A Free Energy Comparison of the Human Sm Protein Complexes B/D_3 and D_3/B

Jian-xin Guo[†] and William H. Gmeiner*[,‡]

Contribution from the Eppley Institute for Cancer Research, University of Nebraska Medical Center, Omaha, Nebraska 68198

Received February 3, 2000. Revised Manuscript Received August 11, 2000

Abstract: Molecular dynamics simulations were carried out for the human Sm proteins B and D_3, and two different configurations of their complexes, B/D_3 and D_3/B, to investigate the structures of possible sub-complexes used for assembly of the ring structure comprised of the seven Sm proteins. Comparison of the free energy for the two sub-complexes revealed that the D_3/B configuration, in which the β4 strand of B interacts with the β5 strand of D_3 protein, is energetically preferred relative to the B/D_3 configuration. The direct interaction energy in vacuo and the solvation energy provided the largest contributions to the free energy. The torsional low-frequency vibration modes provided the largest contribution to the entropic terms due to complex formation, although the single mode also gave a small contribution. Determining the configuration of the D_3/B sub-complex provides powerful support for a proposed binding mechanism of the Sm proteins on snRNA.

Introduction

Pre-mRNA splicing is the process of removing the intronic information from primary transcripts of eukaryotic genes (pre-mRNA).[1] Pre-mRNA splicing occurs in large nuclear RNA:protein complexes called spliceosomes. Although spliceosome formation and pre-mRNA splicing occur in the nucleus, each of the principal snRNAs involved in pre-mRNA splicing, with the exception of U6 snRNA, undergoes several steps of processing prior to spliceosome assembly.[2] Maturation of these snRNAs involves modification of nucleotide bases and sugars and, after entry into the cytoplasm, complexation with snRNP common proteins (e.g. Sm proteins)[3−5] and with snRNA-specific proteins,[6,7] and hypermethylation of the 5′ m^7G cap to 2,2,7-trimethylguanosine (TMG). Sm protein binding, together with formation of the TMG cap, comprise a bipartite nuclear localization signal (NLS) that facilitates re-entry of the snRNPs into the nucleus.

The importance of Sm protein binding for nuclear localization of snRNA has led to a number of investigations of the RNA:protein recognition motif(s) responsible for this interaction. None of the seven individual Sm proteins (B, D_1, D_2, D_3, E, F, G) bind the uracil-rich, single-stranded Sm binding site. Thus, deciphering the mechanism by which sub-complexes of two or three of the Sm proteins form, and how these sub-complexes interact with each other, and with the RNA binding site, is an area of intense interest (Figure 1). A stable subcore Sm complex consisting of E,F,G and D_1,D_2 sub-complexes can bind the Sm

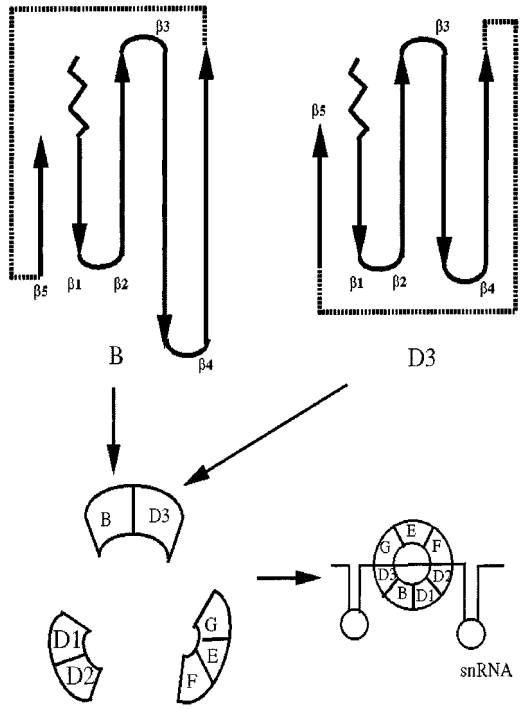

Figure 1. Depiction of the secondary structures for B and D3, two of the Sm (common) proteins. B and D3 form one protein sub-complex while D1, D2, and E,F,G also form protein sub-complexes. The three Sm protein sub-complexes form a ring that encircles the Sm binding site of snRNA. The Sm protein complex, together with the trimethyl guanosine cap, constitutes a bipartite nuclear localization signal for snRNA. In the present manuscript, molecular dynamics simulations are used to show that the D3/B configuration of the sub-complex is more stable than the alternative B/D3 configuration. These results help to establish the topology of the seven-membered ring of Sm proteins around snRNA.

* To whom correspondence should be addressed.
[†] Present address: Camitro Corporation, 4040 Campbell Ave., Menlo Park, California 94025.
[‡] Present address: Biochemistry Department, Wake Forest University School of Medicine, Winston-Salem, North Carolina 27157.

(1) Baserga, S. J.; Steitz, J. A. *The RNA World*; Gesteland, R. F., Atkins, J. F., Eds.; Cold Spring Harbor Laboratory Press: New York, 1993; pp 359−382.
(2) Andersen, J.; Zieve, G. W. *Bioessays* **1991** *13*, 57−64.
(3) Kambach, C.; Walke, S.; Young, R. A.; Avis, J. M.; de la Fortelle, E.; Raker, V. A.; Luhrmann, R.; Li, J.; Nagai, K. *Cell* **1999**, *96*, 375−387.
(4) Plessel, G.; Luhrmann, R.; Kastner, B. *J. Mol. Biol.* **1997**, *65*, 87−94.
(5) Seraphin, B. *EMBO J.* **1995**, *14*, 2089−2098.
(6) Kramer, A. *Annu. Rev. Biochem.* **1996**, *65*, 367−409.
(7) Raghunathan, P. L.; Guthrie, C. *Science* **1998**, *279*, 857−860.

10.1021/ja0004130 CCC: $19.00
Published on Web 09/27/2000

site. The B and B′ proteins arise from a single gene by alternative splicing, and either B or B′ can bind D_3 to form a stable sub-complex. The B,D_3 (or B′,D_3) sub-complex binds to the subcore to complete Sm complex formation.[8] X-ray structures have shown that D_3 may interact with B (or B′) in either of two configurations.[3,9] In one possible configuration (D_3/B), strand β4 of B pairs with strand β5 of D_3, forming an extended antiparallel β sheet. In the second configuration (B/D_3), the β5 strand of B interacts with the β4 strand of D_3. The Sm proteins share common secondary structure motifs, and determining the correct configuration of the B,D_3 sub-complex is essential for elucidating the topology of the ring formed by the seven Sm proteins about the Sm binding site (Figure 1).

MD simulations have proven useful in recent years for calculating the free energies of biological macromolecules including RNA, DNA, and proteins. Recently, Srinivasan et al. developed a method that uses the continuum solvent model to calculate the free energy associated with DNA and RNA structures.[10,11] Massova and Kollman adapted this methodology to permit calculations of the binding free energy of interacting molecules by (computationally) mutating residues at the binding interface.[12] In the present paper, we have applied this methodology to two protein complexes of D_3 and B. The binding free energy was calculated and the stability of these two protein sub-complexes was compared. The results indicate that the D_3/B configuration is more stable than the alternative B/D_3 configuration by more than 77 kcal/mol. Thus, the D_3/B sub-complex probably binds with the D_1,D_2 and the E,F,G sub-complexes to complete formation of the seven-membered-ring structure that encircles the Sm binding site of snRNA.

Computational Methods

The coordinates of the proteins used in the molecular dynamics simulations were taken from the X-ray structures[3] in which three D_3 proteins and three B proteins formed a closed hexamer with a central cavity. The A chain in the X-ray coordinate file consisted of residues Gly4 to Asp75 of D_3, while the B chain consisted of residues Ser7 to Pro87 of the Sm protein, B. The C chain also consisted of residues from D_3, and was identical to the A chain except for consisting of one less residue (Gly4). A total of five systems were considered in our studies: (1) D_3 protein (A chain); (2) B protein (B chain); (3) D_3/B protein complex (A and B chains); (4) D_3 protein (C chain); and (5) B/D_3 protein complex (B and C chains).

Molecular dynamics simulations were carried out using the AMBER5 suite of programs[13] with the Amber 1994 force field. Counterions (K^+ or Cl^-) were added for each of the five different systems analyzed to maintain each system at charge neutrality during the MD simulations. Each system was solvated with two layers of TIP3P water with an inner shell radius of 3 Å and an outer shell radius of 5 Å. The D_3 protein was solvated using 1429 and 1346 water molecules (A and C chains, respectively), the B protein using 1781 water molecules, while the B/D_3 and D_3/B sub-complexes included 2317 and 2337 water molecules. Energy minimization (1000 steps of steepest descent followed by 1000 steps conjugate gradient) was done prior to the molecular dynamics simulations to remove high-energy interatomic contacts. The initial temperature of the simulations was 100 K and the final temperature was 300 K. The Berendsen coupling algorithm was used with a heat bath coupling time constant of 0.2. All of the interactions were considered with a cutoff distance of 12.0 Å. Belly restriction was applied to fix the outer shell water to avoid solvent escape during the MD simulation. The nonbonded pair list was updated every 25 steps. Each of the five systems was run for 600 ps MD with a time step of 0.002 ps. Snapshots were collected every 10 ps during the final 400 ps for each of the MD trajectories.

The normal modes and trajectories were analyzed using expanded versions of the ANAL, CARNAL, and RDPARM modules. The final MD structures were energy minimized initially using the Sander module with explicit water, and then were minimized in a continuum solvent with a dielectric constant of $4r_{ij}$. Alternating conjugate gradient and Newton−Raphson energy minimization methods were used until the RMSD of the Cartesian elements of the gradient was less than 0.0001 kcal/(mol·Å) for all systems. The RMSD of these two energy-minimized structures was verified to be less than 2.2 Å.

The solvation energy for each system was obtained by solving the Poisson−Boltzmann equation using the MEAD program.[14] The parse charges were used.[15] The dielectric constants used were 1.0 and 80.0 for vacuum and water, respectively. A spacing grid of 0.25 Å was used in each system with the grid centered on the geometric center of the respective system. A different grid was used for each system. The nonpolar contribution was calculated based on the solvent accessible surface area (SASA) estimated by MSMS.[16]

Results and Discussion

Analysis of Trajectory Stability. The X-ray structure of the D,B_3 sub-complex of Sm proteins indicated that the sub-complex may form in either of two possible configurations. The D_3/B configuration of the sub-complex was formed by adoption of an antiparallel β-sheet between β4 of B and β5 of D_3, while the alternative configuration of the sub-complex, B/D_3, was formed upon interaction between β4 of D_3 and β5 of B. The D_3/B configuration was adopted by the A and B chains in the X-ray coordinate file while the B and C chains from the X-ray structure formed the B/D_3 configuration. The A chain consisted of residues Gly4 to Asp75 of D_3 while the C chain was identical to the A chain except for lacking one residue (Gly4). The B chain included residues Ser7 to Pro87 of protein B (residue numbering as in ref 3).

Meaningful comparisons of the free energy difference between the two configurations of the B,D_3 sub-complex can be derived from coordinate sets obtained during the course of MD simulations, provided the trajectories for each sub-complex were stable during the course of the simulation. The stability of the MD trajectories for both the B/D_3 and the D_3/B sub-complexes was verified by determining the RMSD during the course of the simulations relative to the initial X-ray coordinates. As is shown in Figure 2, the MD trajectories for both sub-complexes were stable (following an initial equilibration period) during the MD simulations. Coordinate frames were selected only from these stable regions of the trajectory (the final 400 ps) for subsequent free energy calculations. MD simulations were also calculated for the component proteins (B and D_3) in each sub-complex.

The difference in all-atom RMSD is indicative of the reorganization that the system as a whole undergoes due to solvation of the crystallized structure, as well as the inherent flexibility of the system. The single protein B (B chain) had the smallest RMSD (1.9 Å) of the five systems analyzed. The

(8) Lamond, A. I. *Nature* **1999**, *397*, 655−656.
(9) Kambach, C.; Walke, S.; Nagai, K. *Curr. Opin. Struct. Biol.* **1999,** *9*, 222−230.
(10) Srinivasan, J.; Cheatham, T. E., III; Cieplak, P.; Kollman, P. A.; Case, D. A. *J. Am. Chem. Soc.* **1998**, *120*, 9401−9409.
(11) Srinivasan, J.; Trevathan, M. W.; Beroza, P.; Case, D. A. *Theor. Chem. Acc.* **1999**,*101*, 426−434.
(12) Massova, I.; Kollman, P. A. *J. Am. Chem. Soc.* **1999**, *121*, 8133−8143.
(13) Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham, T. E., III; Ross, W. S.; Simmerling, C. L.; Darden, T. A.; Merz, K. M.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Ferguson, D. M.; Radmer, R. J.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman, P. A. AMBER5, University of California, San Francisco, 1997.
(14) Bashford, D.; Gerwert, K. *J. Mol. Biol.* **1992**, *224*, 473−486.
(15) Sitkoff, D.; Sharp, K. A.; Honig, B. *J. Phys. Chem.* **1994**, *98*, 1978−1988.
(16) Sanner, M. F.; Olson, A. J.; Spehner, J. C. *Biopolymers* **1996**, *38*, 305−320.

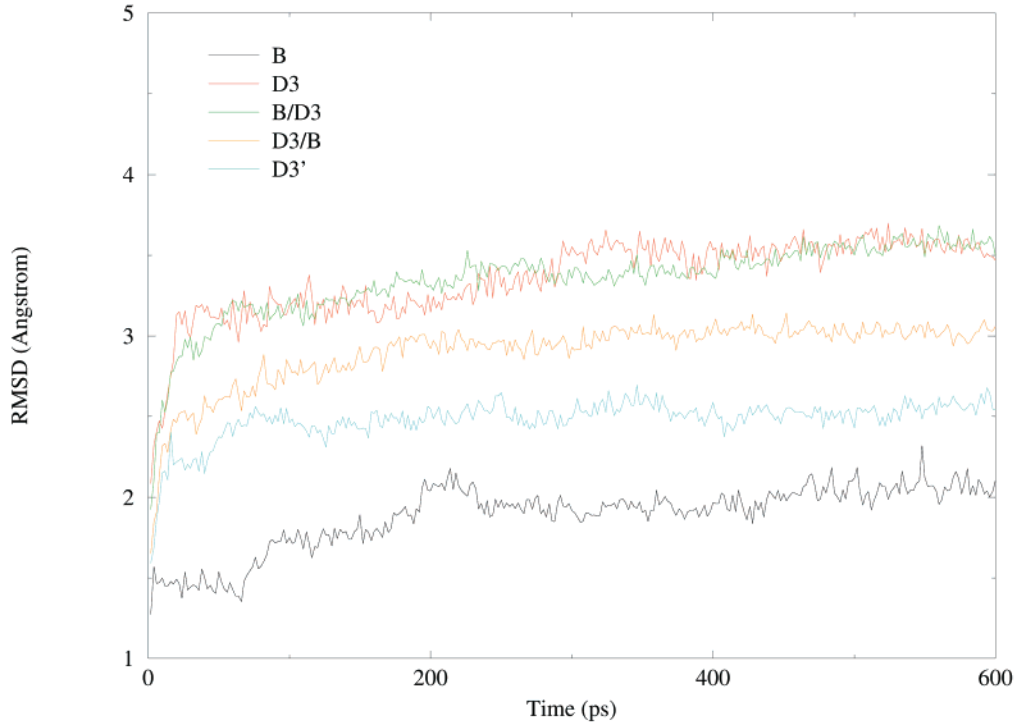

Figure 2. The RMSD of the MD trajectories for the D_3 protein (A chain from the X-ray coordinates, in red), B protein (B chain, in black), D_3 protein (C chain, in blue), B/D_3 protein complex (B, C chains in green), and D_3/B protein complex (A, B chains in yellow). The MD trajectory for each protein or sub-complex was stable following an initial equilibration period. Coordinate frames for free energy calculations were selected only from the final 400 ps of each trajectory.

Table 1. The Free Energy (kcal/mol) of the Sm Proteins B and D_3 and Their Complexes

	B (B chain)	D_3[a] (A chain)	(C chain)	D_3/B (A,B chains)	B/D_3 (B,C chains)
E_{gas}	−1638.1(39.6)[b]	−1520.0(35.2)	−1670.1(33.1)	−3724.5(77.7) [−566.4][c]	−3832.6(69.9) [−524.4]
G_{PB}	−1041.5(17.5)	−807.8(29.0)	−818.7(15.5)	−1501.4(33.2) [+347.9]	−1469.4(37.2) [+390.8]
$G_{nonpolar}$[d]	33.0(0.3)	26.5(0.4)	22.6(0.2)	49.3(0.5) [−10.2]	47.22(0.5) [−8.4]
$-TS$	−964.2	−818.7	−810.9	−1753.2 [+29.7]	−1755.0 [+20.1]
G	−3610.8	−3120.0	−3277.1	−6929.8	−7009.8
ΔG				−199.0	−121.9

[a] The A chain and C chain have a different number of residues. [b] Values in parentheses are standard deviations. [c] Values in square brackets are the differences between the complex and the separate proteins for the individual energy terms. [d] Obtained using the equation $G_{nonpolar} = \gamma \cdot SA + b$, where $\gamma = 0.00542$ kcal/A^2, $b = 0.92$ kcal/mol, and SA is the solvent accessible surface area.

two structures of D_3 have different RMSD, both of which were larger than that for the B protein (3.4 Å for the A chain and 2.5 Å for the C chain). The two configurations of the sub-complex, D_3/B (A,B chains) and B/D_3 (B,C chains), had a small difference in RMSD (3.3 Å versus 2.9 Å). The stabilities of the MD trajectories indicated that both the B/D_3 and D_3/B sub-complexes, and the component proteins (A, B, and C chains), were stable in solvent. None of the systems analyzed underwent any large structural deviations in solvent relative to the resolution of the X-ray structure (2.0 Å).

Free Energy Evaluations. The free energy for both the B/D_3 and the D_3/B configurations was obtained as the sum of the energy in vacuo (E_{gas}), the solvation energy ($G_{PB} + G_{nonpolar}$), and the entropy (TS):[10−12]

$$G = E_{gas} + G_{PB} + G_{nonpolar} - TS$$

The deformation energy, the energy resulting from protein conformational changes due to protein−protein interactions, was considered by comparison of MD simulations for the individual proteins relative to the complexes. The results are summarized in Table 1. The binding free energy of the D_3/B complex was −199.0 kcal/mol, which was 77.1 kcal/mol more stable than the B/D_3 complex. These data indicate D_3 preferentially interacts with B to form the D_3/B sub-complex in which an antiparallel β-sheet is formed between β4 of B and β5 of D_3. The D_3/B sub-complex can further assemble with the D_1,D_2 and E,F,G sub-complexes to form a ring that encircles the Sm-binding site of snRNA (Figure 1). These results are in complete agreement with observations of circular protein complexes obtained by electron microscopy of snRNA complexed by Sm proteins.[17]

In both the D_3/B and B/D_3 sub-complexes, the largest stabilizing contributions to the binding free energy resulted from the solvation energy and the direct interaction energy calculated in vacuo. It is interesting to note that this is not the case for protein−peptide interactions and duplex DNA systems studied previously.[10,12] Components of the free energy calculated in

(17) Plessel, G.; Luhrmann, R.; Kastner, B. *J. Mol. Biol.* **1997**, *265*, 87−94.

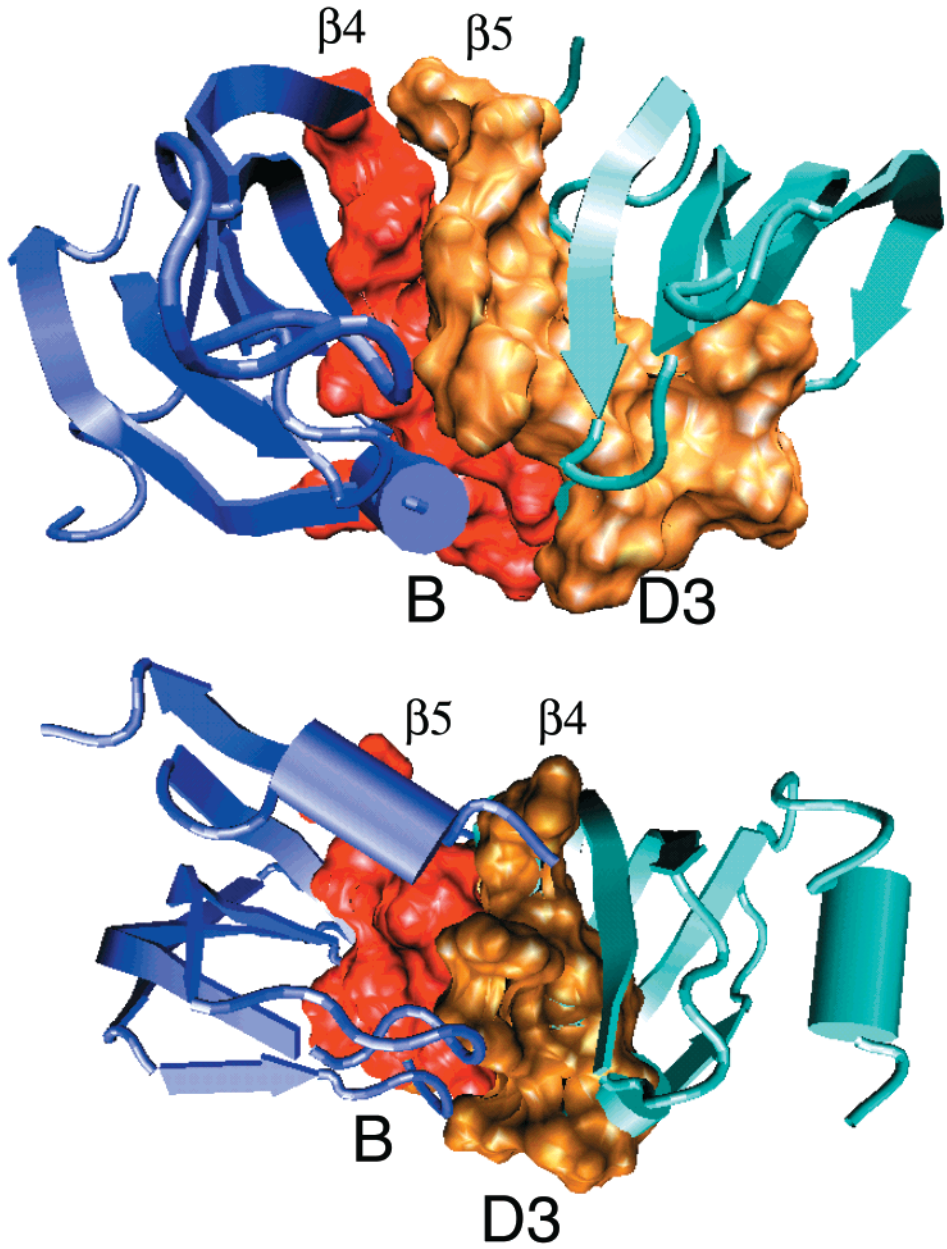

Figure 3. (A, top) Structure of D$_3$/B. The secondary structure for the D$_3$ protein is shown in cyan with residues at the interface of the complex shown with their van der Waals surface in gold. The secondary structure for the B protein is shown in blue with residues at the interface of the complex shown with their van der Waals surface in red. (B, bottom) Structure of B/D$_3$ using exactly the same color scheme for depiction as in part A. The interacting surface is much more extensive in D$_3$/B relative to B/D$_3$ consistent with D$_3$/B being the more stable configuration of the sub-complex.

vacuo included the electrostatic interaction, the van der Waals interaction, and the internal interaction energy including bond, angle, and dihedral angle interactions. The major distinction between the contributions to stability for the B,D$_3$ sub-complexes relative to DNA duplexes and protein−peptide complexes is that, in those systems, there are only short-range intermolecular contacts, i.e., the single strand−single strand contact in DNA or peptide−protein contacts. These types of interactions are fundamentally different from the protein−protein contact (bulk−bulk contact) that occurs in both the B/D$_3$ and D$_3$/B complexes. Due to the large electrostatic interactions between residues that are not in direct contact and the larger surface of interaction between two proteins, the binding free energy should, in general, be larger than that for the corresponding protein−peptide system.

The direct interaction energy calculated in vacuo contributed 42 of the 77 kcal/mol stabilization energy for the D$_3$/B sub-complex relative to the B/D$_3$ sub-complex. These results are consistent with the protein−protein interface in the D$_3$/B sub-complex consisting of substantially more favorable electrostatic interactions, including hydrogen bond formation and van der Waals interactions, relative to the interface for the B/D$_3$ sub-complex. Indeed, analysis of the structural details of the D$_3$/B interface reveals that this is the case (see below).

Differences in solvation energy between protein complexes relative to the component proteins often play a significant role in determining protein complex stability. The solvation energy for each system was obtained by solving the Poisson−Boltzmann equation using the MEAD program.[14] As is shown in Table 1, the Poisson−Boltzmann contribution to the binding free energy disfavored complex formation for both the D$_3$/B and the B/D$_3$ sub-complexes. This unfavorable contribution to the binding free energy largely offset the highly favorable contributions to the binding free energy from the direct interaction energy calculated in vacuo for both sub-complexes.

However, the Poisson−Boltzmann contribution to the binding free energy was 46 kcal/mol less disfavorable for formation of the D_3/B sub-complex relative to the B/D_3 sub-complex. This 46 kcal/mol difference for the Poisson−Boltzmann contribution was the largest of the four components (E_{gas}, G_{PB}, $G_{nonpolar}$, −TS) of the free energy of binding. These data indicate that while significant favorable electrostatic and hydrogen bond contacts between protein and solvent are lost upon complex formation for D_3/B as well as for B/D_3, greater additional protein−protein interactions occur in D_3/B than for B/D_3.

The nonpolar solvation energy is related to the solvent accessible hydrophobic surface area. Large values for this component of the binding free energy are consistent with a significant reduction in the amount of solvent accessible hydrophobic surface area playing a major role in stabilizing the complex relative to the individual protein components. The nonpolar contributions to the binding free energy of both the D_3/B and B/D_3 sub-complexes were relatively moderate, indicating that neither complex was stabilized exclusively by hydrophobic interactions. The D_3/B sub-complex was stabilized to a slightly greater, but statistically significant extent by nonpolar solvation energy than was the B/D_3 sub-complex. This difference probably reflects both the much larger interaction surface overall between B and D_3 in the D_3/B sub-complex and the greater hydrophobic character of the interaction surface (see below).

Entropic components to the binding free energy reflect restrictions in translational, rotational, and vibrational degrees of freedom that occur upon complex formation. Thus favorable electrostatic and hydrogen bonding contacts that stabilize protein−protein interactions most often come with an entropic expense. The energy components calculated in vacuo, as well as the nonpolar and Poisson−Boltzmann contributions to the solvation binding free energy, indicated that the D_3/B sub-complex was more favorable energetically than the B/D_3 sub-complex. It is not surprising, therefore, that the entropic term, −TS, disfavored D_3/B sub-complex formation relative to B/D_3. The overall favorability of complex formation thus depends on the relative balance of entropic and other terms to the binding free energy. The disfavorable entropic terms contributed only about 9 kcal/mol to formation of D_3/B relative to B/D_3 at physiologically relevant temperatures, a value too small to counteract the free energy terms that favor D_3/B sub-complex formation relative to B/D_3.

Structural Basis of Free Energy Difference. The interaction surfaces of the D_3/B and B/D_3 sub-complexes were analyzed to identify particular structural features that enabled rationalization of the substantial difference in free energy between these two configurations. The interaction surface for D_3/B is shown in Figure 3A and that for B/D_3 is shown in Figure 3B. The interaction surface of D_3/B is more than 25 Å in length across its greatest dimension, a value nearly twice that for B/D_3. The D_3/B interface consists of two parts that are approximately equal in dimensions. In one portion of the surface, the interprotein contacts are mainly hydrophobic in nature. This region includes residues Phe 70 through Leu 73 of D_3 and Gly 68 through Leu 72 of B (see ref 3 for residue numbering). Although this region is mainly stabilized by hydrophobic contacts, hydrogen bond formation between Leu 71 of B and Leu 71 of D_3 and electrostatic interactions due to favorable alignment of the amide bond dipoles in this region also stabilizes this portion of the interaction surface. In the second portion of the D_3/B interface, an extensive hydrogen bond network forms involving Lys 67 through Arg 69 of D_3 and Gly 74 and Glu 75 of B.

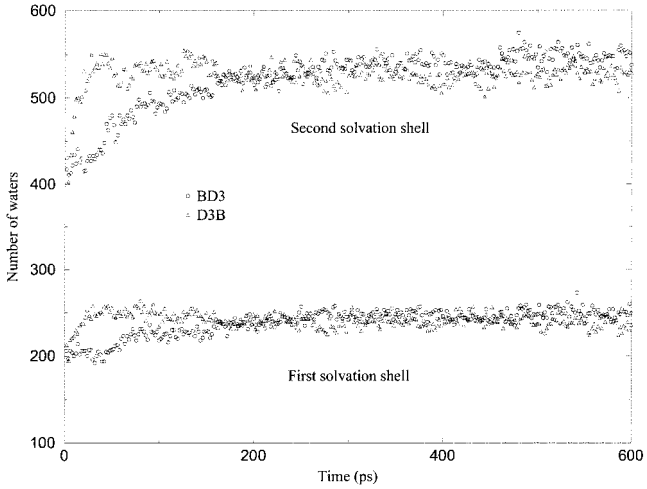

Figure 4. The number of water molecules in the first and second solvation shells of the B/D3 (circles) and D3/B (triangles) complexes as a function of time during the course of the MD simulation. Similar numbers of waters were located in each solvation shell for both sub-complexes. The number of water molecules in each solvation shell remained nearly constant during the trajectory.

Compared to D_3/B, the interaction surface in B/D_3 constituted less surface area, and was comprised almost exclusively of hydrogen bonds. Analysis of the MD trajectory for the B/D_3 complex revealed there are 87 H-bonds formed at the interface of the complex with the average residence time of 87.8 ps. The number of H-bonds observed at the interface of the D_3/B complex in contrast was fewer, with only 64 observed during the MD simulation. Between one-fourth and one-third of the observed hydrogen bonds in both sub-complexes were long-lived, i.e., had a residence time in excess of 100 ps (25 in the B/D_3 complex and 22 for D_3/B). H-bonds were detected mainly at the interface of the complexes, particularly between loop 5 and β5 of the B protein with β4 and loop 5 of the D_3 protein in the B/D_3 complex, and β5 of D_3 with β4 of B in the D_3/B complex.

Solvation of the Complexes. The interaction between macromolecules in a vacuum can be considered either as through-space, i.e., the direct Coulombic interaction without chemical bond connections, and the interaction through chemical bonds (including H-bonds). In solvated systems, the solvent-induced interactions should also be taken into account, as well as the solvation energy. Molecular dynamics simulations provide an excellent tool for insight into these interactions. The number of water molecules in the first and second solvation shells of the D_3/B and B/D_3 sub-complexes is summarized in Figure 4. The average number of first and second shell water molecules is slightly larger for B/D_3 than for D_3/B. This may contribute to the different nonpolar contributions to the solvation energy calculated for the two sub-complexes. The number of waters in both the first and second solvation shells did not change during the simulations, indicating the solvation shells were stable and the protein complexes were well-equilibrated with bulk water. The fluctuations of the water molecules contained in solvation shells during the MD trajectory for the entire system indicated that equilibrium existed between the first and second solvation shells except for a number of water molecules with specific protein interactions that remained fixed in the first solvent shell.

Interfacial water can also occur and contribute to the stability of protein−protein, protein−RNA, protein−DNA besides complexes. In general, bridged water molecules remain in position

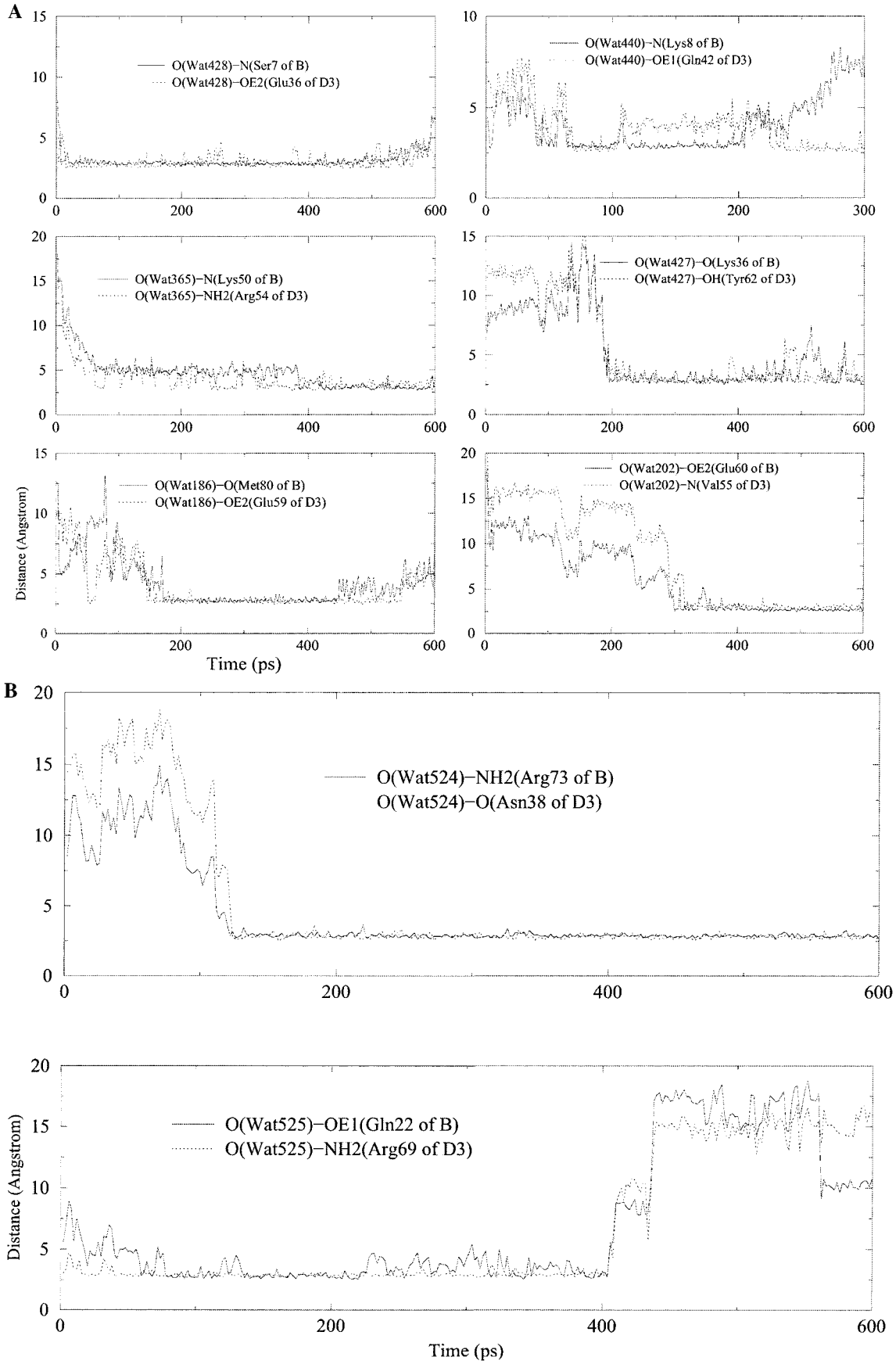

Figure 5. Bridging water molecules that stabilize the protein complexes considered in the present study. (A) The six bridging water molecules at the interface of B/D_3. The distances from the bridging water to each residue are indicated by solid and dashed lines, respectively. (B) The two bridging water molecules at the interface of D_3/B. Distances are indicated for snapshots throughout the trajectory.

longer than surface waters because of stronger interactions with solute. In our trajectories for these two complexes, 149 bridged water molecules occurred with an average H-bonding time of 31.5 ps for B/D_3, while 90 bridged water molecules were

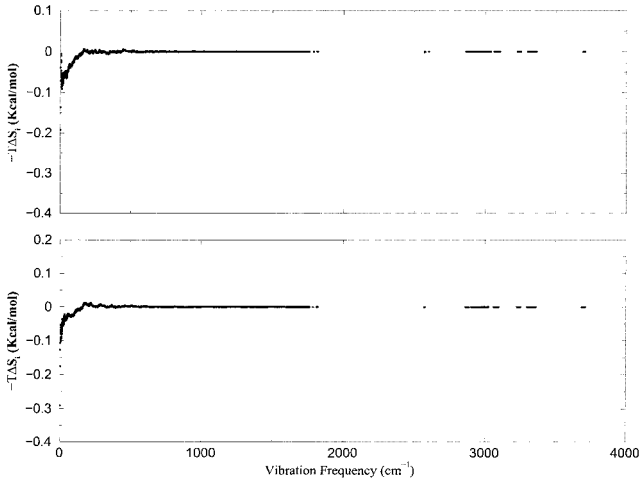

Figure 6. The binding free energy distribution from different vibration modes calculated for protein complexes D_3/B (above) and B/D_3 (below), $T = 300$ K.

observed for D_3/B with an average H-bonding time of 34.0 ps. The increased number of bridging water molecules in B/D_3 may reflect the less precise complementary fit between the two proteins in this configuration presenting opportunities for water molecules to bridge the interaction surface for portions of the trajectory. These bridging water molecules contribute to the stability of the complex, but are not in general integral components of the interfacial surface. For example, only two water molecules in D_3/B and 6 waters in B/D_3 resided longer than 200 ps on both proteins during the trajectories (Figure 5). It should be noted that neither the free energy calculations included in the present work nor those described previously by the groups of Case and Kollman included stabilization due to bridging water molecules. Inclusion of stabilization energy from bridging waters should increase the accuracy of the free energy calculations. In the present studies, the occurrence and duration of bridging waters was based on taking snapshots from the MD trajectories, and averaging the results over the trajectories. A more rigorous approach to accounting for the stabilization energy from bridging waters would involve taking averaged results from many ab intio calculations.

Vibration Modes Analysis. The components of the free energy, S_i , that resulted from vibration modes, ν_i, were calculated using the partition function

$$S_i = R\left\{\frac{(h\nu_i/kT)\exp(-h\nu_i/kT)}{1-\exp(-h\nu_i/kT)} - \ln[1-\exp(-h\nu_i/kT)]\right\}$$

The results for the binding free energy are shown in Figure 6. Due to the change in the number of degrees of freedom during complex formation, six vibration modes in each complex arise from translation or rotation in the respective component proteins. In quantum chemistry, the vibration modes analysis could be used for indicating the saddle point in the potential energy surface for a given structure. In the current instance, the structures taken from snapshots of the MD trajectories were not rigidly defined global minima because the kinetic energy and the position of each atom fluctuated no matter what kind of chemical-bond restriction was placed on it. Therefore, the image frequency may exist in real conformational space. We considered the energy-minimized structure derived from the MD simulation as a good approximation to avoid undertaking an expensive computation of the normal modes analysis for hundreds of conformations in these large protein complex systems. The same procedure has been used previously by the Case and Kollman groups.[10−12] The low-frequency vibration modes that usually correspond to torsional vibrations in proteins contributed to complex stability to a greater extent than higher frequency vibration modes in terms of the entropy. The vibration modes less than 1000 cm^{-1} contributed a total of −29.3 kcal/mol to the free energy of binding, with −24.0 kcal/mol contributed by vibrations lower than 500 cm^{-1}. It is also worth noting that any of the single vibration modes made only a small contribution to the entropic component of the free energy. The translational and rotational contributions to the total entropy of binding were typically only 2−3%, and were not considered further.

Conclusions

In summary, the free energy of interaction for the two possible configurations of the D_3,B Sm protein sub-complex was calculated using MD simulations with initial coordinates derived from the X-ray structure. The calculations indicated that the D_3/B sub-complex was 77.1kcal/mol more stable than the alternative B/D_3 complex. The solvation energy and the direct interaction energy in vacuo provided the largest contributions to the free energy of binding. Analysis of the surface of interaction for the two complexes revealed that the interface for the D_3/B sub-complex was much more extensive than that for B/D_3 and consisted of regions stabilized by both hydrophobic interactions and hydrogen bonds. The calculations also indicated that low-frequency vibration modes contributed most significantly to the entropy change that occurred upon binding. These results suggest the likely topology for formation of the seven-membered ring that occurs upon engagement of the D_3/B sub-complex together with the D_1,D_2 and E,F,G sub-complexes. These results are consistent with the available biochemical experimental data.[18−20]

Acknowledgment. The author thank Professor K. Nagai for providing the crystal structure of the B, D_3 proteins. The calculations were undertaken in the National Center of Supercomputer Application (NCSA) at the University of Illinois Urbana-Champaign (MCB990023N). This work was supported by NIH-NCI 60612 (W.H.G.) and NIH-NCI 36727.

JA0004130

(18) Raker, V. A.; Plessel, G.; Luhrmann, R. *EMBO J.* **1996**, *15*, 2256.
(19) Raker, V. A.; Hartmuth, K.; Kastner, B.; Luhrmann, R. *Mol. Cell. Biol.* **1999**, *19*, 6554−6565.
(20) Camasses, A.; Bragado-Nilsson, E.; Martin, R.; Seraphin, B.; Bordonne, R. *Mol. Cell. Biol.* **1998**, *18*, 1956−1966.