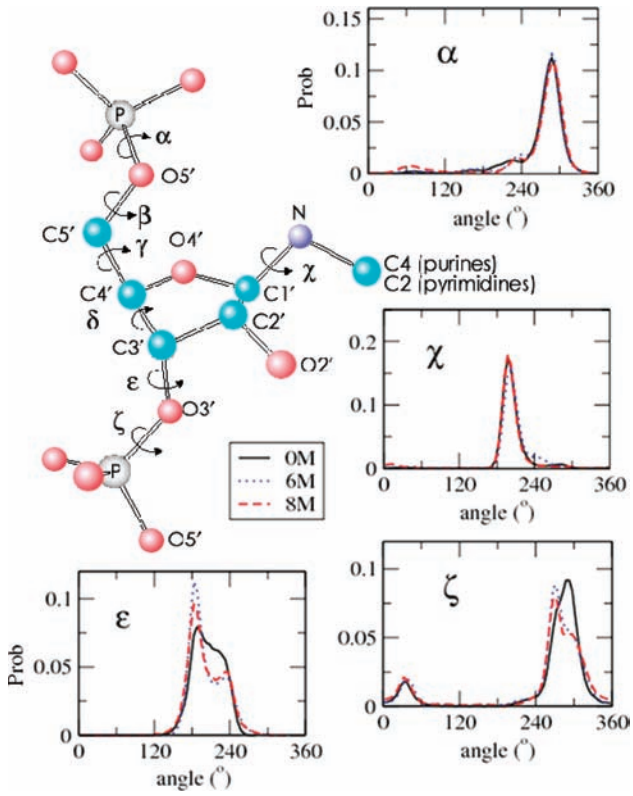

Figure 2. Probability distributions of the dihedral angles along phosphodiester backbone of the RNA hairpin at [C] = 0, 6, 8 M.

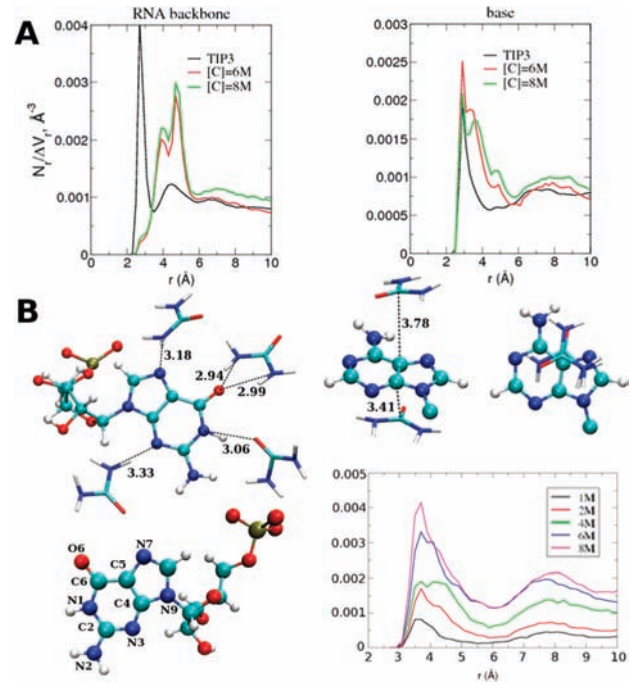

Figure 3. (A) RDFs of OU around the RNA nitrogen atoms at 6 and 8 M urea. The water RDFs are scaled by 5 and 10 for the urea oxygen plots. Results are shown for the RNA backbone atoms (phosphodiester and sugar oxygen) and RNA bases. (B) Structure with multiple hydrogen bonds between urea and RNA base and phosphate group. The panels on the right show the structure of urea base stacking and the corresponding RDFs between the urea carbon and C4, C5, C2 (A and G) and C5, C6 (C and U) atoms. The contributions from individual atoms are in SI Figure 13.

The [C]-dependent values of λ_{DR} also suggests that urea can engage in multiple interactions with the nucleic acid bases, which are reflected in the radial distribution functions (RDF) (Figure 3A). There is an asymmetry in the interaction of N_U and O_U atoms of urea with RNA; N_U atoms compete with water for direct hydrogen bonding interactions with both the bases and the backbone of the RNA as evident from a sharp peak in the RDF around 3 Å at all [C] (SI Figure 4). The urea oxygen RDFs exhibit three distinct peaks at approximately 3, 4, and 5 Å (Figure 3A). Surprisingly, the peak at 3 Å, which corresponds to the direct interactions with the RNA, is absent at [C] $\neq$ 0 indicating that there are only a few direct interactions with the hydroxyl group of the ribose moiety and almost all the direct interactions occur with the bases. The additional peaks at approximately 4 and 5 Å correspond to oxygen atoms in urea molecules that indirectly interact with the RNA via the urea nitrogen atom. Representative examples of common hydrogen bonding interactions of urea with both the bases and backbone of the RNA show (Figure 3B) that N_U donates a hydrogen bond to N7 of a guanine base, with the distance between the hydrogen bond acceptor and the O_U $\approx$ 4 Å corresponding to the second peak of the RDFs of the oxygen (Figure 3A). Multiple urea−RNA interactions, leading to $\lambda_{DR} > 1$, include N7 and O6 of a single guanine base hydrogen bonding simultaneously the N_U atoms (Figure 3B), with the distances between the hydrogen bond acceptors and O_U being ~5 Å, corresponding to the third peak in the RDFs (Figure 3A). Remarkably, urea participates in stacking interactions with the bases (right panels in Figure 3B), which further contributes to the destabilization of the folded RNA. Two urea molecules are positioned parallel to the purine base and the approximate interplanar distances are ~3.5 Å, which is comparable to the distance between the two rings in a benzene dimer that are stacked parallel to each other.[9]

Because of the limitations in the sampling of the conformational space of the P5GA hairpin, we also simulated a smaller RNA duplex made of four complementary base pairs at varying [C] (Figure 4). To establish the robustness of the denaturation mechanism we used a different urea force field (see SI). Since the base pair distances (r_i) are subject to significant fluctuations when the base pairs are not formed, the inverse distance X_i ($= r_i^{-1}$) can be used to better visualize the equilibrium dynamics of the RNA duplex (Figure 4A). From 300 ns trajectories at each [C] (see Figure 4A and SI Figure 5), we calculated the fraction of bound duplex using $f_B = \tau_B/(\tau_B + \tau_U)$ and the change in free energy for (bound)⇔(unbound), $\Delta G[C]/k_BT = \log[(1 - f_B)/f_B] = \log(\tau_U/\tau_B)$. The dwell time in the bound state τ_B satisfies $\Sigma_{i=1}^{4}\,\Theta[X_c - X_i(t)] \neq 0$ and τ_U is the time for $\Sigma_{i=1}^{4}\,\Theta[X_c - X_i(t)] = 0$; $\Theta(x)$ is the Heaviside step function, and $X_c = 0.2(\text{Å})^{-1}$. For the GC pair, X_i is the inverse distance between H1 and N3 atoms in G and C, respectively, while, for GU, X_i is taken between H1 in G and O2 in U (see SI Figure 6). The decrease in f_B, relative to its value in water, as [C] increases (Figure 4B), quantitatively demonstrates the destabilizing effect of urea on the RNA duplex. Because of the small size of the RNA the bound state is unstable even at [C] = 0 M, in accord with an estimated melting temperature in the range 18−35 °C.[10]

Just as in proteins the free energy difference between the bound and the denatured states of the duplex RNA varies as $\Delta G[C]/k_BT = \Delta G[0] + m[C]$ (Figure 4B). The value of m, which is the slope of the aqua line in Figure 4B, is $\approx$ 0.21 kcal/mol·M. It is known that the m-value is a function of the RNA length and ion concentration and depends sensitively on the valence of the counterion. Taking these factors into account, we find that the m-value obtained for the first time using simulations is in reasonable agreement with measurements on small duplexes.[2c,11]

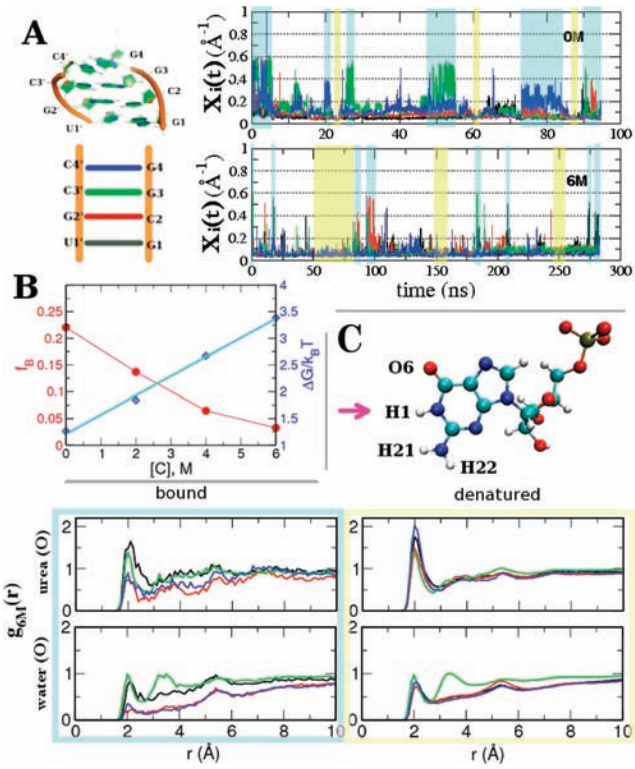

Figure 4. Urea-induced structural transitions. (A) The structure of RNA duplex (left). Inverse N1−N3 distances of the four base pairs (color code: blue, green, red and black are used consistently) as a function of time at [C] = 0 M (top) and [C] = 6 M (bottom). (B) Fraction of bound f_B and change in free energy ΔG as a function of urea concentrations. The fit is made for *m*-value analysis. (C) RDFs of urea oxygen and water oxygen with respect to H1 atom of the G base when the duplex is bound or denatured at [C] = 6 M.

The pair correlation functions involving water around the bound duplex were calculated using only those conformations that satisfy $\Sigma_{i=1}^{4}\,\Theta[X_c - X_i(t)] = 4$ (aqua shadow in Figure 4A). The $g(r)$ for the denatured state is calculated using the time traces in which the RNA duplex is fully denatured, i.e., $\Sigma_i\,\Theta[X_c - X_i(t)] = 0$ (yellow shadow in Figure 4A). The RDFs of O_U and water oxygen relative to various atoms of the nucleic acid at varying urea concentrations (SI Figures 7−13) lead to a number of interesting conclusions: (i) Near the base, the density of water is below the bulk density ($g(r\rightarrow\infty) \approx 1$). As a consequence of the hydrophobic nature of the base, the water distribution around the hydrogen atoms at H1 (Figure 4C and SI Figure 12) and H21 or H22 of the amide group of G (SI Figure 11) is $g(r) < 1$ regardless of the state of the RNA duplex. (ii) When paired bases are disrupted, O_U forms a hydrogen bond with the H1 atom in G (Figure 4C and SI Figure 7), which is reflected in the increase of the RDF peak of urea at $r = 2$ Å. (iii) The $g(r)$ values of urea or water around H21 or H22 (SI Figures 8 and 11), O6 oxygen in guanine base (SI Figure 9), and OP1 or OP2 in the phosphate group (SI Figure 10) are similar between bound and denatured forms. Thus, the disruption of the central hydrogen bond involving H1 of G, which is replaced by hydrogen bonds involving O_U, is the key event for the RNA denaturation. (iv) Comparison of $g(r)$ functions in SI Figures 8 and 11 shows in a dramatic fashion the depletion of water around the bases. More importantly, the ability of O_U to form multiple hydrogen bonds is vividly illustrated (see Figure 8 in the SI). (v) Stacking interactions with urea are reflected in the various RDFs (see Figure 3 and SI).

Both sets of simulations show that destabilization of RNA is due to disruption of base-pair interactions by direct multiple hydrogen bonding with the bases and formation of stacking interactions with the bases. In contrast to proteins, a multitude of favorable interactions largely involving the solvent-exposed bases lead to urea-induced destabilization of the structured RNA. In particular, there is no analogue of the stacking interactions involving urea in proteins, though stacking interactions in GdnHCl have been observed.[13] Finally, the proposed mechanism readily explains the observations[2] that urea-induced destabilization of base pair interactions in misfolded RNA molecules can increase the folding rates, thus acting as "chemical" chaperones.

Acknowledgment. This work was supported in part by grants from the National Science Foundation (CHE 09-10433), the NIH (GM51501), and the National Research Foundation of Korea (NRF) (R01-2008-000-10920-0, KRF-C00142, KRF-C00180, and 2009-0093817).

Supporting Information Available: Simulation methods and figures. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) (a) Fersht, A. *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding*; W. H. Freeman Company: New York, 1998. (b) Tanford, C. *J. Am. Chem. Soc.* **1964**, *86*, 2050–2059. (c) Makhatadze, G. I.; Privalov, P. L. *J. Mol. Biol.* **1992**, *226*, 491–505. (d) Myers, J. K.; Pace, C. N.; Scholtz, J. M. *Protein Sci.* **1995**, *4*, 2138–2148.
(2) (a) Pan, J.; Thirumalai, D.; Woodson, S. A. *J. Mol. Biol.* **1997**, *273*, 7–13. (b) Rook, M. S.; Treiber, D. K.; Williamson, J. R. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *281*, 609–620. (c) Shelton, V. M.; Sosnick, T. R.; Pan, T. *Biochemistry* **1999**, *38*, 16831–16839. (d) Sclavi, B.; Woodson, S. A.; Sullivan, M.; Chance, M. R.; Brenowitz, M. *J. Mol. Biol.* **1997**, *266*, 144–159.
(3) (a) Robinson, D. R.; Jencks, W. P. *J. Am. Chem. Soc.* **1965**, *87*, 2462–2469. (b) Wallqvist, A.; Covell, D. G.; Thirumalai, D. *J. Am. Chem. Soc.* **1998**, *120*, 427–428. (c) Vanzi, F.; Madan, B.; Sharp, K. *J. Am. Chem. Soc.* **1998**, *120*, 10748–10753. (d) Tobi, D.; Elber, R.; Thirumalai, D. *Biopolymers* **2003**, *68*, 359–369. (e) Soper, A. K.; Castner, E. W.; Luzar, A. *Biophys. Chem.* **2003**, *105*, 649–666. (f) Bennion, B. J.; Daggett, V. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 5142–5147. (g) O'Brien, E. P.; Dima, R. I.; Brooks, B.; Thirumalai, D. *J. Am. Chem. Soc.* **2007**, *129*, 7346–7353. (h) Hua, L.; Zhou, R.; Thirumalai, D.; Berne, B. J. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 16928–16933. (i) England, J. R.; Pande, V. S.; Haran, G. *J. Chem. Soc.* **2008**, *130*, 11854–11855.
(4) Mason, P. E.; Neilson, G. W.; Enderby, J. E.; Saboungi, M. -L.; Dempsey, C. E.; MacKerell, A. D.; Brady, J. W. *J. Am. Chem. Soc.* **2004**, *11*, 462–11470, 126.
(5) Rudisser, S.; Tinoco, I., Jr. *J. Mol. Biol.* **2000**, *295*, 1211–1223.
(6) Huang, N.; Banavali, N. K.; MacKerell, A. D., Jr. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 68–73.
(7) Pan, Y. P.; Priyakumar, D.; MacKerell, A. D., Jr. *Biochemistry* **2005**, *44*, 1433–1443.
(8) (a) Banavali, N. K.; MacKerell, A. D., Jr. *J. Mol. Biol.* **2002**, *319*, 141–160. (b) Banavali, N. K.; Huang, N.; MacKerell, A. D., Jr. *J. Phys. Chem. B* **2006**, *110*, 10997–11004.
(9) Sinnokrot, M. O.; Valeev, E. F.; Sherrill, C. D. *J. Am. Chem. Soc.* **2002**, *124*, 10887–10893.
(10) Xia, T. et al. *Biochemistry* **1998**, *37*, 14719–14735.
(11) (a) Scholtz, J. M.; Barrick, D.; York, E. J.; Stewart, J. M.; Baldwin, R. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 185–189. (b) Pace, C. N.; Shaw, K. L. *Proteins: Struct., Funct., Genet.* **2000**, *Suppl 4*, 1–7.
(12) (a) Shelton, V. M.; Sosnick, T. R.; Pan, T. *Biochemistry* **2001**, *40*, 3629–3638. (b) Lambert, D.; Draper, D. E. *J. Mol. Biol.* **2007**, 993–1005.
(13) Mason, P. E.; Brady, J. W.; Neilson, G. W.; Dempsey, C. E., *Biophys. J.* **2007**, *93*, L04–L06.

JA905795V