THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

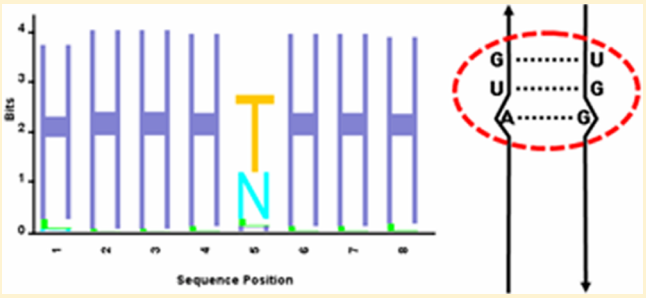
Structural Variations of Single and Tandem Mismatches in RNA Duplexes: A Joint MD Simulation and Crystal Structure Database Analysis

Sukanya Halder[†] and Dhananjay Bhattacharyya*[,‡]

[†]Biophysics Division and [‡]Computational Science Division, Saha Institute of Nuclear Physics, Kolkata, West Bengal, 700 064, India

Supporting Information

ABSTRACT: Internal loops within RNA duplex regions are formed by single or tandem basepairing mismatches with flanking canonical Watson−Crick basepairs on both sides. They are the most common motif observed in RNA secondary structures and play integral functional and structural roles. In this report, we have studied the structural features of 1 × 1, 2 × 2, and 3 × 3 internal loops using all-atom molecular dynamics (MD) simulation technique with explicit solvent model. As MD simulation is intricately dependent on the choice of force-field and these are often rather approximate, we have used both the most popular force-fields for nucleic acids—CHARMM27 and AMBER94—for a comparative analysis. We find that tandem noncanonical basepairs forming 2 × 2 and 3 × 3 internal loops are considerably more stable than the single mismatches forming 1 × 1 internal loops, irrespective of the force field. We have also analyzed crystal structure database to study the conservation of these helical fragments in the corresponding sets of RNA structures. We observe that the nature of stability in MD simulations mimic their fluctuating natures in crystal data sets also, probably indicating reliable natures of both the force fields to reproduce experimental results. We also notice significant structural changes in the wobble G:U basepairs present in these double helical stretches, leading to a biphasic stability for these wobble pairs to release the deformational strains introduced by internal loops within duplex regions.



INTRODUCTION

The RNA macromolecules exhibit extreme diversity in their secondary structures as well as their roles and functions in biological systems. The three-dimensional architecture of folded RNA comprises of canonical double helices interspersed with noncanonical structural regions like hairpin loops, bulges, internal loops, multibranch loops and coaxial stacks. Internal loops are formed with single or tandem basepairing mismatches present within double helical stretches formed with canonical Watson−Crick basepairs. Single mismatches or the (1 × 1) internal loops are the most common motif observed in RNA secondary structures[1] and these play integral functional and structural roles.[2,3] There can be numerous variations of single mismatches depending on the types of pairing bases along with the basepairing patterns and geometries. Nucleic acid bases have the ability to use either of the three hydrogen bonding edges [Watson−Crick, Hoogsteen, and sugar-edges] for forming the mismatched basepairs.[4] Considering this, an A:U or a G:C basepair can also be designated as a mismatch if the bases are hydrogen-bonded following any of the various non-Watson−Crick geometry, e.g.—A:U H:WT, G:C W:SC *etc.* [basepairing nomenclature following Das et al[5]]. A detailed structural classification of single mismatches into 12 geometric families has been done by Leontis and Westhof.[6] We have previously studied the stabilities of (2 × 2) and (3 × 3) internal loops using all-atom molecular dynamics simulations.[7] Here, we focus into the structural features of two different single mismatches from the *cis* Watson−Crick/Watson−Crick geometric family[6] [W:WC]—A:G W:WC as purine−purine mismatch and U:U W:WC as pyrimidine−pyrimidine mismatch. Importance of some of these mismatches was studied earlier by Znosko and co-workers but the (3 × 3) mismatch was not studied.[8−10]

The A:G W:WC is one of the many basepairing schemes that have so far been studied between two purine bases. Former studies report large values of propeller twist and highly buckled structures for A:G basepairs,[11] as its planarity involves stretching of the basepairing strands in opposite direction. It is observed that this basepair, or more commonly known imino form, is rarely found within double helical stretches of RNA. A crystal structure database analysis finds that, among the 14 different basepairing geometries possible for a A:G basepair, only one type is located near the central region of a double helix in 16S rRNA of *Thermus thermophilus*, and none from the other available structures of tRNAs, ribozymes, SRP RNAs *etc.*[11] However, this basepair is quite frequent in other local environments as has been shown by analysis of nonredundant crystal structure database.[12,13] Similarly, many thermodynamic

Received: June 8, 2012
Revised: September 3, 2012
Published: September 6, 2012

Table 1. Description of the Double Helices Containing Single and Tandem Non-Canonical Basepairs

mismatch	PDB	basepair	sequence	residue number[a]	identifier
R:R	1FJG	A:G W:WC	5′-CGCC**A**UGG-3′	1387(1) − 1394(8) [5′ → 3′]	1FJA_CHARMM
			3′-GCGG**G**GUC-5′	1457(9) − 1464(16) [5′ → 3′]	1FJG_AMBER
Y:Y	1J5A	U:U W:WC	5′-GUC**U**GGC-3′	1999(1) − 2005(7) [5′ → 3′]	1J5A_CHARMM
			3′-CGG**U**UCG-5′	2109(8) − 2115(14) [5′ → 3′]	1J5A_AMBER
G:U	1N33	G:U W:WC	5′-GGGC**U**CUACCC-3′	1399(1) − 1409(11) [5′ → 3′]	1N33_CHARMM
			3′-CCCG**G**GAUGGG-5′	1442(12) − 1452(22) [5′ → 3′]	1N33_AMBER
2 × 2	2AW4	G:A S:HT::A:G H:ST	5′-GUGG**GA**GCACG-3′	533(1) − 543(11) [5′ → 3′]	2AW4_CHARMM
			3′-CGUC**AG**UGUGC-5′	550(12) − 560(22) [5′ → 3′]	2AW4_AMBER
3 × 3	1N32	A:A s:hT::A:U H:WT::A:G H:ST	5′-GC**AAA**CCGG-3′	778(1) − 786(9) [5′ → 3′]	1N32_CHARMM
			3′-UG**AUG**GGCC-5′	796(10) − 804(18) [5′ → 3′]	1N32_AMBER

[a]Modified residue numberings as represented in the manuscript are given in parentheses.

studies on uracil−uracil basepairing in RNA duplex have suggested their flexibility and instability.[14,15] Singly occurring U:U W:WC basepair is shown to have destabilizing effect on RNA oligonucleotides owing to their poor stacking ability with the adjacent basepairs.[16−18] They require flexibility in local environment, like helix-termini or the presence of other noncanonical basepairs, for enhanced stability.[15]

Among the various computational approaches based on experimental data, molecular dynamics simulation techniques have widely been established as a useful tool to elucidate biomolecular structural variability. However, MD simulations have been questioned on two significant aspects—limited sampling due to limitation in computational resources and the approximate nature of nonpolarizable and transferable force-fields. Among the existing force-fields, AMBER,[19] and CHARMM27[20] are the most extensively tested ones and are widely used for studying nucleic acids. However, there have been many reports about the shortcomings of both the force-fields based on DNA and RNA simulations.[21−25] AMBER force-field was speculated to generally prefer the B-form, while CHARMM had a tendency to maintain A-DNA structures in their original forms.[23] It is also to be mentioned that simulations of RNA are more problematic than that for DNA and proteins because of its structural variability. RNA can form complicated structural organizations leaving the canonical double helical geometry, while the force-fields were originally designed and tested for regular helices with Watson−Crick basepairs.

We have carried out all-atom explicit solvent model MD simulations of four double helical segments extracted from rRNA crystal structures, including two helices from our previous report,[7] using both AMBER and CHARMM force fields for 50 ns. We have selected the helices containing single mismatch considering larger length of the double helices and location of the mismatches near central region of the helices. We notice higher stabilities of the tandem noncanonical basepairs under both the force fields whereas the single noncanonical basepairs show tendency to alter significantly. We observe significant variation in the structures of wobble G:U basepairs present in these helices, which are known to be nearly isosteric with the usual Watson−Crick A:U or G:C basepairs and thus, are very frequently observed in RNA double helix structures. To elucidate the importance of structural fluctuation of G:U basepairs, we have also performed 50 ns long MD simulations of a helical stretch containing a single wobble G:U basepair in the central region. We have further analyzed all the crystal structures of the class of RNA from which the initial structures are taken.[13] We observe that the tandem non-canonical basepairs are more conserved in the corresponding set of similar crystal structures also while the single non-canonical ones are found to be less conserved. This gives us the confidence that limited sampling by the approximate force-fields also can give reliable results.

MATERIALS AND METHODS

1. Initial Structures. Coordinates of the double helices have been obtained from Protein Data Bank [PDB[26]]. We have picked up helices containing single mismatches of the purine−purine [A:G W:WC] and pyrimidine−pyrimidine [U:U W:WC] types as shown in Table 1. The other two helices, containing successive noncanonical basepairs in the middle, are taken from our previous study[7] [Table 1] − (A:A s:hT)::(A:U H:WT)::(A:G H:ST) in 1N32 and (G:A S:HT)::(A:G H:ST) in 2AW4 [":" represents basepairing and "::" represents stacking] (schematic diagrams of the mismatches are given in Figure 1). The helix containing wobble G:U basepair is taken from PDB-ID 1N33. All the helices are renumbered starting from "1" for simplicity of representation [see Table 1].

2. MD Simulations. In all the simulations, we have used the CHARMM27[20] and AMBER94[19] force-fields to describe the biomolecular interactions. Although different groups have put significant effort in improving the AMBER force-field recently and quite a few variants of this force-field are available now, we have used AMBER94 as the newer versions essentially address anomalies of DNA structural variations. Each of the double-helical stretches have been solvated in an orthorhombic water-box containing TIP3P water molecules in such a way that there remains enough solvent layer [~15 Å] surrounding the double helices in all directions to reduce interaction between the solute molecules in adjacent periodic cells. The periodic boxes also contain Na^+ ions to maintain electroneutrality of the whole systems. The positions of these counterions have been generated by Monte Carlo simulation in absence of water, only considering interaction between the ions and the double helix in concern.[27] These solvated systems have been subjected to minimizations and, finally, MD simulations under no force-constraint using periodic boundary condition. We have used particle mesh Ewald [PME[28]] method for treating the long-range electrostatic interactions with a kappa value of 0.35. A force-switch method has been applied for nonbonded interactions with 12 Å cutoff. Energy minimizations have been performed followed by slow heating of the systems from 0 to 300 K in 30 ps. All these calculations were performed by CHARMM c30b1[29] and sander module of Amber 8.[30] Constant pressure−temperature MD simulations [CPT Dynamics] were carried out by NAMD[31,32] software using

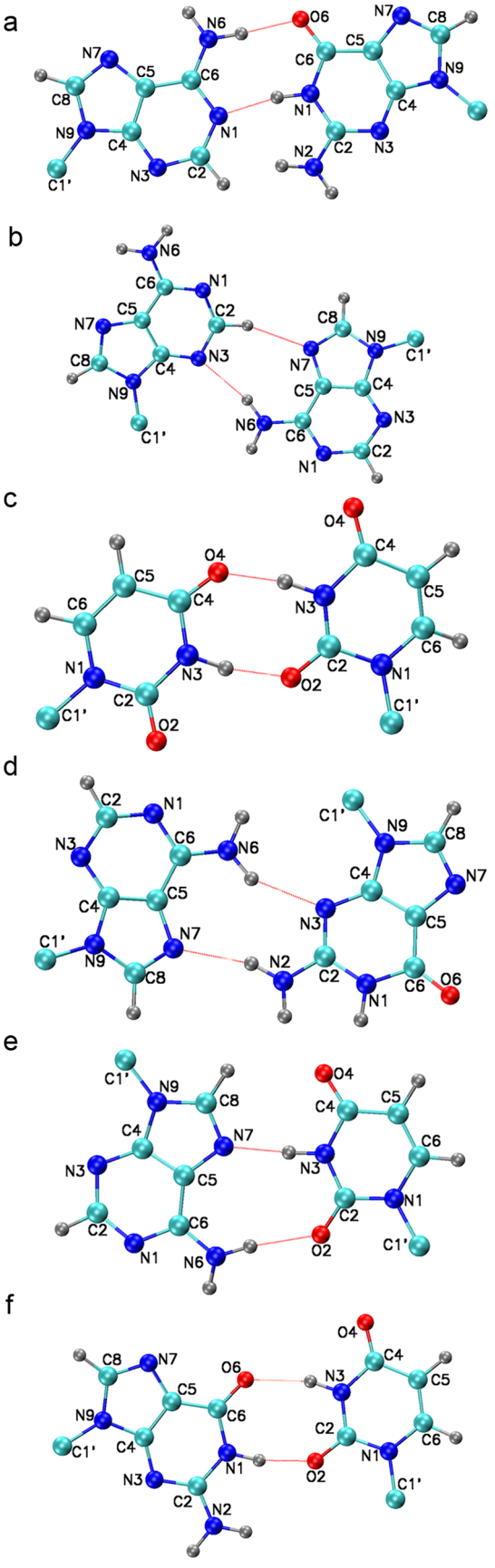


Figure 1. Schematic diagrams of noncanonical basepairs (a) A:G W:WC [1FJG], (b) U:U W:WC [1J5A], (c) A:G H:ST [1N32 and 2AW4], (d) A:A s:hT [1N32], (e) A:U H:WT[1N32], and (f) G:U W:WC [1N33].

extended system algorithm without coupling with 1 fs integration time step. Translational and rotational movements of the center of mass have been removed at an interval of 5 ps. Production runs of 50 ns have been carried out at 300 K and 1 atm pressure [piston mass of 400 amu]. We have also extended our previously reported simulations on (2 × 2) and (3 × 3) internal loops.[7] The helical stretches from PDB entries 1N32 and 2AW4 have been simulated further up to 50 ns under both AMBER94 and CHARMM27 force-fields, in order to compare the stabilities of single mismatches with tandem noncanonical basepairs.

Trajectory analyses have been performed using CHARMM and the intra- and interbasepair parameters have been calculated by NUPARM[33,34] software at an interval of 1 ps. The basepairing patterns during the MD run are determined by BPFind software,[5] with a cutoff distance of 3.8 Å and cutoff angle of 120°. Basepair overlap values are calculated by 3DNA software.[35] Solvent-mediated interactions have been determined by calculation of residence times for the solvent molecules, monitoring the uninterrupted time-span of trajectories through which a solvent molecule remains within 4 Å of two specified atoms.

3. Crystal Structure Analyses. We have analyzed the crystal structure database for similar double helical stretches containing the single mismatches under study. The helices containing A:G mismatch [taken from PDB entry 1FJG] and wobble G:U basepair [taken from PDB entry 1N33] are obtained from crystal structures of 16S subunit of rRNA from *T. thermophilus*, whereas the helix containing U:U mismatch is extracted from PDB entry 1J5A, which is a 23S subunit of *Deinococcus radiodurans* rRNA. We have picked up all the structures of *T. thermophilus* 16S rRNA and *D. radiodurans* 23S rRNA to analyze the structural features of these single mismatches from our HD-RNAS database.[13] We have also analyzed the crystallographic ensembles for 1N32 and 2AW4 helices[7] containing (2 × 2) and (3 × 3) noncanonical mismatches. The helical stretches from each group have been compared based on stability of the basepairs as well as structural parameters of the mismatches. The basepair propensity plots have been generated using the *seqlogo*[36] tool of MATLAB (R2009a).

RESULTS AND DISCUSSION

1. Overall Simulation Behavior. We have analyzed 50 ns long trajectories from all five simulations and found that overall behaviors displayed by the solvated helical systems are essentially same in both force-fields; though there remain some subtle differences between them. As have been formerly observed for RNA simulations, the trajectories reach equilibrium phase at an early stage of simulation. Fluctuations of the entire helices are found to be more in CHARMM27 than AMBER94 force-field. We also find that terminal basepairs of all the double helices show opening after a certain period of MD run under CHARMM force-field even at physiological temperature, sometimes with reannealing. However, no such melting is observed in simulations by AMBER, and terminal basepairs remain stable throughout the run. This force-field dependent behavior is in contrary to the previous results shown by different groups for both DNA[37] and RNA.[25,38]

We have analyzed fluctuations of the sugar−phosphate backbone and base moieties of each nucleotide separately. We find from the rmsf plots [Figure 2] that some bases, apart from those situated in terminal regions, have higher fluctuations

dx.doi.org/10.1021/jp305628v | J. Phys. Chem. B 2012, 116, 11845−11856

Table 2. Non-Canonical Basepairs[a]

	buckle	open	propeller	stagger	shear	stretch
5A:12G (1FJG)	4.8(10.9)	39.4(34.2)	−4.7(10.2)	0.1(0.9)	−5.0(2.5)	0.2(3.2)
	11.4(12.1)	67.0(28.1)	−10.3(10.7)	0.5(1.1)	−3.0(1.5)	3.4(0.7)
	−2.6(6.6)	9.8(10.4)	−1.0(12.2)	−0.4(0.9)	−0.1(0.7)	2.8(0.4)
4U:11U (1J5A)	−4.7(13.2)	−66.9(45.8)	−11.6(18.8)	0.2(2.8)	1.0(3.7)	4.0(2.4)
	4.6(10.0)	−6.4(11.9)	−19.3(8.1)	−0.1(0.4)	2.5(0.4)	3.0(0.2)
	7.3(9.4)	−0.7(6.7)	−17.3(9.0)	0.2(0.7)	1.6(0.3)	3.3(0.7)
5U:18G (1N33)	−2.0(9.5)	18.5(26.0)	−9.6(9.0)	−0.1(0.6)	2.8(1.2)	3.6(1.1)
	0.5(9.2)	−1.9(5.3)	−13.1(7.7)	−0.0(0.4)	2.5(0.3)	2.9(0.1)
	1.1(7.5)	2.4(8.3)	−3.0(7.2)	−0.2(0.7)	1.6(0.9)	2.9(0.7)
5G:18A (2AW4)	−2.5(8.3)	−7.2(5.4)	3.2(7.7)	−0.2(0.4)	2.2(0.3)	3.4(0.2)
	−4.1(8.8)	−7.4(6.7)	1.8(7.8)	−0.2(0.4)	2.2(0.3)	3.3(0.2)
	−3.5(5.4)	−12.2(2.6)	−1.1(3.9)	−0.1(0.3)	2.4(0.2)	3.3(0.1)
6A:17G (2AW4)	−5.7(9.0)	6.0(6.6)	6.0(6.5)	−0.2(0.4)	2.0(0.4)	3.3(0.2)
	−2.5(9.1)	9.9(6.0)	5.7(7.1)	−0.3(0.4)	2.3(0.3)	3.3(0.1)
	0.8(5.5)	15.1(4.1)	7.5(7.1)	0.2(0.4)	2.6(0.2)	3.3(0.1)
3A:16A (1N32)	−17.4(8.0)	1.5(4.3)	5.3(5.9)	0.5(0.4)	2.3(0.2)	2.9(0.2)
	−13.2(9.3)	−2.6(3.9)	3.6(6.7)	0.3(0.4)	2.5(0.2)	3.0(0.2)
	−24.2(8.8)	0.7(3.9)	5.5(4.3)	0.5(0.7)	2.4(0.3)	2.8(0.1)
4A:15U (1N32)	−11.3(9.1)	−0.4(4.6)	−7.4(7.0)	0.2(0.4)	0.1(0.3)	2.8(0.1)
	−11.0(15.7)	−16.6(10.4)	−6.7(9.3)	0.3(0.6)	−0.2(0.4)	2.8(0.1)
	−5.6(7.8)	0.7(5.3)	−10.9(5.4)	0.3(0.6)	0.0(0.3)	2.8(0.2)
5A:14G (1N32)	−13.4(8.0)	6.8(4.5)	0.5(6.1)	0.4(0.4)	2.7(0.3)	3.4(0.2)
	−12.8(8.2)	9.3(4.6)	−0.7(7.4)	0.4(0.4)	2.5(0.3)	3.4(0.2)
	−17.5(7.0)	12.4(3.3)	−10.6(5.9)	0.0(0.4)	2.7(0.3)	3.3(0.2)

[a]Key: 1st line, CHARMM; 2nd line, AMBER; 3rd line, crystal structure database average (standard deviation values are given within parentheses).

during simulation time. The 12-Gua residue of the A:G basepair in 1FJG shows higher movement for both CHARMM and AMBER, whereas its pairing residue 5-Ade, and other residues having noncanonical basepairing, e.g., 4-Ura and 11-Ura of 1J5A, 5-Ura and 18-Gua of 1N33 show noticeable movements only in CHARMM. The large difference of rmsf values between 1J5A_CHARMM and 1J5A_AMBER is suggestive of distinct behavior of the two systems, the fluctuation being higher for 11-Ura in 1J5A_CHARMM. Again, the shift of 5-Ade from its average position is higher than that of 12-Gua for 1FJG_CHARMM, whereas 1FJG_AMBER shows little movement for this base. The rmsf plots also indicate larger movements in CHARMM force-field for few other nucleotides participating in Watson−Crick basepairing like 4-Cyt of 1FJG, 7-Ura of 1N33, and 8-Ade of 1N33.

2. Variations in Backbone Dihedrals. In general, backbone dihedrals of the residues involved in noncanonical basepairs show regular behaviors under both the force fields. Nucleotides with higher rmsf have unusual values of ε and ζ with larger standard deviations, indicating that the rotations around C3′−O3′ and O3′−P linkages are more allowed to attain nonstandard conformations, especially in simulations with CHARMM27 force-field. For example, the standard deviation values for ε and ζ dihedrals of 5-Ade (1FJG_CHARMM) are ±25.9° and ±40.8°, respectively [see Supporting Information, part S1]. Likewise, anomalous conformations are observed for 11-Ura (1J5A) with ε of −135.4°(±21.1°) and ζ of −53.5°(±32.2°). All the other torsion angles take up values within their respective normal ranges. Whereas α and γ maintain the g^-/g^+, or sometimes the less common t/t conformations, β and glycosidic χ angles display stable *trans*-orientations in both the force-fields. Lesser fluctuations are seen for the torsion angles in 2 × 2 and 3 × 3 noncanonical motifs, though some average values are anomalous. A detailed discussion about their torsion angle variations can be found in our previous report.[7] The sugar puckers also displays regular A-RNA behavior, C3′-endo being majority of the population in both the force-fields.

3. Structural Variations in Canonical and Non-canonical basepairs. Apart from the terminal basepairs of double helical stretches, the canonical Watson−Crick basepairs are quite stable in both CHARMM and AMBER force-fields. However, the variation of basepairing geometry as represented by translational [stagger, shear, stretch] and rotational [open, buckle, propeller] parameters suggest that CHARMM allows larger movements of the basepairs in all directions. The stacking parameters also show their usual trends. It could be noted here that the hydrogen bonding edge specific axis system, used for calculation of intrabasepair parameters by NUPARM,[34] generates values of the parameters—buckle, open angle, propeller twist, stagger and shear—close to zero for a strong, stable and planar basepair, irrespective of its basepairing geometry. This helps to retain consistency of parameter values with canonical basepairs unlike Curves and 3DNA.[35,39] The stretch values calculated by NUPARM are, however, around 3 Å for both canonical and noncanonical basepairs.

As for the noncanonical basepairs, we find that 4U:11U basepair of 1J5A is stable in AMBER, but fluctuates significantly in CHARMM from the beginning of simulation run. This basepair is ruptured from the minor groove side and thus has a high negative open-angle value (−66.9° ± 45.8°) [see Supporting Information, parts S2, S3, and S4]. However, the 5A:12G basepair of 1FJG shows disruption of H-bonding at major groove side between 6 and 7 ns in both CHARMM and AMBER force fields, as confirmed from their open angle values of 39.4°(±34.2°) and 67.0°(±27.1°) in CHARMM and AMBER, respectively [Figure 3].

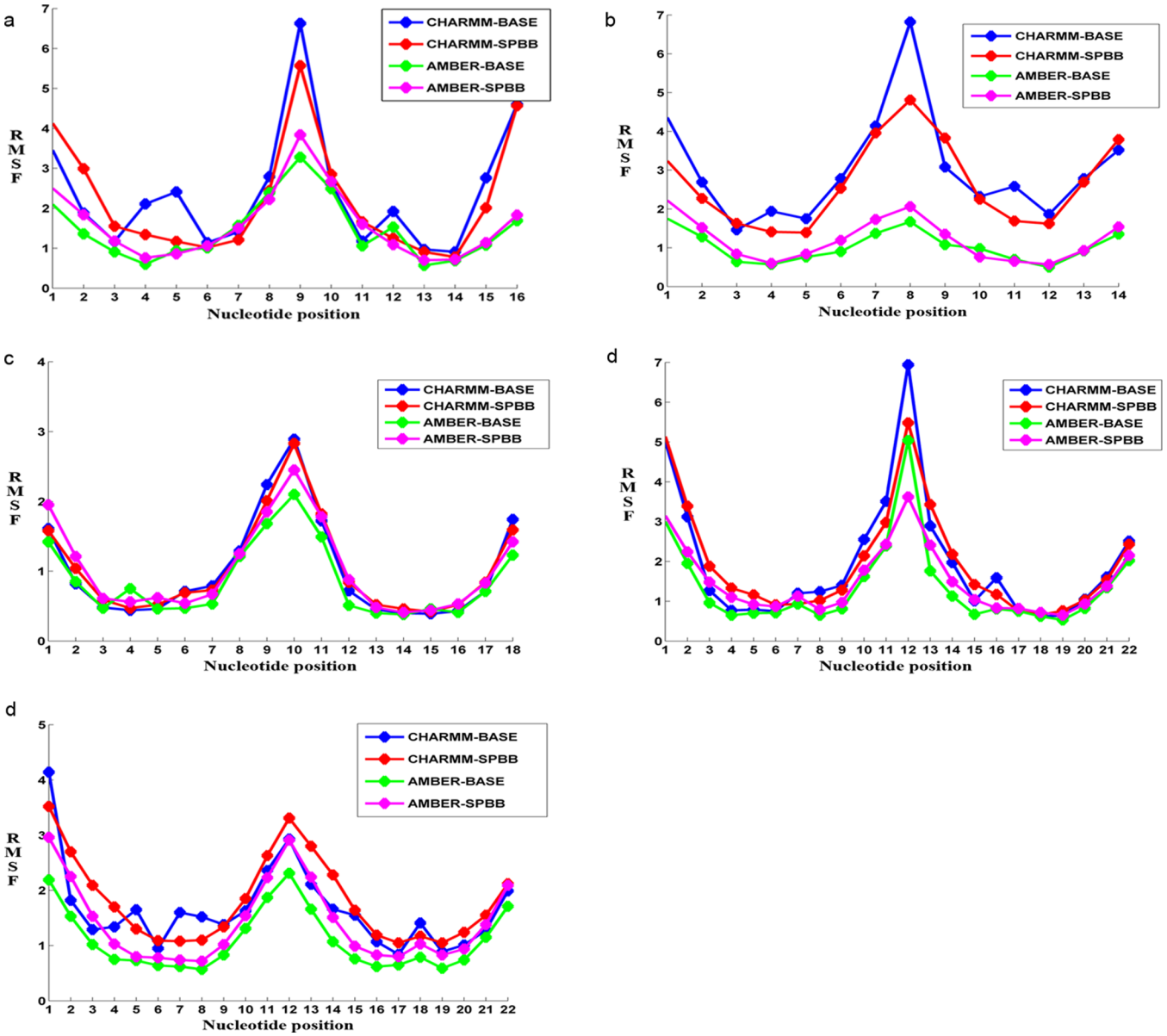


Figure 2. Nucleotide-wise RMSF values for (a) 1FJG, (b) 1J5A, (c) 1N32, (d) 2AW4, and (e) 1N33.

3.1. A:G W:WC Basepair. The out-of-plane motions of the A:G W:WC, leading to large magnitude of propeller twist or buckle as observed by Sponer et al.,[11] are restricted as the basepair in our system is located centrally within a double helix, stacked by canonical Watson−Crick basepairs on either side. Instead, it tries to reduce the strain introduced in double helical architecture, due to its larger strand separation, by deviating from its initial head-on orientation and maintaining a sheared structure where the base centers are located closer to each other. Plot a in Figure 3 gives us the open angle variation for 5A:12G basepair from 1FJG. It shows a harmonic open-angle variation centered about zero at the beginning with fluctuations being restricted within ±5° about the average, indicating the existence of a stable structure up to ∼6.9 ns. The structure consists of two polar hydrogen bonds formed between N6A-H6A···O6G and N1A···H1G-N1G, with both the bases using their Watson−Crick edges to form a stable imino-basepair (Figure 1a). This steady basepairing geometry is disrupted with a sudden shearing motion of 12-Gua toward the major groove side, leading to a higher open-angle value and a negative shear [see Supporting Information, part S2]. This motion leads to rupture of hydrogen bonding and movement of one base out of the helical stack manifested in lowered overlap area with adjacent 4C:13G and 6U:11G basepairs, which is energetically unfavorable. The interaction energy profile for this basepair thus shows a destabilization due to breakage of stable H-bonds and disruption in stacking interactions. We have calculated the approximate interaction energies for the two structures considering the energy values of different types of H-bonds.[40] The initial structure consists of one N−H···O and one N−H···N type of polar interactions, and hence, the contribution from H-bonding becomes −12.8 kcal/mol [including the term for nonspecific interactions], which is comparable to the energy obtained from MD data before transition, −12.1 kcal/mol. However, this value differs from the energy of an A:G W:WC basepair in its optimized geometry in gas phase [−15.0 kcal/mol][40] or solvated state [−15.66 kcal/mol].[12] The discrimination arises because of the G(NH2)-A(H2C) steric repulsion

dx.doi.org/10.1021/jp305628v | J. Phys. Chem. B 2012, 116, 11845−11856

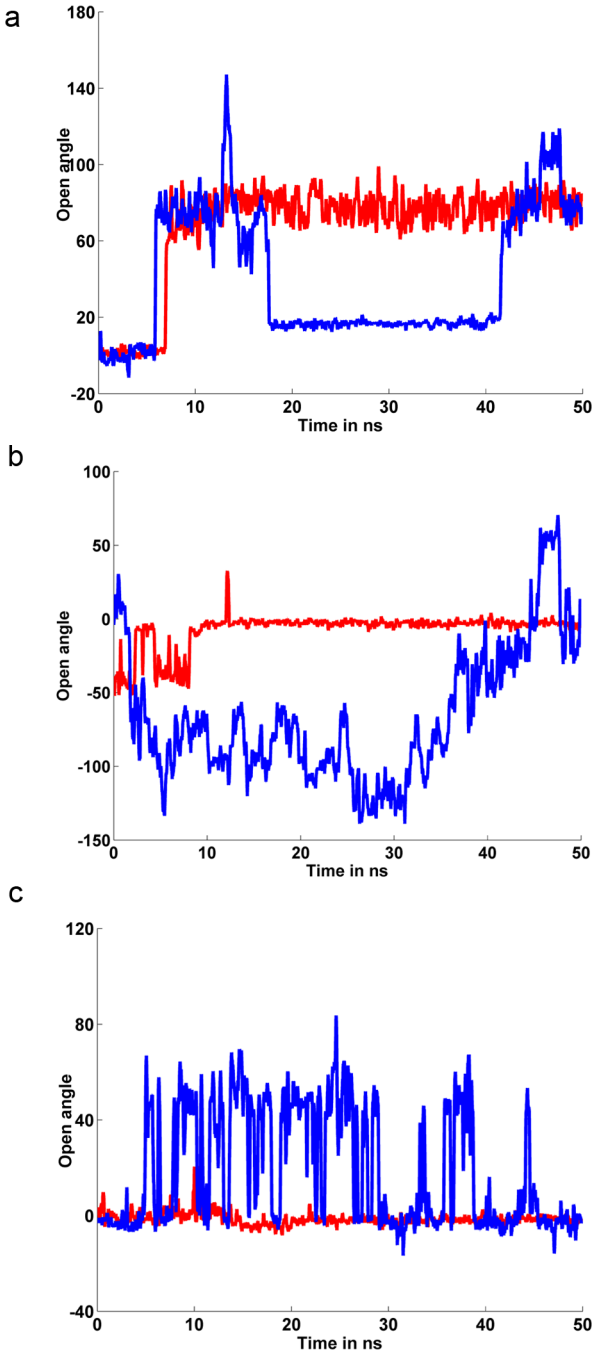


Figure 3. Open angle plots for (a) 5A:12G of 1FJG, (b) 4U:11U of 1J5A, and (c) 5U:18G of 1N33: red, AMBER; blue, CHARMM.

present in A:G basepairs within a double helix, but absent in optimized isolated basepair. As quantum chemical studies for geometry optimization are generally carried out with isolated basepairs representing the gaseous phase, 2-amino group in guanine can adopt a nonplanar conformation by partial sp^3 hybridization of nitrogen to diminish steric clash with H2C atom of the pairing adenine residue.[11,12,41] Moreover, extra stabilization of the sheared structure comes from release of steric strain due to larger separation between G(NH2) and A(H2C), as well as water mediated interactions between the polar atoms. The interatomic distance between these atoms is increased from ∼3 to ∼8 Å during this transition. This sheared structure of A:G basepair is maintained throughout the simulation in 1FJG_AMBER. We find that both the bases further adjust their relative orientations in 1FJG_CHARMM to result in a stable A:G H:ST basepaired structure, indicated by the second plateau region during 20−40 ns in Figure 3a. A:G H:ST has a well-defined basepairing geometry and is more favorable than the sheared intermediate structure. It is also one of the most frequent noncanonical basepairs available in crystal structure database. Because of this conformational change, the bases are further rotated leading to an increased twist and better overlap in the dinucleotide steps involving the A:G basepair, signifying enhanced stacking compared to the water-mediated structure. It is also important to note that the 6-amino group of adenine has more polar N−H bonds in AMBER94 than CHARMM27 and thus, more available for hydrogen bonding. However, the transition is found for CHARMM rather than AMBER indicating its force-field independent nature. This structure is maintained for a stretch of ∼20 ns before returning to its previously observed water-mediated hydrogen bonded state.

3.2. U:U W:WC Basepair. The 4U:11U basepair in 1J5A behaves distinctly in two force fields (Figure 3b). It shows a feature of partial instability from the early stage of simulation in 1J5A_CHARMM, whereas maintains an organized basepaired structure in 1J5A_AMBER. In the crystal, the basepair is formed by hydrogen bonding between two uracil residues through potential donor and acceptor atoms located along the Watson−Crick edges. The O2 and N3 atoms of one uridine form direct hydrogen bonds with N3 and O4 atoms of the other uridine (Figure 1b). Generally, the bases move closer toward each other and also toward the helical axis to form this basepair. The geometry suggests that one of the two uracil bases is rotated toward major groove, with O4 atom protruding out. This has been referred to as an "open" conformation, whereas the other base is said to be in "closed" form.[42] Interestingly, this type of basepairing scheme can be achieved by keeping either of the two bases in "open" conformation.[12] In 1J5A_AMBER, we find that the 4-Ura residue is in open conformation, while 11-Ura maintains a closed conformation.

Many studies on uracil−uracil basepairing in RNA duplex have suggested their flexibility and instability.[14,15] Such basepairs with free carbonyl (>C═O) or amino ($-NH_2$) group was shown to be capable of acting as conformational switch.[12,43] We observe that the centrally located U:U basepair in 1J5A_CHARMM undergoes frequent vibrations between unpaired and hydrogen-bonded states, manifested by the large average and standard deviation values of intrabasepair parameters like open angle [−66.9° ± 45.8°], shear [1.0 Å ± 3.7 Å], and stretch [4.0 Å ± 2.4 Å], which denote basepairing strength and stability. Further analyses reveal that the basepair attains an unstable transition state where only one hydrogen bond between the two bases can be formed through N3(4-Ura)-O4(11-Ura) by rotating either or both the bases around helix axis [see Supporting Information, part S2]. The structure can be considered as an intermediate between two basepairing schemes − W:WC with two polar H bonds (between O2U−N3U and N3U−O4U) and w:hT with one polar and one C−H···O-mediated weakly polar interaction (between N3U−O4U and O4U−C5U) − and any of these two stable basepairing states is possible to achieve by rotating the bases to either side. According to the force-field descriptions, the C5−H5 bond of Uracil is more polar in case of AMBER94 than CHARMM27. Still the formation of observed C−H···O-type of weakly polar hydrogen bond between the bases in CHARMM force-field

rather than AMBER suggests a force-field independent behavior of the system during simulation. This U:U w:hT pair is not compatible within an RNA duplex because of the conformational strain imparted in the nucleic acid backbone via nucleobase rotation.[42] The basepair is unstable in gas phase as well as in aqueous environment due to low interaction energies.[12,40] We have estimated the basepairing energies for both the basepairing geometries and for the intermediate structure also based on their H-bonding patterns.[40] The interaction energy values for W:WC pair (with two N−H···O type of interactions), h:wT pair (with one N−H···O and one C−H···O types of interactions) and intermediate (with one N−H···O type of interaction) are −14.1, −8.8, and −1.7 kcal/mol, respectively. Water-mediated interactions are found to stabilize all the three structure variants found during simulation. However, the basepair remains stable throughout the simulation in AMBER force-field maintaining the original two hydrogen-bonded W:WC form.

3.3. G:U W:WC Basepair. The G:U wobble basepair is present in the five double helices at several different locations and this basepair shows quite flexible and adaptable nature. Their shear values are significantly different from the canonical A:U or G:C W:WC basepairs.[12] These large shear values are generally found to be accompanied by larger open angle values during the simulations, suggestive of a possible conformational alteration. For example, 5U:18G of 1N33 shows a basepair opening at the major groove side as indicated by its high positive open-angle (18.5° ± 26.0°), compared to ∼0° open-angle value for ideal Watson−Crick basepair. The average open angle value for G:U W:WC basepairs in nonredundant crystal structure database is also close to 0°.[12,34] The stagger, buckle or propeller twist values, however, remain small, indicating that the altered conformation maintains stacking between neighboring basepairs [see Supporting Information, part S3]. Such high open angle values are results of a transition from W:WC to W:SC structure, forming a hydrogen bond between the 2′-OH of pyrimidine and the Watson−Crick edge of guanine residue. The transition initiates with disruption of H-bonding between the bases in the major groove side. This breaking of H-bond is compensated by electrostatic interaction with surrounding water molecules, thus resulting in a G:U basepair with large value of open-angle. The two bases are more displaced with respect to each other along the basepair short axis, to allow interaction between the polar hydrogen of pyrimidine 2′-hydroxyl and 2-amino group of guanine [see Supporting Information, part S2]. Thus, such transition is not possible in DNA for equivalent G:T pairs due to the absence of the 2′-OH group. It is a known fact that basepairs like G:U which have free amino or carbonyl groups on either base can lead to basepair alteration and may play important roles as conformational switches.[12,43,44] Moreover, the 2-amino hydrogens in guanine are available for polar interactions and can form strong hydrogen bonds with the 2′-hydroxyl group according to both the force-field parameters. The G:U pairs in W:SC form also try to revert back to the original W:WC form, thus suggesting that the interconversion between canonical and noncanonical forms involves a low energy barrier especially in water environment, which makes the transition feasible even at 300 K. However, the transition energy cannot be determined accurately due to large number of interacting solvent molecules present in the surrounding. This transition effect is not seen for the 5U:18G basepair in 1N33_AMBER while 1N33_CHARMM shows a tendency to undergo structural transition to achieve sugar-mediated basepairing. However, it fails to maintain a stable conformation in the altered state and majority of the ensemble population remains as canonical W:WC form. Table 3 shows the frequencies of transition for the wobble G:U basepairs under study.

Table 3. Transition of G:U Basepairs in Different Double Helices in Terms of % of Occurrence during Simulation Time Period

		CHARMM, %	AMBER, %
1FJG	6U:11G	91.4	80.2
	7G:10U	89.5	78.5
1J5A	2U:13G	66.6	1.0
	5G:10U	0	82.0
1N33	5U:18G	37.4	0.2
1N32	1G:18U	5.6	2.8
2AW4	2U:21G	60.1	8.2
	3G:20U	67.6	7.8
	7G:16U	54.5	4.5

3.4. Tandem Noncanonical Basepairs. In contrast to what we observe in case of single mismatches, noncanonical basepairs forming (2 × 2) and (3 × 3) internal loops in 1N32 and 2AW4, respectively, show high stability in MD simulations under both the force-fields. We have already discussed about the detailed structural features of (2 × 2) and (3 × 3) noncanonical mismatches in our previous report.[7] They remain stable in their initial orientations throughout the simulation runs of 50 ns in CHARMM force-field and the structural parameters usually are small and have lower standard deviation values [Table 2]. This structural stability appears to be force-field independent as the helices remain stable in AMBER force fields also. However, some local variations can be observed between the two force-fields, although the basic natures remain same.

4. Isostericity—Role of Wobble G:U Basepairs. We observe that the noncanonical mismatches induce deformations within the helical stack, thus causing deviation from canonical A-form. While the purine−purine mismatches like A:G W:WC require larger strand separation, a pyrimidine−pyrimidine basepair like U:U would require a narrowing in the helical architecture. Such deformations are perhaps required for proper molecular recognition. The (2 × 2) and (3 × 3) noncanonical basepairs also introduce a cleft in the central regions of helices 1N32 and 2AW4 having lower strand separations than canonical Watson−Crick basepairs. On the other hand, the A:G W:WC basepair has an initial distance of 13.0 Å between the C1′ atoms, whereas the values for normal Watson−Crick basepairs are 10.69 Å (A-U/U-A) and 10.77 Å (G═C/C═G).[12] Considering the distance between C1′ atoms as an estimation of helical strand separation, we can easily assume a swelled region within 1FJG helix due to the presence of A:G W:WC pair. The relative shearing motion of the bases near ∼6.9 ns makes the base-centers to move toward each other; as a result, the C1′−C1′ distance is decreased to an average value of ∼11.5 Å and the bulge is reduced [Figure 4a,b and Figure 5]. This distance is closer to that of canonical Watson−Crick basepairs and more suitable to maintain an A-form helical stack. We also observe that adjacent G:U basepairs, 6U:11G and 7G:10U, undergo a concurrent transition from W:WC to W:SC form. Now, G:U in its original W:WC form is nearly isosteric with canonical Watson−Crick basepairs having a strand

dx.doi.org/10.1021/jp305628v | J. Phys. Chem. B 2012, 116, 11845−11856

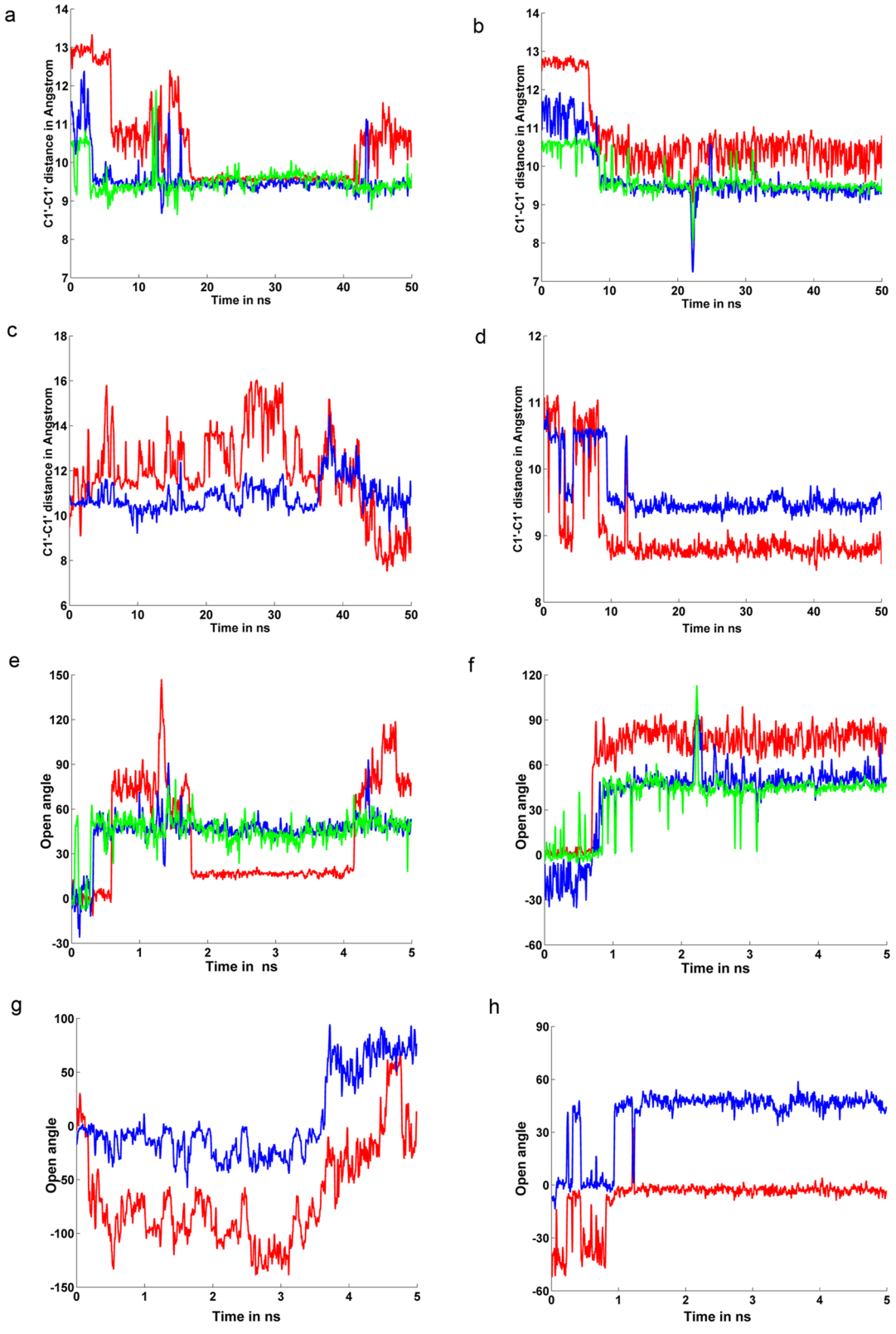


Figure 4. C1′−C1′ Distance for (a) 1FJG_CHARMM, (b) 1FJG_AMBER [red, 5A:12G; blue, 6U:11G; green, 7G:10U], (c)1J5A_CHARMM, and (d) 1J5A_AMBER [red, 4U:11U; blue, 5G:10U]. Basepair opening for (e) 1FJG_CHARMM, (f) 1FJG_AMBER [red, 5A:12G; blue, 6U:11G; green, 7G:10U], (g) 1J5A_CHARMM, and (h) 1J5A_AMBER [red, 4U:11U; blue, 5G:10U].

separation of 10.30 Å, whereas transition to W:SC form leads to a decrease in its C1′−C1′ distance to 9.7 Å. This matches strand separation of the altered conformation of A:G basepair (Figure 4, parts a and b). We find concomitant transitions of 5A:12G with 6U:11G and 7G:10U between 6 and 7 ns in both 1FJG_AMBER and 1FJG_CHARMM. This is also represented by the change of basepair opening in the helices. In Figure 4, parts e and f, the G:U basepairs of helix 1FJG open up along

dx.doi.org/10.1021/jp305628v | J. Phys. Chem. B 2012, 116, 11845−11856

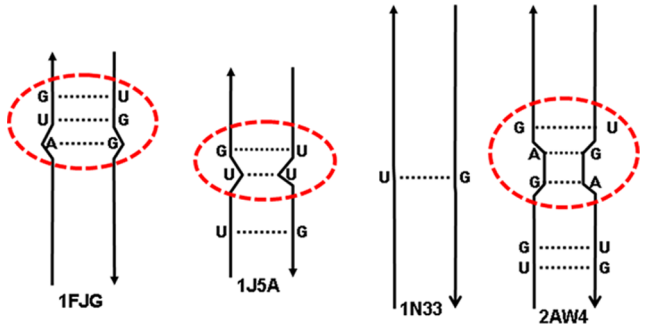


Figure 5. Schematic representation of helices with the role of adjacent G:U basepairs in maintaining the helical architecture.

with the A:G mismatch and undergoes transition to W:SC form, which has a open angle of ~60°. It can also be assumed that 5U:18G basepair of 1N33 stacked within Watson–Crick basepairs require no transition as they can be accommodated within the canonical A-form without much deviation in the backbone direction. Again, for 2U:13G of 1J5A_CHARMM, 1G:18U of 1N32_CHARMM, 2U:21G and 3G:20U of 2AW4_CHARMM, the variability possibly arises due to their proximity to the partially unpaired terminal region.

On the other hand, the helix 1J5A contains a pyrimidine–pyrimidine pair in its central region, which has a C1′–C1′ distance smaller than Watson–Crick basepairs (the initial average for C1′ separation in 4U:11U is ~9.5 Å). Clearly, this is not sufficient to maintain isostericity with the flanking canonical basepairs. So, a structural transition to increase the C1′–C1′ distance is obvious. As mentioned before, this U:U basepair shows different behaviors under the two force-fields. In 1J5A_CHARMM, it tries to increase its C1′–C1′ distance by breaking the W:WC geometry and rotating either or both bases toward the minor groove (Figure 4c). Thus, no well-defined basepairing pattern can be detected and the destabilization is extended to the adjacent 5G:10U basepair, which also fails to

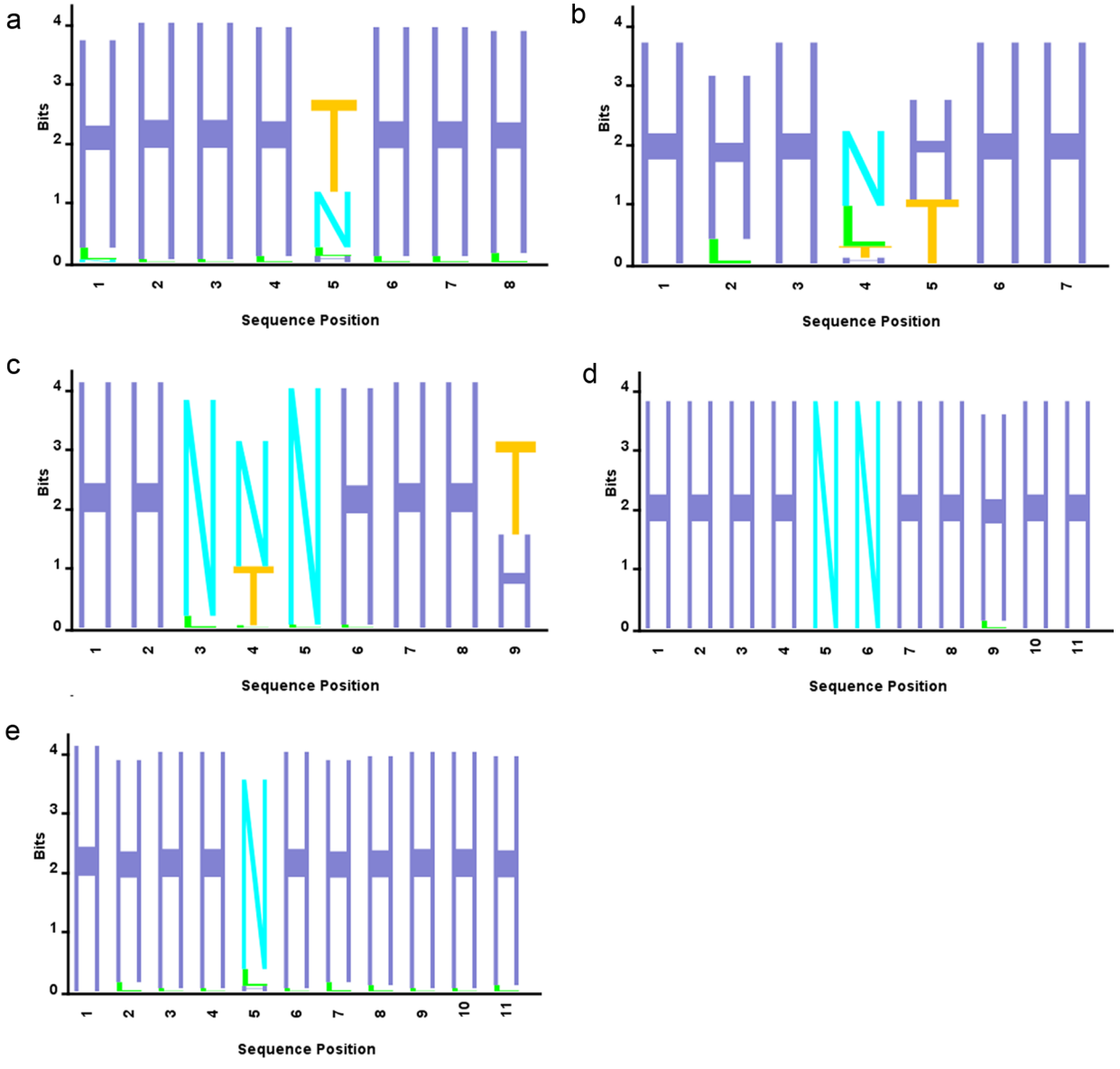


Figure 6. Basepair stability of the helices (a) 1FJG, (b) 1J5A, (c) 1N33, (d) 1N32, (e) 2AW4. Plots show the pairing propensities represented by alphabets: H, Watson–Crick basepair; N, noncanonical basepair; L, unpaired residues; T, triplets.

dx.doi.org/10.1021/jp305628v | J. Phys. Chem. B 2012, 116, 11845–11856

maintain stable hydrogen bonding in either W:WC or W:SC form. On the contrary, the 4U:11U of 1J5A_AMBER retains its initial W:WC form throughout the simulation with its smaller strand separation [Figure 5]. This local strain is released by distributing the deformation to the next 5G:10U basepair, which reduces its C1′−C1′ distance through conversion to sugar-mediated W:SC form and maintains the helical architecture [Figure 4(d)]. However, in either case, a correlated basepair opening pattern of U:U mismatch and its adjacent G:U basepair becomes distinct from the open angle plots in Figure 4, parts g and h. We also find that 7G:16U basepair of 2AW4 (next to 2 × 2 noncanonical motif) undergoes significant transition in CHARMM force-field to maintain the strand separation of ∼9.5 Å similar to neighboring A:G H:ST pairs [see Supporting Information, part S4]. However, this basepair shows significantly less transition in AMBER force-field, probably due to limited sampling or stronger basepairing energy. The transition to W:SC form for 5U:18G (1N33) is much less pronounced as it is located within the central region flanked by Watson−Crick basepairs on both sides and hence it tries to revert back to its initial W:WC form.

5. Crystal Structure Analysis. We have observed from MD simulations that the tandem noncanonical basepairs present within RNA duplex regions (1N32 and 2AW4) are more stable than the singly occurring ones like A:G W:WC in 1FJG and U:U W:WC in 1J5A. To validate our observations against experimental results, we have also analyzed the crystal structure database to check for basepairing patterns in these helices. Now, it is a general fact that many of the Watson−Crick as well as noncanonical basepairs in RNA are rather flexible as all the atoms fluctuate at physiological temperature. The overall basepair motions also take place at a frequency of about 10 ps[40] giving rise to paired and unpaired bases in a large sampling. Finding out a basepair in a RNA crystal structure depends on finding specific atoms within short distances as is the case for locating α-helices or β-sheets in a protein structure.[45] In some structures, such H-bonds are longer due to contextual pressure, which are not detected by programs like DSSP, Stride or BPFind.[5,45,46] This indicates lack of conservation of the helical form.

We have analyzed the complete crystal structure database as available in PDB for the double helical stretches containing noncanonical basepairs. We find that A:G and U:U basepairs have frequencies 194 and 374 respectively within double helical regions, while tandem mismatches have frequencies 829 (for 2 × 2) and 97 (for 3 × 3). Both the (2 × 2) motif and U:U mismatch are found in small ribosomal subunits (16S rRNA) of *T. thermophilus* and *E. coli* and large ribosomal subunits (23S rRNA) of *Haloarcula marimortui*, *D. radiodurans*, *T. thermophilus*, and *E. coli*, sometimes occurring more than once within same molecule. The A:G mismatch is found within small subunits of *T. thermophilus* and *E. coli* as well as a large subunit of *E. coli*. The (3 × 3) motif is less populated and is found only within small ribosomal subunits (16S rRNA) of *T. thermophilus* and *E. coli*. However, this motif is highly conserved among all the crystal structures of 16S rRNA of *T. thermophilus* and *E. coli*. On the other hand, although U:U and A:G mismatches are found in more types of molecules spanning both large and small ribosomal subunits, the basepairs are not conserved among all the structures in the respective classes and thus have lower propensity within double helical regions. In order to detect conservations of the five double helices in their respective crystallographic ensembles, we have used the structural classifications given by HD-RNAS database in Feb. 2012.[13] According to this classification, we obtain 35 structures in the 1J5A class (23S rRNA of *D. radiodurans*) and 52 structures in the 2AW4 class (23S rRNA of *E. coli*). All the other three helices, 1FJG having A:G basepair, 1N32 having 3 × 3 noncanonical motif and 1N33 having G:U basepair, are taken from the 16S rRNA class of *T. thermophilus*, having 102 structures. A nucleotide residue in RNA has been considered to be in one of the four states—(i) in a Watson−Crick base-paired state (H), (ii) in a noncanonical base-paired state (N), (iii) in triplet state (T) which also implies paired state in N or H forms, and (iv) unpaired looped out state (L). Figure 6 shows the propensity of the residues in the helices as a function of alphabet size. The 5A:12G basepair of 1FJG and the 4A:15U basepair of 1N32 are often found to form triplets while maintaining the original basepairing patterns. It is found from crystal structure analysis that none of the two single mismatches remain conserved in their crystallographic ensembles. They show a tendency to open up, U:U basepair being the most unstable one. The isolated wobble G:U basepair in 1N33 also display fluctuations in the crystallographic ensemble of *T. thermophilus* 16S rRNA. On the other hand, tandem noncanonical basepairs comprising the 2 × 2 and 3 × 3 internal loops [helix 1N32 and helix 2AW4] are retained in almost all the crystal structures of their corresponding groups. The structural parameters also show large deviations for single mismatches, indicating their labile nature in the crystal structure database [Table 2]. In case of triplet formation, we find some structural anomaly in crystal structures. We have previously discussed the effect of triplet formation on the structure of 4A:15U H:WT in 1N32.[7] Here we observe that most of the crystal structures in 1FJG class maintain a triplet form involving the 5A:12G W:WC basepair. This introduces fluctuation in the structure of this basepair in crystal database. However, such tertiary interactions within the macromolecule are supposedly required for stability of the structure, as observed for the 1N32 helix (4A:15U H:WT) previously,[7] and their absence in our solvated helices may cause instability of the 5A:12G basepair during MD simulations.

CONCLUSION

We have performed 50 ns long simulations of three double-helical stretches of RNA containing single mismatches in the central regions under both CHARMM and AMBER force-fields. We observe that CHARMM allows more fluctuations of the basepairs than AMBER. Even Watson−Crick basepairs are also departed from their stable structures by CHARMM. It is also clear from our studies that both purine−purine and pyrimidine−pyrimidine basepairs try to fit themselves in the helical scaffold. It is found that reduced stability of the singly occurring A:G and U:U basepairs are also in agreement with the crystallographic database analysis of the respective structural classes. The occurrences of single mismatches in their respective structural classes are lower than that of the 2 × 2 and 3 × 3 mismatches. Helices containing tandemly occurring noncanonical basepairs, 1N32 and 2AW4, are more conserved in crystal structures. This perhaps indicates that both the force fields are reasonably reliable though some fine-tuning may improve their performances. Wobble G:U basepairs adjacent to these single mismatches are seen to induce required flexibility in the local environment through their biphasic stability. The energy barrier required for transition from original wobble form

to sugar-mediated form is also low enough to be allowed at room temperature.

The importance of G:U basepairs in wobble position of codon−anticodon recognition was first proposed by Francis Crick[47] and, later on, was studied and modified by many others. They have also been shown to have free energy comparable to A:U Watson−Crick basepairing through thermodynamic as well as theoretical studies. According to the wobble hypothesis, G:U basepairs are known to populate the third position of codon−anticodon recognition, along with other wobble basepairs containing Ionosine, during the protein synthesis process within ribosome. Thus, these basepairs often appear at the termini of mini helices between codon of mRNA and anticodon of tRNA, giving rise to distortion. It is also speculated that the overall geometry of codon−anticodon minihelix is, in many cases, determined by that of the wobble position.[47] The wobble position is already known to display variation in terms of both basepairing geometry and chemical modification of the bases. We have seen that the structural transition of G:U basepairs is crucial to induce flexibility at the junction of canonical and noncanonical regions within a double helix. Similarly, the transition to sugar mediated form can have an important role in determining local geometry of the minihelix formed between anticodon-loop of aminoacyl tRNA and the cognate mRNA. Again, protein translation mechanism requires flexibility at this region because the tRNA binds to mRNA with high specificity and leave the binding site of mRNA once peptide bond formation is complete. The biphasic stability of G:U basepairs at wobble position may be playing an important role in the regulation of short-lived codon−anticodon recognition by acting as a conformational switch. However, it requires an extensive experimental and theoretical study of G:U and additional wobble basepairs to substantiate this hypothesis.

■ ASSOCIATED CONTENT

Supporting Information

Table S1, average and standard deviation values of backbone dihedrals for the bases involved in noncanonical basepairing, Figure S2, structural transition of A:G W:WC, U:U W:WC and G:U W:WC, Table S3, average and standard deviation values of G:U basepairs present in the helices, Figure S4, plot of C1′−C1′ distances for different G:U basepairs present in the double helical stretches, and Table,S5, a list of structures present in crystallographic classes 1FJG, 1J5A, 1N33, 1N32, and 2AW4. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Telephone: +91 33 2337 0379, Extn. 3506. Fax: +91 33 2337 4637. E-mail: dhananjay.bhattacharyya@saha.ac.in.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Ms. Stephannie Kaypee for her technical support. This project was partially funded by Department of Biotechnology, Government of India. We thank the Supercomputing Facility (BRAF), Center for Development of Advanced Computing, Pune, India, for computational facilities used.

■ REFERENCES

(1) Peritz, A. E.; Kierzek, R.; Sugimoto, N.; Turner, D. H. *Biochemistry* **1991**, *30*, 6428−6436.
(2) Calin-Jageman, I.; Nicholson, A. W. *Biochemistry* **2003**, *42*, 5025−5034.
(3) Saito, H.; Richardson, C. C. *Cell* **1981**, *27*, 533−542.
(4) Leontis, N. B.; Stombaugh, J.; Westhof, E. *Nucleic Acids Res.* **2002**, *30*, 3497−3531.
(5) Das, J.; Mukherjee, S.; Mitra, A.; Bhattacharyya, D. *J. Biomol. Struct. Dyn.* **2006**, *24*, 149−161.
(6) Leontis, N. B.; Westhof, E. *RNA* **2001**, *7*, 499−512.
(7) Halder, S.; Bhattacharyya, D. *J. Phys. Chem. B* **2010**, *114*, 14028−14040.
(8) Davis, A. R.; Znosko, B. M. *Biochemistry* **2008**, *47*, 10178−10187.
(9) Davis, A. R.; Znosko, B. M. *Abstr. Pap. (Am. Chem. Soc.)* **2009**, *237*, x.
(10) Christiansen, M. E.; Znosko, B. M. *Nucleic Acids Res.* **2009**, *37*, 4696−4706.
(11) Sponer, J.; Mokdad, A.; Sponer, J. E.; Spackova, N.; Leszczynski, J.; Leontis, N. B. *J. Mol. Biol.* **2003**, *330*, 967−978.
(12) Panigrahi, S.; Pal, R.; Bhattacharyya, D. *J. Biomol. Struct. Dyn.* **2011**, *29*, 541−556.
(13) Ray, S. S.; Halder, S.; Kaypee, S.; Bhattacharyya, D. *Front. Genet.* **2012**, *3*, 59.
(14) Alkema, D.; Hader, P. A.; Bell, R. A.; Neilson, T. *Biochemistry* **1982**, *21*, 2109−2117.
(15) Santalucia, J.; Kierzek, R.; Turner, D. H. *Biochemistry* **1991**, *30*, 8242−8251.
(16) Alkema, D.; Bell, R. A.; Hader, P. A.; Neilson, T. In *Biomolecular Stereodynamics*; Sarma, R. H., Ed.; Adenine Press: New York, 1981.
(17) Alkema, D.; Bell, R. A.; Hader, P. A.; Neilson, T. *J. Am. Chem. Soc.* **1981**, *103*, 2866−2868.
(18) Everett, J. R.; Hughes, D. W.; Bell, R. A.; Alkema, D.; Neilson, T.; Romaniuk, P. J. *Biopolymers* **1980**, *19*, 557−573.
(19) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(20) Foloppe, N.; MacKerell, A. D. *J. Comput. Chem.* **2000**, *21*, 86−104.
(21) Cheatham, T. E.; Kollman, P. A. *J. Mol. Biol.* **1996**, *259*, 434−444.
(22) Feig, M.; Pettitt, B. M. *Biophys. J.* **1998**, *75*, 134−149.
(23) Feig, M.; Pettitt, M. *J. Phys. Chem. B* **1997**, *101*, 7361−7363.
(24) Gong, Z.; Zhao, Y. J.; Xiao, Y. *J. Biomol. Struct. Dyn.* **2010**, *28*, 431−441.
(25) Romanowska, J.; Setny, P.; Trylska, J. *J. Phys. Chem. B* **2008**, *112*, 15227−15243.
(26) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235−242.
(27) Samanta, S.; Mukherjee, S.; Chakrabarti, J.; Bhattacharyya, D. *J. Chem. Phys.* 2009, *130*.
(28) Darden, T.; Perera, L.; Li, L. P.; Pedersen, L. *Struct. Fold. Des.* **1999**, *7*, R55−R60.
(29) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.
(30) Case, D. A.; Darden, T. A.; Cheatham, T. E. I.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Wang, B.; Pearlman, D. A.; Crowley, M.; Brozell, S.; Tsui, V.; Gohlke, H.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Schafmeister, C.; Caldwell, J. W.; Ross, W. S.; Kollman, P. A. *Amber 8*; University of California, San Francisco, San Francisco, CA, 2004.
(31) Kale, L.; Skeel, R.; Bhandarkar, M.; Brunner, R.; Gursoy, A.; Krawetz, N.; Phillips, J.; Shinozaki, A.; Varadarajan, K.; Schulten, K. *J. Comput. Phys.* **1999**, *151*, 283−312.
(32) Nelson, M.; Humphrey, W.; Kufrin, R.; Gursoy, A.; Dalke, A.; Kale, L.; Skeel, R.; Schulten, K. *Comput. Phys. Commun.* **1995**, *91*, 111−133.

(33) Bansal, M.; Bhattacharyya, D.; Ravi, B. *Comput. Appl. Biosci.* **1995**, *11*, 281−287.
(34) Mukherjee, S.; Bansal, M.; Bhattacharyya, D. *J. Comput. Aided Mol. Des.* **2006**, *20*, 629−645.
(35) Lu, X. J.; Olson, W. K. *Nucleic Acids Res.* **2003**, *31*, 5108−5121.
(36) Schneider, T. D.; Stephens, R. M. *Nucleic Acids Res.* **1990**, *18*, 6097−6100.
(37) Pan, Y. P.; MacKerell, A. D. *Nucleic Acids Res.* **2003**, *31*, 7131−7140.
(38) Reblova, K.; Lankas, F.; Razga, F.; Krasovska, M. V.; Koca, J.; Sponer, J. *Biopolymers* **2006**, *82*, 504−520.
(39) Lavery, R.; Moakher, M.; Maddocks, J. H.; Petkeviciute, D.; Zakrzewska, K. *Nucleic Acids Res.* **2009**, *37*, 5917−5929.
(40) Roy, A.; Panigrahi, S.; Bhattacharyya, M.; Bhattacharyya, D. *J. Phys. Chem. B* **2008**, *112*, 3786−3796.
(41) Mukherjee, S.; Majumdar, S.; Bhattacharyya, D. *J. Phys. Chem. B* **2005**, *109*, 10484−10492.
(42) Lietzke, S. E.; Barnes, C. L.; Berglund, J. A.; Kundrot, C. E. *Structure* **1996**, *4*, 917−930.
(43) Sharma, P.; Mitra, A.; Sharma, S.; Singh, H.; Bhattacharyya, D. *J. Biomol. Struct. Dyn.* **2008**, *25*, 709−732.
(44) Chawla, M.; Sharma, P.; Hader, S.; Bhattacharyya, D.; Mitra, A. *J. Phys. Chem. B* **2011**, *115*, 1469−1484.
(45) Kabsch, W.; Sander, C. *Biopolymers* **1983**, *22*, 2577−2637.
(46) Heinig, M.; Frishman, D. *Nucleic Acids Res.* **2004**, *32*, W500−W502.
(47) Crick, F. H. C. *J. Mol. Biol.* **1966**, *19*, 548−555.