Molecular Dynamics Simulation of the Human U2B″ Protein Complex with U2 snRNA Hairpin IV in Aqueous Solution

Jian-xin Guo and William H. Gmeiner
Eppley Institute, University of Nebraska Medical Center, Omaha, Nebraska 68198-6805 USA

ABSTRACT A 2200-ps molecular dynamics (MD) simulation of the U2 snRNA hairpin IV/U2B″ complex was performed in aqueous solution using the particle mesh Ewald method to consider long-range electrostatic interactions. To investigate the interaction and recognition process between the RNA and protein, the free energy contributions resulting from individual amino acids of the protein component of the RNA/protein complex were calculated using the recently developed glycine-scanning method. The results revealed that the loop region of the U2 snRNA hairpin IV interacted mainly with three regions of the U2B″ protein: 1) β1-helix A, 2) β2-β3, and 3) β4-helix C. U2 snRNA hairpin IV bound U2B″ in a similar orientation as that previously described for U1 snRNA with the U1A′ protein; however, the details of the interaction differed in several aspects. In particular, β1-helix A and β4-helix C in U2B″ were not observed to interact with RNA in the U1A′ protein complex. Most of the polar and charged residues in the interacting regions had larger mutant free energies than the nonpolar residues, indicating that electrostatic interactions were important for stabilizing the RNA/protein complex. The interaction was further stabilized by a network of hydrogen bonds and salt bridges formed between RNA and protein that was maintained throughout the MD trajectory. In addition to the direct interactions between RNA and the protein, solvent-mediated interactions also contributed significantly to complex stability. A detailed analysis of the ordered water molecules in the hydration of the RNA/protein complex revealed that bridged water molecules reside at the interface of RNA and protein as long as 2100 ps in the 2200-ps trajectory. At least 20 bridged water molecules, on average, contributed to the instantaneous stability of the RNA/protein complex. The stabilizing interaction energy due to bridging water molecules was obtained from ab initio Hartree-Fock and density functional theory calculations.

INTRODUCTION

The interactions that stabilize complexes between ribonucleic acids (RNA) and proteins are of great interest for understanding splicing, translation, and other RNA-mediated processes. At present, there is relatively little information available concerning RNA/protein interactions, especially for the single-stranded RNA/protein systems (Antson, 2000). The relative lack of information concerning RNA/protein complex structure and stability results, in part, from the technical difficulty in determining the structures of RNA/protein complexes. RNA/protein complexes are often difficult to crystallize, and determining RNA structures by NMR is challenging due to poor spectral dispersion and problems with isotopic enrichment. For these reasons, only a few structures of RNA/protein complexes become available each year (Cusack, 1999). Even for those RNA molecules for which structural information is available, identifying the interactions that stabilize RNA/protein complexes is challenging because of the complexity of interactions that contribute to complex stability (Saenger, 1983; Nolan et al., 1999). In many cases, RNA duplexes are not involved in the interactions that stabilize RNA/protein complexes because the RNA bases in helical regions are largely unavailable for hydrogen bond formation. Thus, many of the specific contacts that stabilize RNA/protein complexes involve ribonucleotides in single-stranded or loop regions of RNA (Oubridge et al., 1994; Price et al., 1998; Honda et al., 1999).

As a component of the spliceosomal U2 small nucleic ribonucleoprotein particle (RNP), the U2B″ protein interacts with U2 snRNA hairpin IV (Fig. 1). Recently, an x-ray structure of the U2 snRNA hairpin IV/U2B″ complex was reported (Price et al., 1998). The structure of this complex revealed U2B″ interacted with U2 hairpin IV in a similar manner to that previously observed to occur for the U1A protein in complex with U1 snRNA hairpin II. The U2A′/U2B″/U2 snRNA forms a sandwich-like ternary complex in which U2A′ interacts strongly with U2B″, and interacts weakly with U2 snRNA hairpin IV. This strong interaction between U2A′/U2B″ alters the conformation of U2B″, allowing U2B″ to bind U2 snRNA (Scherly et al., 1990). The U2 snRNA hairpin IV includes a stem consisting of five Watson-Crick basepairs and one non-canonical U-U basepair (Gmeiner and Walberer, 2000; Fig. 1). The U-U basepair occurs at the base of an 11-nucleotide loop. Although the U2A′ protein is essential for U2B″ binding to U2 snRNA hairpin IV, the direct interaction between U2A′ and U2 snRNA hairpin IV is very weak, and is limited to the stem region of the RNA in the x-ray structure (Price et al., 1998; Boelens et al., 1990).

Received for publication 10 July 2000 and in final form 20 April 2001.

Address reprint requests to (present address) Dr. W. H. Gmeiner, Dept. of Biochemistry, Wake Forest University School of Medicine, Medical Center Boulevard, Winston-Salem, NC 27157. Tel.: 336-716-6216; Fax: 336-716-7671; E-mail: bgmeiner@wfubmc.edu.

J.-x. Guo's present address is Camitro Corporation, 4040 Campbell Ave., Menlo Park, CA 94025.

0006-3495/01/08/630/13 $2.00

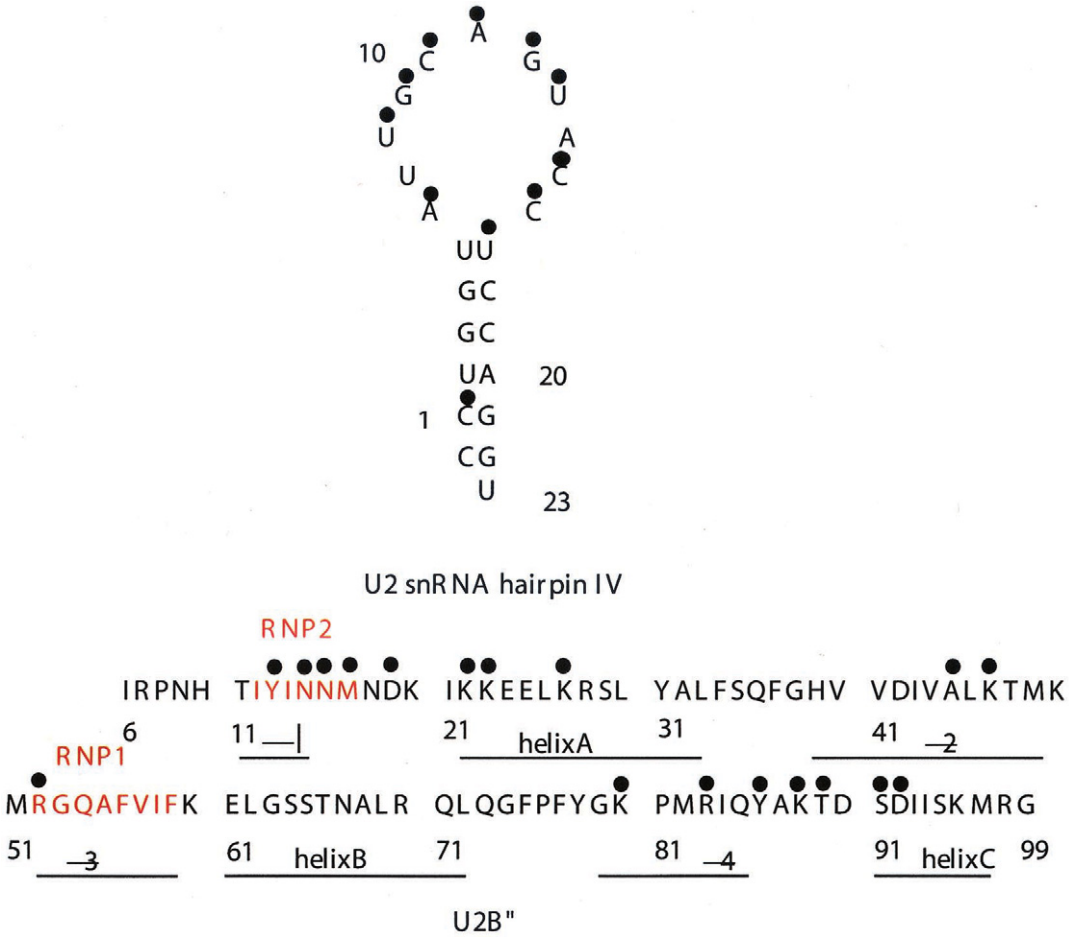

FIGURE 1 Secondary structure and sequence numbering of the U2 snRNA hairpin IV/U2B″ protein complex. The conserved RNP1 and RNP2 sequences for the U2B″ protein are shown in red. The filled circles indicate sites of hydrogen bond formation between protein and RNA.

Understanding the physical basis for formation of the U2 snRNA hairpin IV and U2B″ complex in vivo requires detailed information about complex structure and stability under fully-hydrated conditions. X-ray structures of RNA/protein complexes provide useful structural information; however, x-ray structures may differ from fully hydrated structures. Furthermore, the crystalline state cannot fully reflect the contribution of mobile water molecules to RNA/protein complex stability. Although evidence for some water molecules can be seen in the experimental electron density map, water molecules that are in fast exchange on the surface of RNA/protein complexes, and those present in cavities of the complex, are not well accounted for in x-ray studies (Schwabe, 1997). Molecular dynamics simulations are one of the best methods available to investigate the dynamic behavior of the RNA/protein interface.

In the present study, the interaction between U2B″ protein and U2 snRNA hairpin IV in solution has been investigated using molecular dynamics (MD) simulations to identify the residues of U2B″ that provide the greatest contribution to complex stability (Fig. 2). These MD simulations also provide insight into the role of ordered water in complex stability. To perform MD on a system as large as the hydrated U2B″/U2 snRNA hairpin IV complex, the long-range electrostatic energy of the periodic box was calculated using the particle mesh Ewald (PME) method (Darden et al., 1993; Essman et al., 1995). The MD trajectory and the structure in solution for the complex were analyzed, and the mutant free energy was calculated to determine the contribution from each amino acid in the protein to the stability of the RNA/protein complex. Finally, the location and the residence times of waters residing at the interface of the RNA/protein complex were identified, and the stabilizing interaction for those bridging water molecules that play a significant role in the stability of the RNA/protein complex have been characterized using ab initio calculation. The results indicate that the U2B″/U2 snRNA hairpin IV complex has a novel interaction surface stabilized mainly by electrostatic interactions, and that bridging water molecules contribute significantly to complex stability.

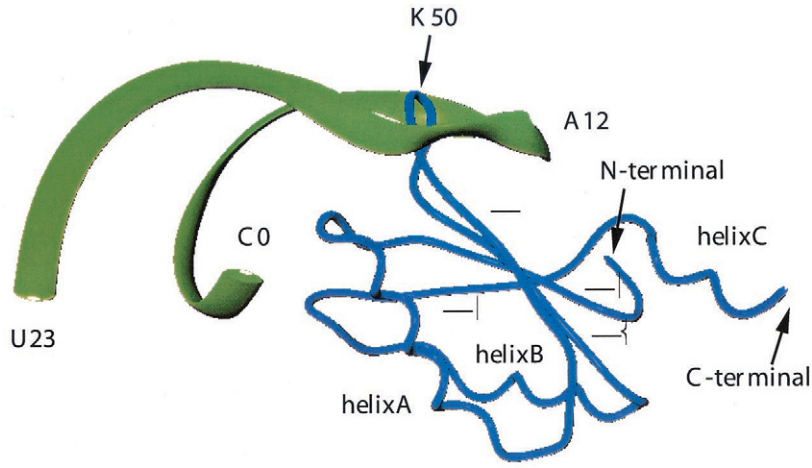

FIGURE 2 Ribbon diagram of the U2 snRNA hairpin IV/U2B″ protein complex. The protein is shown in purple and the RNA in green. Residues that link β2 and β3 of U2B″ (e.g., K50 as shown) protrude into the interior of the RNA loop.

METHODS

The structure of the U2 snRNA hairpin IV/U2B″ protein was obtained from the x-ray crystallographic coordinates at 2.38 Å resolution (Price et al., 1998) in the Brookhaven protein database (identifier 1A9N). Fourteen K^+ ions were added to the RNA/protein complex by using a Coulombic grid potential using the LEAP module of Amber 6.0 to maintain the system at charge neutrality (Case et al., 2000). The system was then solvated using TIP3P water (Jorgensen et al., 1983) in a periodic box. The RNA/protein complex had at least 10 Å buffer in every direction of the box to permit substantial fluctuations of the conformation during the course of the MD simulation. The MD simulation was performed using the AMBER suite of programs, version 6.0 (Case et al., 2000). Energy minimization (500 steps steepest descent followed by 500 steps conjugate gradient) was undertaken before initiating the MD simulation to remove initial steric clashes. The complete interaction energy was calculated. The constant pressure MD simulation was calculated using anisotropic diagonal position scaling. The time step used was 0.002 ps. The temperature of the system was increased gradually from 100 K to 300 K with 20-ps NPT reassemble. The target pressure was 1 atm. The Berendsen algorithm (Berendsen et al., 1984) was used with a scaling factor time constant of 0.2. The Lennard-Jones cutoff value used was 8 Å. SHAKE constraints were applied to all bonds involving hydrogen atoms. The PME (Darden et al., 1993; Essman et al., 1995) was used in all calculations to consider the long-range electrostatic energy with grid spacing around 1.0 Å. The size of the charge grid was chosen to be a product of powers of 2, 3, and 5 for each dimension so that the fast Fourier transform could be applied to increase the speed of the calculation of the reciprocal sum. Finally, the 2200-ps MD simulation was run under the same conditions as the equilibration procedure. The density of the system was maintained near 1 g/cm^3. In all calculations the AMBER94 force field (Cornell et al., 1995, 1996) was used. The center of mass translation was removed periodically (at every restart or 100 ps) during the production dynamics (Cheatham and Kollman, 1997) to avoid energy drains (Harvey et al., 1998; Chiu et al., 2000). The analysis for the trajectory was carried out using Carnal, Ptraj, and in-house scripts (shell, awk, perl) similar to our previous work (Guo and Gmeiner, 2000; Guo et al., 2000).

Mutations were introduced by using hydrogen with a C-H distance of 1.09 Å to replace the side chain of other residues (Xxx→Gly). A detailed description of the method can be found on studies of protein/peptide systems (Massova and Kollman, 1999). The mutant binding free energy was calculated using these same methods. The stable and well-equilibrated final 1800 ps of the 2200-ps trajectory was used for the mutant analysis with snapshots each 5 ps resulting in highly converged solutions with low standard deviations. A total of 360-snapshot conformations were used for every mutant residue system. The internal energy was calculated using the ANAL module in AMBER6. The solvent-accessible surface area was obtained from the program MSMS with a probe radius of 1.4 Å (Sanner et al., 1996). The solvation energy was calculated by using the generalized Born method. (Still et al., 1990; Srinivasan et al., 1998). Previous studies have found that the relative entropic contribution was negligible, and the results are in good agreement with experimental measurement for the mutant free energy between wild-type and mutant complexes in other systems (Reyes and Kollman, 2000a; Massova and Kollman, 2000). In addition, the computational cost for normal mode analysis (NMA) is expensive, considering all snapshots for large systems. Therefore, we did not consider the entropic contribution in the mutant free energy calculations.

The ab initio calculations were carried out using the Gaussian98 program (Frisch et al., 1998). One specific bridging water molecule (water 624), which interacts with cytosine 10 and threonine 89, was calculated as a model system. The starting geometry was obtained from the minimization of the MD averaged structure over the residual periods in the trajectory. The fully geometry optimizations were finished on the Hartree-Fock level with the 6–31G* basis set for the complex, and for each molecule separately. The Becke's three parameter hybrid function using the LYP correlation function (B3LYP) was evaluated for the density functional theory (DFT) calculation (Becke, 1993; Lee et al., 1988). The interaction energy was corrected for basis set superposition errors (BSSE) using the full counterpoise correction (Boys and Bernardi, 1970). The vibrational zero-point energies (ZPE) correction was considered from the frequency calculations without scaling. In the calculation of interaction energy the deformation energy was also taken into account.

RESULTS AND DISCUSSION

Stability of the trajectory

Root-mean-square deviations (RMSD) from the starting structure coordinates during the MD simulation were used to assess the stability and the reliability of the simulation. RMSD from the starting structure over the course of the MD simulation is shown in Fig. 3 for the U2 snRNA hairpin IV/U2B″ complex, and also separately for the RNA and protein components of the complex. Both the RNA and protein components of the com-

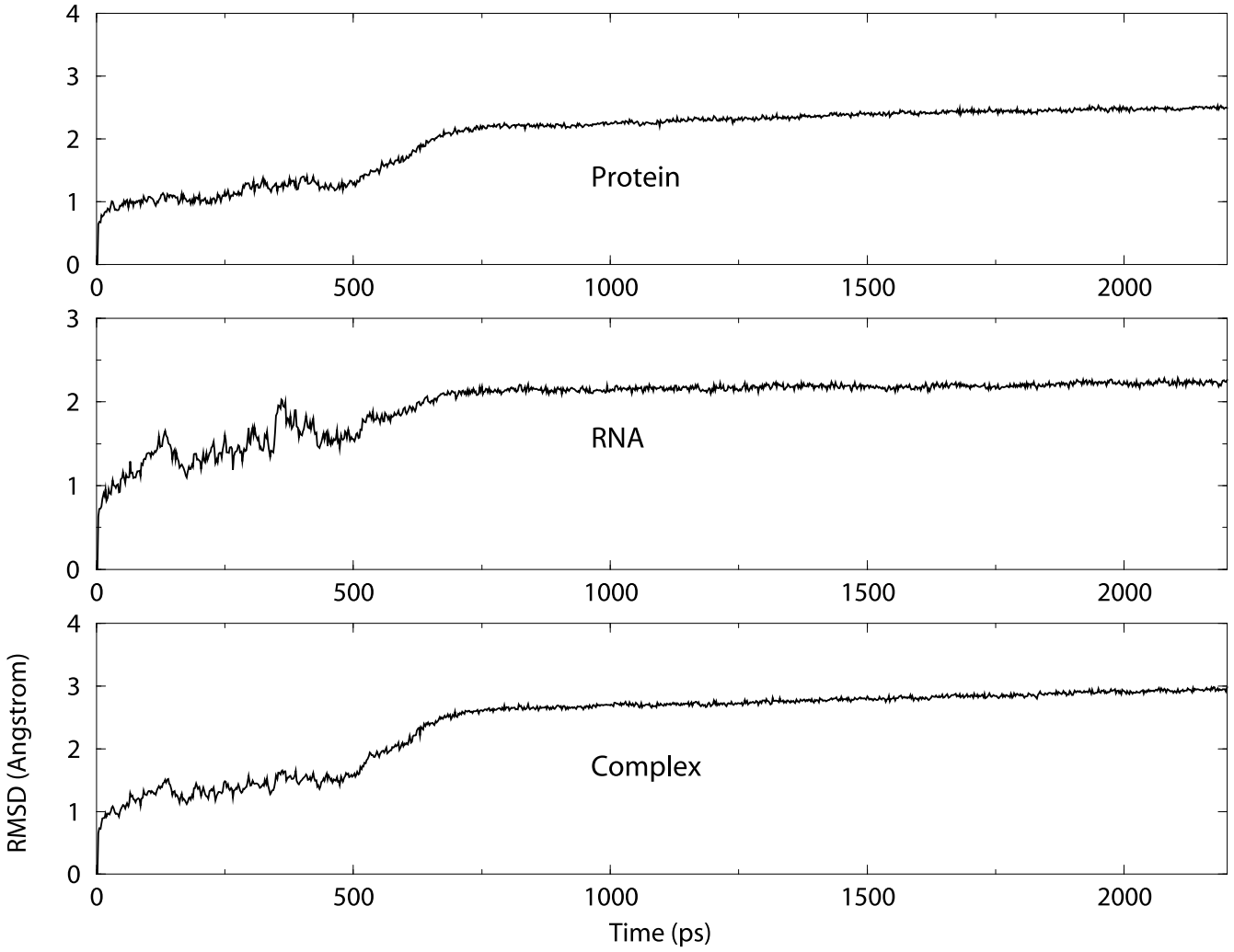

FIGURE 3 RMSD, relative to the initial coordinates, during the 2200-ps molecular dynamics trajectory. RMSD for the U2B″ protein is shown in the top panel, while those for the U2 snRNA hairpin IV and whole system are shown in the middle and bottom panels, respectively. The trajectory is stable after the initial 750 ps of MD.

plex underwent structural fluctuations during the first 500 ps of the trajectory, although the amplitude of the fluctuations was larger for the RNA component. Equilibration of the structure for the complex was achieved after the initial 750 ps of the trajectory. The all-atom RMSD of the U2 snRNA hairpin IV averaged over the remainder of the trajectory was 2.0 ± 0.3 Å, while that for the U2B″ protein was 2.0 ± 0.5 Å. The RMSD of the whole system was 2.4 ± 0.60 Å. The RMSD of the complex during the MD simulation indicates that the x-ray crystal structure did not undergo a large change in aqueous solution.

U2 snRNA hairpin IV

The average values for the pseudorotation phase angle of the ribose sugars (P), the glycosidic bond torsional angle χ, and the torsional angles in the phosphodiester backbone for U2 snRNA hairpin IV over the course of the trajectory were, with several notable exceptions, similar to those characteristic of A-form helical geometry. The pseudorotation phase angles P adopted by the ribonucleotides in the stem region of hairpin IV (C0-U5 and U17-G22), indicated that these ribose sugars uniformly adopted C3′-endo conformations throughout the trajectory. The ribonucleotides in the 5′ portion of the loop region of hairpin IV alternated between C3′-endo and C2′-endo sugar puckers with U7, G9, and A11 adopting C3′-endo sugar puckers and U8, C10, and G12 adopting C2′-endo sugar puckers. U13 also adopted a C2′-endo sugar pucker, breaking the strict alternating pattern, and A14 also adopted a C2′-endo sugar pucker. The UGCA sequence is common to both U1 snRNA hairpin II and U2 snRNA hairpin IV, and these residues were found in the x-ray structures (Price et al., 1998) of both complexes to

be engaged in nearly identical interactions with residues conserved between the U1A and U2B″ proteins.

The loop region of U2 snRNA hairpin IV was stabilized by a number of intramolecular hydrogen-bonding and base-stacking interactions. The 5′ portion of hairpin IV (A6–G12) was stabilized by three pairs of base-stacking interactions. A6 stacked efficiently on U7 and U8 stacked efficiently on G9, but little base overlap occurred between U7 and U8. A single intramolecular hydrogen bond, between U7 O4 and G9 NH_2, stabilized this region of the loop structure. The torsion angles for A6–G9 were near canonical A-form values except for U8, which adopted a C2′-endo sugar pucker and anomalously high values of δ, ϵ, and ζ. A6 in this region also adopted an anomalously low value of β. The two purines, A11 and G12, also stacked efficiently on one another, and this region of the loop was stabilized by a bifurcated hydrogen bond from the 2′-OH of A11 with both O4′ and O5′ of G12. This region of the loop was also stabilized by an intraresidue hydrogen bond between G12 NH_2 and O1P oxygen of this residue. C10 stacked neither with the A11/G12 purine stack nor the U8/G9 stack, and made no intramolecular hydrogen bonds. Rather, C10 was involved in a number of contacts with the U2B″ protein that stabilized the complex (see below).

The pairwise stacking arrangements in the 5′ portion of the loop region (A6–G12) initiated the reverse in directionality of the phosphodiester backbone of the RNA, and the A11/G12 purine stack was roughly perpendicular to the axis of the helical stem region. The four nucleotides in the 3′-portion of the loop formed a "step-ladder structure" that was stabilized by intramolecular base-stacking between A14, C15, and C16. The RNA conformation in this region was also stabilized by intramolecular hydrogen bond formation between O2P of A14 and the N4 amino group of C16, and between A15 O1P and the 2′-OH of U13. Base-stacking did not continue from C16 into the stem region, rather U17, the first residue in the 3′ portion of the stem, region stacked with A6/U7 in the 5′ portion of the loop region. U13 and C16 adopted C2′-endo sugar puckers and anomalously high values of δ and ϵ. A14 adopted a *syn* orientation about χ and the rotation about C4′-C5′ bond for A14 was in the *ap* range, while other nucleotides in the loop were in the $+sc$ range. Obviously, O5′ and the phosphate of A14 were rotated away from the ribose ring. This opened up the nucleotide exposing its sugar, base, and phosphate groups to the surface of the protein to facilitate recognition and binding contacts (Saenger, 1983). Other torsion angles in this region were similar to A-form values.

The non-Watson-Crick U5-U17 basepair (Gmeiner and Walberer, 2000) in the U2B″/U2 snRNA hairpin IV complex closes the RNA loop. This U-U basepair is one of the points of dissimilarity with the U1A/RNA complex, which includes a C-G basepair to close the loop. The average hydrogen bond distances were 2.87 ± 0.11 Å (U5:N3-U17:O4) and 2.96 ± 0.11 Å (U5:O_2-U17:N3) with an ideal planer geometry. The angles are 169.4 ± 6.2° and 167.9 ± 6.6°, respectively. The non-Watson-Crick U-U basepair displayed backbone torsional angles in the same range as those for other ribonucleotides in the stem, except the α torsion angle that is slightly larger than average for U17. The remainder of the stem, except for U2, was nearly canonical A-form duplex with C3′-endo sugar puckers. The rotation about C4′-C5′ bond for U2 is in the *ap* range, extruding the U2 nucleobase, and allowing for U2 to contact the protein.

RNA/protein interaction and mutagenesis

The U2B″ protein recognizes hairpin IV of U2 snRNA using an RNP motif (Uhlenbeck et al., 1997; Nagai, 1996; Burd and Dreyfuss, 1994; Mattaj, 1993). The conserved RNP1 and RNP2 sequences that occur in the two middle strands of a four-stranded anti-parallel β-sheet each contribute one highly stabilizing base-stacking interaction with the RNA (Y13 in RNP2 and F56 in RNP1). The loop region of hairpin IV of U2 snRNA was involved in most of the contacts that stabilized the complex with U2B″. U2B″ mainly interacted with U2 snRNA in three regions. Region 1 involved β1 and helix A (Asn-15–Lys-23). Region 2 involved residues near the C-terminus of β2 and the N-terminus of β3, and those residues that link these β-strands (V44–F56). Region 3 involved residues near the C-terminus of β4 and those residues that link this strand to helix C (Q85–D92). These regions were selected by comparison of the contact area of the U2B″ protein and the U2 snRNA hairpin IV based on the x-ray structure and the structure derived from the MD simulations.

Mutagenesis is an important experimental method used to determine the residues important for macromolecular association and has been used to explore protein/protein (Bogan and Thorn, 1998; Jensen et al., 2000), DNA/protein (Lundquist et al., 1997; O'Neill et al., 1998), and RNA/protein (Gordon et al., 1999; Harada et al., 1998) interfaces. Computational mutagenesis is conceptually similar to experimental mutagenesis; however, it does not require expression of mutant proteins. Rather, the difference in binding free energies between the native complex and a mutant complex of identical structure, except lacking one amino acid side chain, are compared. The mutant free energy, $\Delta\Delta G$, is indicative of the contribution provided by the one amino acid side chain to the stability of the native complex. The mutant free energy consists of the solvation energy, $\Delta\Delta G_{solv}$, and the molecular mechanics energy in vacuo, $\Delta\Delta E_{gas}$

$$\Delta\Delta G = \Delta\Delta E_{gas} + \Delta\Delta G_{solv} \quad (1)$$

The solvation free energy $\Delta\Delta G_{solv}$ was calculated using the general Born approach based on the continuum solvent model (Still et al., 1990), while the nonpolar contribution

TABLE 1 The mutant free energy (kcal/mol) by glycine scanning (β1-helix A)

	Asn-15		Asn-16		Met-17		Asn-18		Asp-19		Lys-20		Ile-21		Lys-22		Lys-23	
Contribution	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
$\Delta\Delta E_{electrostatic}$	−0.16	2.62	6.98	1.16	−5.26	0.38	9.84	0.90	495.84	3.01	−593.09	3.63	1.08	0.11	−539.48	3.65	−559.87	3.22
$\Delta\Delta E_{vdw}$	−4.54	0.61	−0.78	0.69	−2.77	0.24	−0.46	0.03	−4.76	0.72	−4.83	0.85	−0.30	0.02	−2.81	0.78	−2.23	1.24
$\Delta\Delta E_{gas}$	−4.70	2.76	6.21	1.02	−8.03	0.38	9.38	0.91	491.09	2.98	−597.93	3.40	0.78	0.10	−542.29	3.26	−562.10	2.62
$\Delta\Delta G_{nonpolar}$	−0.12	0.02	0.02	0.01	0.25	0.03	0.00	0.00	0.04	0.01	−0.99	0.07	0.00	0.00	−0.98	0.05	−0.42	0.05
$\Delta\Delta G_{GB}$	−3.86	2.69	−9.88	0.72	5.52	0.32	−8.57	0.92	−496.04	2.88	584.99	3.21	1.50	0.25	534.09	2.98	546.04	2.38
$\Delta\Delta G_{subtotal}$	−8.69	0.84	−3.65	0.61	−2.19	0.33	0.81	0.19	−4.91	1.25	−13.93	0.85	2.27	0.21	−9.19	0.59	−16.48	0.97

was calculated from the solvent-accessible surface area (Sanner et al., 1996)

$$\Delta\Delta G_{solv} = \Delta\Delta G_{GB} + \Delta\Delta G_{nonp} \quad (2)$$

The molecular mechanics energy in vacuo, $\Delta\Delta E_{gas,}$ can be obtained from the internal energy $\Delta\Delta E_{int}$ (including bonds, angle, and torsional terms), the electrostatic energy, $\Delta\Delta E_{elec}$, and the van der Waals energy $\Delta\Delta E_{vdw}$,

$$\Delta\Delta E_{gas} = \Delta\Delta E_{int} + \Delta\Delta E_{elec} + \Delta\Delta E_{vdw} \quad (3)$$

Computational glycine mutagenesis was used to identify the amino acids that provided the greatest stabilizing contributions to the RNA/protein complex. The binding free energy (difference between complex and free RNA and protein) for the wild-type and each mutant protein was calculated in terms of the components to the free energy listed in Eqs. 1–3. Here, we assume the unbound RNA and protein have the same conformation as in the complex. In the free energy calculations for the mutated protein, the side chain of the mutated amino acid was replaced by a hydrogen atom (glycine), so that the mutant exclusively reflected the interaction between the side chain of the mutated residue with all other components of the complex. Substantial values for the mutant free energy indicate the side chain plays an important role in complex stability, but does not assure that the side chain is involved in direct RNA/protein interactions. Negligible values of the mutant free energy, however, indicate that the mutated amino acid does not contribute significantly to complex stability, and is unlikely to be important in complex formation. The mutant free energy must be calculated over a substantial region of the MD trajectory during which the structure is stable. In the present case, the final 1.8 ns of the 2.2-ns MD trajectory were used for the mutant free energy analysis.

The mutant free energies for amino acids in the three regions implicated in RNA/protein complex formation are listed in Tables 1–3. Negative values correspond to unfavorable changes in the free energy of the complex as a consequence of the mutation. In region 1, the amino acids with large unfavorable mutant free energies were N15, N16, D19, K20, K22, and K23. N15 had an unfavorable mutant free energy of 8.69 kcal/mol, which derived mainly from loss of a hydrogen bond between N15, R83, and U9 of the RNA (N15:OD1-U9:O2′, N15:OD1-R83:NH2). Similarly, the unfavorable mutant free energies for N16, D19, K20, and K22 resulted from loss of hydrogen bonds. Unfavorable mutant free energies in region 2 resulted mainly from lost van der Waals interactions. For example, V44 mainly interacted with A12 and U14 of the RNA by van der Waals interactions that were absent in the G44 mutant. K47 did not form a hydrogen bond with any nucleotide or residue in the native complex, and the unfavorable mutant free energy for this residue resulted from the loss of electrostatic interactions with C17 and U18 of the RNA. F56 interacted with C11 and stacked with A12 in the native complex; however, these interactions were absent in the G56 mutant. The loss of hydrogen bonds resulted in unfavorable mutant free energies for amino acids in region 3. K88 formed several hydrogen bonds with C11, while S91 formed a hydrogen bond with A12. These hydrogen bonds were one of the most important sources of the large mutant free energy.

Many of the residues with larger mutant free energies were charged residues (e.g., K20, K22, K23, K47, K50, K88, R52), or had a polar side chain (e.g., N15, N16). These results are consistent with the unfavorable mutant free energy arising from loss of an ion pair or hydrogen bond present in the native complex. Similar results have been obtained in protein/protein complexes (Sheinerman et al., 2000). Besides the hydrogen bond and ion pair interaction, charged residues such as Lys and Arg can also form cation-π interactions with the aromatic side chains of proteins or bases of nucleic acids, and loss of these interactions can also result in an unfavorable mutant free energy (Gallivan and Dougherty, 1999).

Hydration of the RNA/protein complex

Nucleotides in the loop region of hairpin IV formed a series of hydrogen bonds with residues in U2B″ that stabilized the complex. Hydrogen bond strength is not readily predictable and can vary between 1 and 5 kcal/mol, depending on the heavy atom pair, internuclear distance, angle, local electrostatic environment, and solvent accessibility (Fersht, 1987; Erion et al., 2000). Direct formation of hydrogen bonds between RNA and protein are summarized in Table 4. The occupied time is the total time that the hydrogen bond

TABLE 2 The mutant free energy (kcal/mol) by glycine scanning (β2-β3)

	Val-44		Ala-45		Leu-46		Lys-47		Thr-48		Met-49		Lys-50		Met-51		Arg-52		Gln-54		Ala-55		Phe-56	
Contribution	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
$\Delta\Delta E_{electrostatic}$	−0.28	0.14	−0.01	0.04	0.10	0.09	−552.72	4.44	3.09	0.50	4.44	0.32	−382.61	2.52	−1.46	0.42	−544.07	3.44	−3.19	0.55	0.01	0.05	−0.14	0.36
$\Delta\Delta E_{vdw}$	−4.41	0.37	−0.20	0.02	−3.91	0.23	−3.91	0.34	−2.46	0.38	−3.79	0.27	−0.19	0.01	−1.80	0.20	−4.58	0.90	−5.07	0.44	−0.92	0.08	−9.73	0.54
$\Delta\Delta E_{gas}$	−4.69	0.40	−0.22	0.04	−3.81	0.24	−556.63	4.64	0.63	0.66	0.65	0.44	−382.80	2.53	−3.25	0.41	−548.65	3.04	8.25	0.63	−0.91	0.07	−9.87	0.70
$\Delta\Delta G_{nonpolar}$	−0.07	0.02	0.00	0.00	−0.23	0.02	0.14	0.13	−0.36	0.01	−0.44	0.03	0.00	0.00	−0.28	0.02	−0.16	0.06	0.01	0.03	0.02	0.00	0.22	0.04
$\Delta\Delta G_{GB}$	1.41	0.24	1.22	0.10	3.02	0.32	551.68	4.50	−1.91	0.55	−4.95	0.44	378.47	2.51	2.55	0.43	534.71	2.61	8.58	0.61	0.88	0.12	−3.48	0.32
$\Delta\Delta G_{subtotal}$	−3.36	0.44	1.01	0.10	−1.02	0.33	−4.80	0.59	−1.63	0.51	−4.74	0.37	−4.33	0.04	−0.99	0.19	−14.11	0.97	0.34	0.68	0.00	0.13	−13.12	0.63

existed during the 2200-ps MD simulation, and is indicative of hydrogen bond stability. Only hydrogen bonds with residence times longer than 1100 ps are included in the table. In Table 4 there are some hydrogen bonds listed that involve backbone atoms; for example, the N4 amino group of C10 formed hydrogen bonds with the backbone amide oxygen of Y86. O2 and N3 of C10 formed a bifurcated hydrogen bond with the backbone amide proton of K88, and this hydrogen bond persisted during the course of the entire MD trajectory of 2200 ps.

Formation of hydrogen bonds and salt bridges between protein residues and nucleotides in the 3′ portion of the loop also showed the role played by the ion pair and polar residues in the RNA/protein complex. The number of hydrogen bonds and salt bridges formed between U2B″ and this region of RNA was limited, although several hydrophobic contacts between RNA and protein stabilized this region of the complex. The salt bridge formed between the ϵ amino group of K23 and O2P of C16 was, however, one of the RNA/protein interactions sustained for >2000 ps during the MD simulation (Table 4). The R52 guanidinium group also formed a salt bridge with U17 O1P stabilizing the complex in this region. Formation of these salt bridges reduced the occupancy of water molecules associated with the phosphodiester backbone at these positions. One other salt bridge, involving the phosphodiester at C1 in the stem region and K22, also stabilized the complex. A number of residues in the β1-helix A of B2″, including K22, also form hydrogen bonds with nucleotides in the stem region of hairpin IV, stabilizing the complex.

Water at the RNA/protein interface

The stability of the RNA/protein complex is derived not only from direct interactions between protein and RNA, but also from interfacial water molecules. Water in the U2B″/U2 snRNA hairpin IV complex can be divided into three types: 1) rapidly exchanging bulk water; 2) surface water forming the first solvation shell; and 3) bridged water molecules that form hydrogen bonds simultaneously to both RNA and protein. Bridging water molecules frequently reside in the cavities of macromolecular complexes for relatively long durations. Several studies have described the occurrence of bridging water molecules at the DNA/protein interface (Schwabe, 1997). Relatively few studies on the importance of bridging water molecules in stabilization of RNA/protein complexes have been described, however, due in part to the lack of structural data for RNA/protein complexes.

Water molecules that form hydrogen bonds simultaneously with both U2 snRNA hairpin IV and U2B″ in the RNA/protein complex (bridging water molecules) contribute significantly to complex stability. The criteria used for hydrogen bonding of water molecule was $X \cdots Y < 4.0$ Å between hydrogen bond donating atoms in RNA (X), and

TABLE 3 The mutant free energy (kcal/mol) by glycine scanning (β4-helix C)

	Gln-85		Tyr-86		Ala-87		Lys-88		Thr-89		Asp-90		Ser-91		Asp-92	
Contribution	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
$\Delta\Delta E_{electrostatic}$	9.23	0.92	−0.24	0.23	0.05	0.10	−415.00	2.86	−2.19	1.13	266.40	1.49	−14.35	1.23	310.08	2.86
$\Delta\Delta E_{vdw}$	−0.35	0.05	−0.07	0.01	−0.59	0.15	−3.13	0.87	−0.76	0.49	−0.09	0.01	0.29	0.97	−1.53	0.23
$\Delta\Delta E_{gas}$	8.89	0.93	−0.31	0.23	−0.54	0.17	−418.13	2.84	−2.94	0.98	266.32	1.49	−14.05	0.66	308.55	2.87
$\Delta\Delta G_{nonpolar}$	−0.10	0.02	0.00	0.00	0.03	0.05	−0.64	0.04	−0.06	0.03	0.00	0.01	0.10	0.01	−0.13	0.02
$\Delta\Delta G_{GB}$	−8.17	0.91	0.57	0.24	0.42	0.12	409.93	2.86	0.80	0.75	−263.15	1.49	9.05	0.73	−306.52	2.86
$\Delta\Delta G_{subtotal}$	0.62	0.06	0.26	0.05	−0.09	0.18	−8.83	0.81	−2.20	0.58	3.17	0.05	−4.90	0.61	1.90	0.24

protein (Y) and an X · · · H · · · Y angle >120°. To identify those water molecules that were truly involved in bridging hydrogen bond formation, only those water molecules that simultaneously bonded with both RNA and protein were considered as bridging water molecules. During the course of the 2200-ps MD trajectory, there were 1153 water molecules with residence times >1 ps within the inner solvation shells of both the RNA and the protein. A total of 45 bridging water molecules were identified in the trajectory for the U2B″:U2 snRNA hairpin IV complex, all with residence times longer than 200 ps (Table 5). Additional bridging water molecules that underwent fast exchange were also identified at the RNA/protein interface. The bridging water molecule having the longest residence time in the trajectory formed hydrogen bonds with O2P of C16 on U2 snRNA hairpin IV (C16:O2P) and the side chain of E24 in U2B″ (E24:OE2). Distance fluctuations for two of the most persistent interfaced water molecules are shown in Fig. 4. Water 624 had large fluctuations of position during the initial several hundred picoseconds of the trajectory, but was stable during the remainder of the trajectory. In contrast, water 2115 was stable nearly from the beginning of the trajectory.

Due to the high mobility of most waters solvating the complex, it is not too surprising to note that different water molecules can reside in the same position of the RNA/protein interface at different times during the trajectory. Conversely, a single water molecule can reside at different sites in the complex during the trajectory. For example, WAT258 formed a hydrogen bond with the side chain of D42 and the phosphate oxygen of U13 of duration 66 ps. Later in the trajectory, this same water molecule formed a hydrogen bond with the side chain of D92 and the sugar ring of G12 of 14-ps duration. Both WAT277 and WAT623 were found to hydrogen-bond to the base of A11 and the side chain of T89 at different instants during the trajectory. Thus, hydration of the RNA/protein interface is dynamic, with multiple water molecules contributing to the stability of the complex throughout the trajectory.

Defining the number of water molecules that reside at the RNA/protein interface is difficult because this number is time-dependent. A simple estimate is to use the time average over the trajectory. i.e.,

$$N = \frac{1}{\Delta T}\sum_{i}\Delta t_i \tag{4}$$

where Δt_i is the residence time of each bridging water and ΔT is the trajectory time, which is 2200 ps in the present simulation. Using this approach, the average number of bridging water molecules residing at the RNA/protein interface at any instant is 20 (considering only the water with residence time >200 ps). Because in this estimate only water molecules that resided for at least 200 ps were con-

TABLE 4 Observed hydrogen bonds in the U2B″:U2 snRNA hairpin IV complex

RNA	Protein	Time*	RNA	Protein	Time
C1:O1P	K22:NZ	1872	A11:N6	T89:OG1	2130
A6:N6	D19:OD1	2156	A11:N6	S91:OG	2172
U8:N3	N16:OD1	2200	A11:N1	S91:OG	2200
U8:O2′	N15:OD1	1576	G12:O6	D92:N	2186
U8:O4	K80:NZ	1704	U13:N3	A45:O	2198
U8:O2′	R83:NH2	1806	U13:O4	A45:N	2200
G9:O6	M17:N	1608	U13:O4	K27:NZ	1696
G9:O6	N16:N	2200	C15:O2P	K23:NZ	2052
C10:O4′	Y13:OH	1606	C16:O2′	K47:O	2066
C10:O5′	Y13:OH	1548	C16:O2P	K23:N	1852
C10:N4	Y86:O	2058	C16:O1P	K23:NZ	2198
C10:O2	K88:N	2200	U17:O5′	R52:NH1	1800
C10:N3	K88:N	2200	U17:O1P	R52:NH1	2190
C10:O2′	K88:NZ	1866			

*Indicates time in picoseconds the hydrogen bond persisted during the 2200 ps MD trajectory.

TABLE 5 Bridging water molecules in the U2B″:U4 snRNA hairpin IV complex

RNA	WAT	Protein	Time*	RNA	WAT	Protein	Time
U13:O2′	1227	T48:OG1	214	C10:O2P	1853	K88:NZ	824
U13:O2	1313	K47:N	352	U8:N3	3031	R83:NH1	962
U7:O4	2196	K80:NZ	246	A11:N7	968	T89:OG1	836
A6:O4′	2317	E25:OG1	226	C10:O2′	1344	T89:OG1	1116
U2:O4	2416	K20:NZ	264	C16:O2	2082	M49:N	1194
U8:O2′	2813	R83:NE	214	G9:O1P	1064	I14:O	1576
A11:O1P	3609	K88:NZ	350	U8:O2′	1102	R83:NH2	1584
U8:O2′	3647	R83:NE	208	C1:O1P	1112	K22:NZ	1666
U5:O4	3732	K20:NZ	208	C10:O1P	1331	K88:NZ	1656
G4:N7	3825	K20:NZ	385	C12:O1P	1375	T48:OG1	1662
U7:O4	1309	N18:OD1	422	C1:O2P	1546	K22:NZ	1698
G9:O4′	1524	R52:O	466	G12:O6	1929	D90:O	1748
U2:O1P	1948	K20:NZ	442	U8:O2P	2345	E24:OE1	1644
G12:O2′	2626	D61:OD2	410	U7:O2	2407	M17:O	1672
U13:O2′	2932	E24:OE1	584	G22:O2P	2791	R7:N	1694
G9:O2P	895	N15:OD1	592	G12:O2′	342	D92:OD2	1612
U17:O2P	1664	K47:NZ	746	C16:O1P	1303	K23:NZ	1900
C10:O2P	2201	Q54:NE2	710	C16:O2P	2008	K23:N	1906
U7:O2	2479	D19:OD1	698	U17:O2P	2167	R52:NH2	1896
C16:O2′	2628	M49:N	722	C16:O2P	2115	E24:OE2	2088
C1:O2P	2704	E25:OE1	644	C10:O2	624	T89:OG1	2030
U2:O1P	3092	K20:NZ	708	U5:O4	3593	N18:OD1	610
U5:O4	3522	K20:NZ	668				

*Duration (ps) the bridged water persisted during the 2200 ps MD trajectory.

sidered, the estimate is the lower limit of the number of bridging water molecules that occurred during this simulation. The average value of the maximum observed residence time of water molecules (Qian et al., 1993; Sunnerhagen et al., 1998) residing longer than 200 ps during the MD simulation was 979 ps.

It is also interesting to note that most residues with a large mutant free energy can hydrogen-bond with RNA, either directly or indirectly, through bridged water, or both (Tables 4 and 5). For example, OD1 of N15 hydrogen-bond with O_2' of U8 on RNA for 1576 ps, and also indirectly interact with O2P of G9 through bridging water WAT895 for 592 ps. NZ of K20 does not form hydrogen bonds with RNA that persist for long times, but three bridging water molecules (WAT2416, WAT3092, and WAT3522) do hydrogen-bond with O4 of U5, O1P of U2, and O4 of U5. Mutation at this residue would cause a reduction in the number of hydrogen bonds between RNA/protein and protein/bridging water, as well as the direct interaction.

Interaction energy associated with water

Based on the above analysis of the MD trajectory, water molecules reside on the interface of the RNA/protein complex. However, it is difficult to calculate the free energy associated with the bridging water molecules directly from the MD calculation in this case. In this paper, the interaction energy associated with those residual water molecules was estimated using ab initio calculations for a model system. Water molecule 624, associated with Cyt-10 and Thr-89, as identified from the analysis above, were used as a model system for the ab initio calculations (Table 6). The geometry of the starting complex was taken from the minimized structure for the complex averaged over the residence time during the trajectory. The electronic structure was fully optimized at the HF/6–31G* level. The optimized geometries of the Cyt-10 and Thr-89 were taken directly for the single-point calculations.

The interaction energy ΔE can be understood as the sum of pairwise dimer contributions and a three-body term, ΔE^3, which accounts for cooperative effects (Sponer et al., 1997; Brandl et al., 1999, 2000) in the water 624/C10/T89 system.

$$\Delta E^{\text{water624/C10/T89}} = \Delta E^{\text{water624/C10}} + \Delta E^{\text{water624/T89}} + \Delta E^{\text{C10/T89}} + \Delta E^3 \quad (5)$$

The direct interaction between C10 and T89 without water is straightforward with consideration of the BSSE correction. The interaction energy was further corrected by consideration of the deformation energy and the zero-point energy. The deformation energy in the C10/T89 dimer is 0.7 kcal/mol for both HF and DFT calculations. The zero-point energy correction is 0.4 kcal/mol for HF and 0.3 kcal/mol for the DFT calculation. The calculated interaction energy of 1.3 kcal/mol (HF) and 1.6 kcal/mol (DFT) is much smaller than the value in the water-mediated trimer (Table 6). The total deformation energy from each component in the trimer is 0.8 kcal/mol (HF) and 1.0 kcal/mol (DFT). The

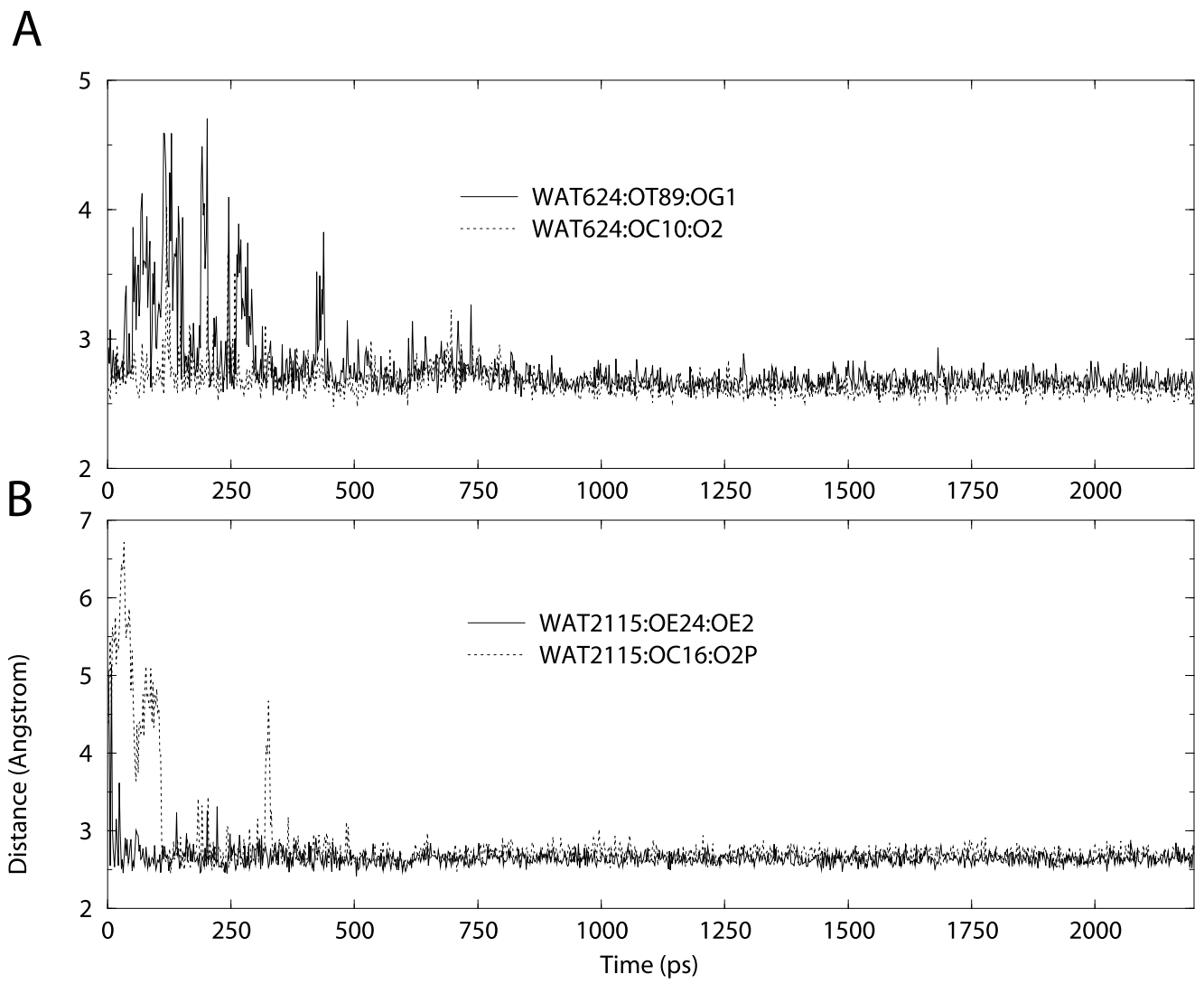

FIGURE 4 The distance between the bridging water molecule and U7:O4 and D19:OD1 during the course of the 2200-ps trajectory. The structure of this bridging water molecule was remarkably stable during the trajectory. (*A*) The residence time is 1800–2000 ps. (*B*) The residence time is 2000–2200 ps.

zero-point energy is increased from 0.4 kcal/mol (dimer) to 4.1 kcal/mol (trimer) in the HF level calculation, and from 0.3 kcal/mol (dimer) to 2.6 kcal/mol (trimer) in the DFT level calculation. The ΔE^3 value of −0.3 kcal/mol (HF) and −2.3 kcal/mol (DFT) indicated that DFT indicated a stronger cooperativity than HF calculation in the water-mediated trimer. The difference of the 7.2 kcal/mol (HF) and the 10.5 kcal/mol (DFT) between the dimer and trimer is the stabilization energy associated with the water-mediated trimer compared with the direct interaction in the dimer (Table 6). The stabilization energy is derived from the additional hydrogen bonds due to the bridged water and orbital interac-

TABLE 6 The interactions energies (kcal/mol) for water-mediated (C10/T89/WAT624) and the corresponding direct (C10/T89) calculated according to the HF and DFT methods (X = C10, Y = T89, Z = WAT624)

	ΔE^{xy}	ΔE^{xz}	ΔE^{yz}	ΔE^3	ΔE	$\Delta E^{DEF(X)}$	$\Delta E^{DEF(Y)}$	$\Delta E^{DEF(Z)}$	ΔE^{ZPE}	ΔE_0*
C10/T89/WAT624										
HF	−2.4	−6.0	−4.7	−0.3	−13.4	0.2	0.5	0.1	4.1	−8.5
DFT	−2.6	−5.9	−4.9	−2.3	−15.7	0.1	0.6	0.3	2.6	−12.1
C10/T89										
HF	−2.4				−2.4	0.2	0.5		0.4	−1.3
DFT	−2.6				−2.6	0.1	0.6		0.3	−1.6

*E_0 is the interaction energy considering the deformation energy and vibrational zero-point energy correction.

tions among those three components in the trimer. The results of the ab initio calculation indicated that water 624 greatly stabilized the direct interaction between C10 and T89. To explore the overall effects from all bridged waters at the protein/RNA interface, it would be necessary to do a statistical average for the total stabilizing energy including all bridging waters. However, considering the large number of the bridged water in the MD trajectory, it would be time-consuming to complete calculations at the ab initio level for this system. Therefore, the ab initio calculations reported here do not provide a definite value for the entire stabilization energy from all of the bridged waters in the RNA/protein complex, rather they only provide a possible range for the stabilization interaction energy for each water-associated trimer at the RNA/protein interface. We expect future studies to provide additional understanding of this kind of interaction.

Comparison with U1A′/U1 snRNA hairpin II

The molecular dynamics of the U1A′/U1 snRNA hairpin II complex (Reyes and Kollman, 1999) were investigated by Reyes and Kollman using both alanine scanning mutagenesis (Reyes and Kollman, 2000a), and binding thermodynamics (Reyes and Kollman, 2000b). The computational results from these studies were in excellent agreement with the available experimental results. The loop structure of U1 snRNA hairpin II is very similar to U2 snRNA hairpin IV, the subject of the present investigation. The major differences between these stem loops are that C12 in U1 snRNA hairpin II was replaced by the G12 in U2 snRNA hairpin IV, and there is one more nucleotide (A14) in U2 snRNA hairpin IV. Although U2 snRNA hairpin IV binds U2B″ in an orientation similar to the way U1 snRNA binds to the U1A′ protein, the details of the interaction differ in several ways (Price et al., 1998). In the U1A′/RNA complex, the most important residues for binding were in the β2-β3 loop and RNP regions, which is similar for U2B″/U2 snRNA hairpin IV. The other two interaction areas (β1-helix A and β4-helix C) in U2B″ were not, however, observed in the U1A′ protein complex.

MD simulations of the U1/U1A′ complex RNA/protein complex also analyzed hydration of the complex and the role of bridging water molecules in complex stability (Tang and Nilsson, 1999). In this RNA/protein complex, which is highly homologous to the U2B″/U2 snRNA hairpin IV complex considered in the present study, the lifetime of the bridging water molecules identified was typically $<$20.0 ps (Tang and Nilsson, 1999). It is surprising to note such a large difference in the stabilities of bridging waters between these two complexes. One possible reason for the shorter residence time for bridging water molecules in the previous study compared to the present results may derive from differences in computational methods. In the present study, PME has been used to properly account for long-range electrostatic interactions in the RNA/protein complex, especially hydrogen bond formation. Appropriate accounting of long-range electrostatic interactions is very important for deriving meaningful results for MD simulations involving nucleic acids, which should routinely be considered using the PME sum (Essman et al., 1995). Other studies also observed residence time as long as 5.1 ns in zipper-like DNA (Spackova et al., 2000).

CONCLUSIONS

In summary, a 2200-ps MD simulation of the solvated U2 snRNA hairpin IV/U2B″ protein complex was completed and the trajectory has been analyzed. The MD simulation in water provided several insights into the factors that provided important contributions to complex stability. Analysis of the mutant free energies indicated the most important regions contributing to the stability of the RNA/protein complex. Charged and polar residues played an important role due to favorable electrostatic interactions. For example, several lysine residues showed a large mutant free energy in the RNA/protein interface that arose due to interaction with the phosphodiester backbone of the RNA. The formation of direct hydrogen bonds and salt bridges between the RNA and protein contributed to complex stability. The overall interaction between U2 snRNA hairpin IV/U2B protein differed from the U1 snRNA/U1A RNA/protein complex, although the two complexes had similar secondary structures; 1153 waters were observed in bridging between RNA and protein in this 2200-ps trajectory. Only 45 of them had a residence time $>$200 ps ($\sim$9% of the trajectory). Other bridge waters exchanged rapidly at the interface. On average, 20 bridging water molecules occurred at the interface of the complex, and the average residence time of those bridged waters was 979 ps. Calculation for the interaction energy resulting from bridging water molecules was determined for the C10/T89 interaction using *ab initio* calculations. These results indicate that bridged water play an important role in stabilization of the RNA/protein interface.

Computation was supported by the research computing facility of the University of Nebraska-Lincoln, National Computational Science Alliance (MCB990023N) in the National Center for Supercomputing Application at the University of Illinois at Urbana-Champaign, and National Cancer Institute's Advanced Biomedical Computing Center (991108JG46). J. Guo thanks Prof. Stuchebrukhov at UC, Davis for kindly providing computers for part of the ab initio calculations. This work was supported by National Institutes of Health-NCI Grant 60612 (to W.H.G.) and National Institutes of Health-NCI Grant 36727.

REFERENCES

Antson, A. A. 2000. Single stranded RNA binding proteins. *Curr. Opin. Struct. Biol.* 10:87–94.

Becke, A. D. 1993. Density-functional thermochemistry. 3. The role of exact exchange. *J. Chem. Phys.* 98:5648–5652.

Berendsen, H. J. C., J. P. M. Postma, W. F. V. Guntersen, A. Dinola, and J. R. Haak. 1984. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* 81:3684–3690.

Boelens, W., D. Scherly, R. P. Bejer, E. J. R. Jansen, N. A. Dathan, I. W. Mattaj, and W. J. Venrooij. 1990. A weak interaction between the U2A′ protein and U2 snRNA helps to stabilize their complex with the U2B″ protein. *Nucleic Acids Res.* 19:455–460.

Bogan, A. A., and K. S. Thorn. 1998. Anatomy of hot spots in protein interfaces. *J. Mol. Biol.* 280:1–9.

Boys, S. F., and F. Bernardi. 1970. The calculations of small molecular interaction by the difference of separate total energies. Some procedures with reduced error. *Mol. Phys.* 19:553–566.

Brandl, M., M. Meyer, and J. Suhnel. 1999. Quantum-chemical study of a water-mediated uracil-cytosine basepair. *J. Am. Chem. Soc.* 121: 2605–2606.

Brandl, M., M. Meyer, and J. Suhnel. 2000. Water-mediated basepairs in RNA: a quantum-chemical study. *J. Phys. Chem.* 104:11177–11187.

Burd, C. G., and G. Dreyfuss. 1994. Conserved structures and diversity of functions of RNA-binding proteins. *Science.* 265:615–621.

Case, D. A., D. A. Pearlman, J. W. Caldwell, T. E. Cheatham III, W. S. Ross, C. L. Simmerling, T. A. Darden, K. M. Merz, R. V. Stanton, A. L. Cheng, J. J. Vincent, M. Crowley, D. M. Ferguson, R. J. Radmer, G. L. Seibel, U. C. Singh, P. K. Weiner, and P. A. Kollman. 2000. AMBER 6. University of California, San Francisco.

Cheatham, T. E. III, and P. A. Kollman. 1997. Molecular dynamics simulations highlight the structural differences among DNA:DNA, RNA:RNA, and DNA:RNA hybrid duplexes. *J. Am. Chem. Soc.* 119: 4805–4825.

Chiu, S. W., M. Clark, S. Subramaniam, and E. Jakobsson. 2000. Collective motion artifacts arising in long-duration molecular dynamics simulations. *J. Comp. Chem.* 21:121–131.

Cornell, W. D., P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman. 1995. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* 117: 5179–5197.

Cornell, W. D., P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman. 1996. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* 117:2309.

Cusack, S. 1999. RNA-protein complexes. *Curr. Opin. Struct. Biol.* 9:66–73.

Darden, T., D. York, and L. Pedersen. 1993. Particle mesh Ewald: an N.log(N) method for Ewald sums in large systems. *J. Chem. Phys.* 98:10089–10092.

Erion, M. D., P. D. Van Poelje, and M. R. Reddy. 2000. Computer-assisted scanning of ligand interactions: analysis of the fructose 1,6-bisphosphatase-amp complex using free energy calculations. *J. Am. Chem. Soc.* 122:6114–6115.

Essman, U., L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen. 1995. A smooth particle mesh Ewald method. *J. Chem. Phys.* 103:8577–8593.

Fersht, A. R. 1987. The hydrogen bond in molecular recognition. *Trends Biochem. Sci.* 12:301–304.

Frisch, M. J., G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M. W. Gill, B. G. Johnson, W. Chen, M. W. Wong, J. L. Andres, M. Head-Gordon, E. S. Replogle, and J. A. Pople. 1998. Gaussian 98, Gaussian, Inc., Pittsburgh, PA.

Gallivan, J., and D. A. Dougherty. 1999. Cation-π interactions in structural biology. *Proc. Natl. Acad. Sci. U.S.A.* 96:9459–9464.

Gmeiner, W. H., and B. J. Walberer. 2000. Basepairing in DNA: unusual patterns. Encyclopedia of Life Sciences, Macmillan Reference Ltd., London.

Gordon, J., T. K. Sengupta, C. A. Phillips, S. M. O'Malley, K. R. Williams, and E. K. Spice. 1999. Identification of the RNA binding domain of T4 RegA protein by structure-based mutagenesis. *J. Biol. Chem.* 274: 32265–32273.

Guo, J. X., I. Daizadeh, and W. H. Gmeiner. 2000. Molecular dynamics simulation and the structure of the U4 snRNA sm site. *J. Biomol. Struct. Dyn.* 18:335–344.

Guo, J. X., and W. H. Gmeiner. 2000. A free energy comparison of the human Sm Protein complexes B/D_3 and D_3/B. *J. Am. Chem. Soc.* 122:9926–9932.

Harada, N., K. Maemura, N. Yamasaki, and M. Kimura. 1998. Identification by site-directed mutagenesis of amino acid residues in ribosomal protein L2 that are essential for binding to 23S ribosomal RNA. *Biochim. Biophys. Acta.* 1429:176–186.

Harvey, S. C., R. K. Z. Tan, and T. E. Cheatham. 1998. The flying ice cube: velocity rescaling in molecular dynamics leads to violation of energy equipartition. *J. Comp. Chem.* 18:726–740.

Honda, N., O. Nureki, K. Kurimoto, I. Kim, H. Sakamoto, Y. Shimura, Y. Muto, and S. Yokoyama. 1999. Structural basis for recognition of the *tra* mRNA precursor by the sex-lethal protein. *Nature.* 398:579–585.

Jensen, A. A., T. A. Spalding, E. S. Burstein, P. O. Sheppard, P. J. O'Hara, M. R. Brann, P. Krogsgaard-Larsen, and H. Brauner-Osborne. 2000. Functional importance of the Ala116-Pro136 region in the calcium-sensing receptor. Constitutive activity and inverse agonism in a family C G-protein coupled receptor. *J. Biol. Chem.* 276:10110–10118.

Jorgensen, W. L., J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926–935.

Lee, C., W. Yang, and R. G. Parr. 1988. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B.* 37:785–789.

Lundquist, A. J., R. D. Beger, S. E. Bennett, P. H. Bolton, and D. W. Mosbaugh. 1997. Site-directed mutagenesis and characterization of uracil-DNA glycosylase inhibitor protein—role of specific carboxylic amino acids in complex formation with *Escherichia coli* uracil-DNA glycosylase. *J. Biol. Chem.* 272:21408–21419.

Massova, I., and P. A. Kollman. 1999. Computational alanine scanning to probe protein-protein interactions: a novel approach to evaluate binding free energies. *J. Am. Chem. Soc.* 121:8138–8143.

Massova, I., and P. A. Kollman. 2000. Combined molecular mechanical and continuum solvent approach (MM-PBSA/GBSA) to predict ligand binding. *Persp. Drug Disc. Des.* 18:113–135.

Mattaj, I. W. 1993. RNA recognition: a family matter? *Cell.* 73:837–840.

Nagai, K. 1996. RNA-protein complexes. *Curr. Opin. Struct. Biol.* 6:53–61.

Nolan, S. J., J. C. Shields, J. B. Tuite, K. L. Cecere, and A. M. Baranger. 1999. Recognition of an essential adenine at a protein-RNA interface: comparison of the contributions of hydrogen bonds and a stacking interaction. *J. Am. Chem.* Soc. 121:8951–8952.

O'Neill, M., D. T. F. Dryden, and N. E. Murray. 1998. Localization of a protein-DNA interface by random mutagenesis. *EMBO J.* 17: 7118–7127.

Oubridge, C., N. Ito, P. R. Evans, C. H. Teo, and K. Nagai. 1994. Crystal structure at 1.92 Å resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. *Nature.* 372: 432–438.

Price, S. R., P. R. Evans, and K. Nagai. 1998. Crystal structure of the spliceosomal U2B″-U2A protein complex bound to a fragment of U2 small nuclear RNA. *Nature.* 394:645–650.

Qian, Y. Q., G. Otting, and K. Wuthrich. 1993. NMR detection of hydration water in the intermolecular interface of a protein-DNA complex. *J. Am. Chem. Soc.* 115:1189–1190.

Reyes, C. M., and P. A. Kollman. 1999. Molecular dynamics studies of U1A-RNA complexes. *RNA.* 5:235–244.

Reyes, C. M., and P. A. Kollman. 2000a. Investigating the binding specificity of U1A-RNA by computational mutagenesis. *J. Mol. Biol.* 295:1–6.

Reyes, C. M., and P. A. Kollman. 2000b. Structure and thermodynamics of RNA-protein binding: using molecular dynamics and free energy analyses to calculate the free energies of binding and conformational change. *J. Mol. Biol.* 297:1145–1158.

Saenger, W. 1983. Principles of Nucleic Acid Structure. Springer-Verlag, New York.

Sanner, M. F., A. J. Olson, and J. C. Spehner. 1996. Reduced surface: an efficient way to compute molecular surfaces. *Biopolymers.* 38:305–320.

Scherly, D., W. Boelens, N. A. Dathan, W. J. V. Venrooij, and I. W. Mattaj. 1990. Major determinants of the specificity of interaction between small nuclear ribonucleoproteins U1A and U2B″ and their cognate RNAs. *Nature.* 345:502–506.

Schwabe, J. W. R. 1997. The role of water in protein DNA interactions. *Curr. Opin. Struct. Biol.* 7:126–134.

Sheinerman, F. B., R. Norel, and B. Honig. 2000. Electrostatic aspects of protein-protein interactions. *Curr. Opin. Struct. Biol.* 10:153–159.

Spackova, N., Berger, I., and J. Sponor. 2000. Nanosecond molecular dynamics of zipper-like DNA duplex structure containing sheared G.A mismatch pairs. *J. Am. Chem. Soc.* 122:7564–7572.

Sponer, J., J. V. Burda, P. Mejzlik, J. Leszczynski, and P. Hobza. 1997. Hydrogen-bonded trimers of DNA bases and their interaction with metal cations: Ab initio quantum-chemical and empirical potential study. *J. Biomol. Struct. Dyn.* 14:613–628.

Srinivasan, J., T. E. Cheatham, P. Cieplak, P. A. Kollman, and D. A. Case. 1998. Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate-DNA helices. *J. Am. Chem. Soc.* 120:9401–9409.

Still, W. C., A. Tempczyk, R. C. Hawley, and T. Hendrickson. 1990. Semi-analytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.* 112:6127–6129.

Sunnerhagen, M., V. P. Denisov, K. Venu, A. Bonvin, J. Carey, B. Halle, and G. Otting. 1998. Water molecules in DNA recognition 1: hydration lifetimes of trp operator DNA in solution measured by NMR spectroscopy. *J. Mol. Biol.* 282:847–858.

Tang, Y., and L. Nilsson. 1999. Molecular dynamics simulations of the complex between human U1A protein and hairpin II of U1 small nuclear RNA and of free RNA in solution. *Biophys. J.* 77: 1284–1305.

Uhlenbeck, O. C., A. Pardi, and J. Feigon. 1997. RNA structure comes of age. *Cell.* 90:833–840.