Stability and Dynamics of Domain-Swapped Bovine-Seminal Ribonuclease

by **Kalyan S. Chakrabarti**, **B. S. Sanjeev**, and **Saraswathi Vishveshwara***

Molecular Biophysics Unit, Indian Institute of Science, Bangalore-560012, India
(phone: +91-80-2293-2611; fax: + 91-80-2360-0535; e-mail: sv@mbu.iisc.ernet.in)

The proteins of the ribonuclease-A (RNase-A) family are monomeric, with the exception of bovine-seminal ribonuclease (BS-RNase). BS-RNase is formed by swapping the N-terminal helices across the two monomeric units. A molecular-dynamics (MD) study has been performed on the protein for a simulation time of 5.5 ns to understand the factors responsible for the stability of the dimer. Essential dynamics analysis and motional correlation of the protein atoms yielded the picture of a stabilising, yet flexible, interface. We have investigated the role of intermolecular H-bonding, protein/water interaction, and protein/water networks in stabilising the dimer. The networks of interchain H-bonds involving side-chain/side-chain or side-chain/main-chain (ScHB) interactions between the two chains have also been studied. The ability of protein atoms in retaining particular H_2O molecules was investigated as a function of the *accessible surface area* (*ASA*), *depth*, and *hydration* parameters, as well as their participation in protein/water networks.

1. Introduction. – Three-dimensional domain swapping between two protein molecules has been observed in a number of dimeric proteins. A comprehensive review of this topic, with a summary of the known domain-swapped proteins, has been presented recently by *Liu* and *Eisenberg* [1]. In most cases, either the N- or the C-terminal fragments are swapped [1]. The mechanism of domain swapping has been proposed for the evolution of protein oligomers [2][3]. The same mechanism has also been suggested for reversible oligomerisation, and for pathological oligomerisation, as in amyloids [4]. In bovine-seminal ribonuclease (BS-RNase), the flexibility of hinged loops appears to be responsible for domain swapping.

Biologically, oligomers have several advantages over monomers [5]. BS-RNase is a classic example of a domain-swapped protein [6]. It has been a protein of interest in view of the evolution of oligomeric proteins, and this is the only protein of the ribonuclease-A (RNase A) superfamily, expressed as a dimer in *Bos taurus* [5][6]. Although all members of the RNase-A superfamily perform the obligatory role of catalysing RNA cleavage, they differ in their substrate specificity and carry out diverse biological functions [6]. BS-RNase is highly specific to double-stranded RNA and DNA/RNA hybrids. It exhibits activities such as aspermatogenicity, embryotoxicity, immunosuppression, antitumor and antiviral functions [5–7]. The concentration of this protein is very high in seminal fluid, and its existence as a dimer seems to prevent its inactivation by RNase-A inhibitors in the intracellular fluid [4–6]. Interestingly, BS-RNase is highly homologous to RNase A, significant differences in the sequences being confined mainly in two regions [5]. The mutations in the loop region comprising the residues Gly^{16} to Ser^{20} in BS-RNase facilitate the swapping of the N-terminal helix, and

the mutations in the region Leu[28] to Cys[40] strengthens the dimer interface through disulfide bonds. These features have been established by the crystal structure of BS-RNase [8][9], as shown in *Fig. 1*, which was created with MOLSCRIPT [10].

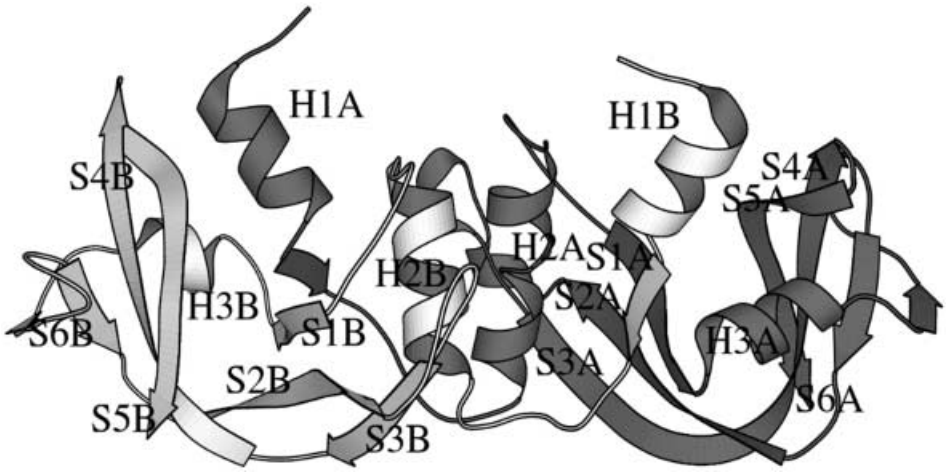

Fig. 1. *Structure of BS-RNase in ribbon representation*. The two chains *A* (dark) and *B* (light) are shaded differently, and the secondary structural elements are shown.

In the present study, we have carried out extensive molecular-dynamics (MD) simulations to investigate the factors imparting stability to the domain-swapped BS-RNase dimer. Essential dynamics analysis was employed for detailed study of each mode of atomic motion of the protein. The protein atoms showing motional correlation were identified by correlation analysis. We investigated the H-bonding, hydration patterns, and protein/water networks and examined the role of H_2O molecules in influencing the stability and dynamics of BS-RNase. A comparison of these features was also made with that of monomeric RNase A [11].

2. Results and Discussion. – 2.1. *Simulation Profile.* MD Simulation was performed on BS-RNase in aqueous medium. The root-mean-square deviation (RMSD) as a function of time is shown in *Fig. 2*. The RMSD for the C_α-atoms fluctuated by *ca*. 1.0 Å with respect to the MD-averaged structure of BS-RNase. The corresponding value with respect to the crystal structure was *ca*. 1.3 Å for the initial 1.5 ns, and then increased to *ca*. 2 Å by the end of the simulation. The residue-wise RMSD for the two chains *A* and *B* are plotted in *Fig. 3*. As expected, the RMSD of the loop regions in general are higher than those of the secondary structural regions in both chains. However, the residues Tyr[25] to Val[47], constituting a helix, a loop, and a strand in both chains, were relatively stable, with the exception of Gly[38] in chain *B* (*Fig. 3*). It is to be noted that the residues Cys[31] and Cys[32], linking the two chains by disulfide bonds, are part of this region. Further, there are some differences in the RMSD behaviour between the two chains. For example, the RMSD of the linker region Gly[16] to Ser[22] and the loop before the C-terminal strand (Gly[112] to Pro[114]) is higher for chain *A* than that of the corresponding

chain *B* regions. The residues His[105] and Ile[106] show a 0.2-Å deviation from the average MD structure in chain *B*, whereas the corresponding RMSD is *ca*. 0.5 Å in chain *A*.

2.2. *Hydrogen Bonds.* The H-bonds within and between the two chains of BS-RNase are indicated in a 2-D plot in *Fig. 4*, showing the donor/acceptor and main-chain/side-chain atoms. The H-bonds present for more than 80% of the simulation time are indicated along with those found in the crystal structure. The stable main-chain H-bonds within helices appear close to the diagonal, and those within the β-strands occur as consecutive points perpendicular to or parallel to the diagonal. Isolated points are mainly due to the interactions in nonregular structures such as those within loops or across secondary structures. It is interesting to note that there are several strong H-bonds involving side-chain atoms (ScHB), which stabilise the interchain interactions. The details of the interacting atoms are presented in *Table 1*. Strikingly, none of these interactions involve the residues mutated from RNase A, although several side-chain atoms participate in the interchain H-bonds. Mainly, there are two regions that are stabilised by inter-chain H-bonds. One of them is the region between the C-terminal end of the first helix (His[12] to Gly[16]) and the strand below that helix (Asn[44] to Glu[49]). These two regions contain the catalytically important residues His[12] and Thr[45]. The interchain interaction of Arg[33] with the main-chain atoms of the first helix assists in

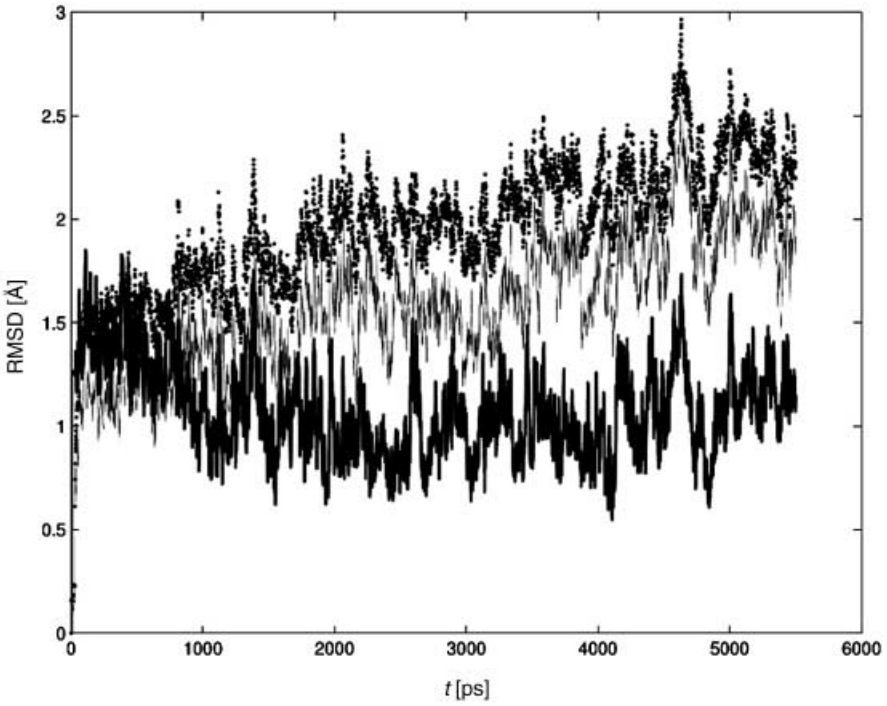

Fig. 2. *Root-mean-square deviation* (RMSD; in Å) *of BS-RNase as a function of time.* The RMSD was calculated for both all the protein atoms (doted curve; highest values) and the C_α-atoms only (bold curve; lowest values). For comparison, the RMSD data (C_α-atoms) from the crystal structure is also shown (light curve; intermediary values).

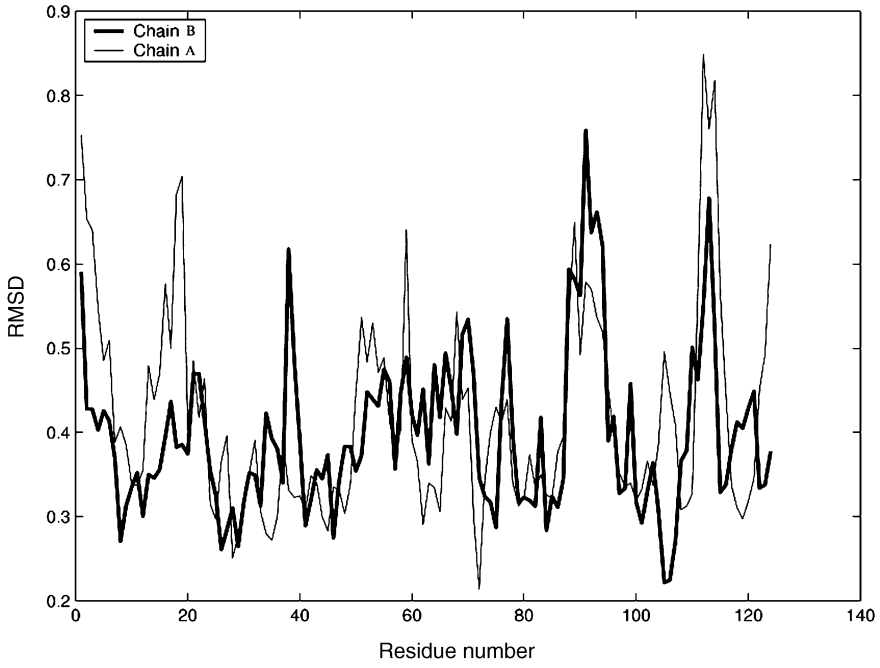

Fig. 3. *Root-mean-square deviation* (RMSD; in Å) *of BS-RNase chains* A (bold curve) *and* B (light curve) *as a function of residue number*

further stabilisation of the dimer. This includes the interaction of the interface-region residues Asp[14] of chain *A* to Val[47] of chain *B*, and of Val[47] of chain *A* to His[12] of chain *B*, and *vice versa*. All these interactions involve the main-chain atoms and were present for more than 98% of simulation time. This region below the first helix of chain *A* (Ala[4] to Met[13]) and the strand (Pro[42] to Val[47]) from chain *B* (and *vice versa*) was also stabilised by H-bonds within the backbone atoms, by H-bonds between backbone and side-chain atoms, as well as among side-chain atoms. The interchain side-chain/main-chain and side-chain/side-chain (ScHB) interactions are listed in *Table 1*.

In addition to the above-mentioned pairwise interactions, networks of stable H-bonds involving side chains (ScHB) have also been detected by a clustering algorithm. The identified interchain clusters are presented in *Table 2*. It is evident that the interactions between the first helix and the residues from the strand below this helix (clusters *2 a,b*) and Arg[33] with the first helix (clusters *1 a,b*) are further strengthened by networks of H-bonds. Such interchain interactions not only stabilise the dimeric structure, but are also significant from an 'active-site point of view'. Similar interactions involving the active-site residues (His[12], Thr[45]) were also detected during the MD calculations of monomeric RNase A and other members of this superfamily [11].

2.3. *Protein/Water Interaction.* 2.3.1. *Maximum Residence Time.* BS-RNase contains 710 polar atoms. The capacity of these polar atoms to retain a particular H_2O molecule for a maximum length of time can be evaluated by the so called *maximum residence time* (*MRT*) parameter, as plotted in *Fig. 5* for the BS-RNase chains *A* and *B*. The

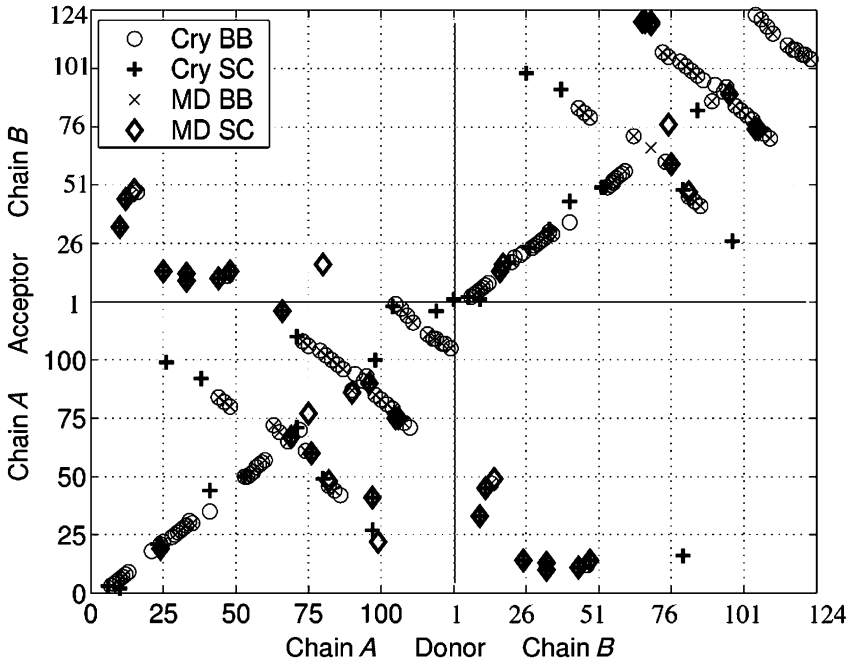

Fig. 4. *Map of calculated stable* (>80%) *hydrogen bonds of BS-RNase.* The acceptor and donor atoms are shown along the two axes. The numbers along these axes correspond to the residue numbers of chains *A* and *B*, resp. H-Bonds between backbone atoms are marked as 'BB', those between side-chain/side-chain or side-chain/backbone atoms as 'SC'. The H-bonds found in the crystal structure (Cry) are also shown.

Table 1. *Side-Chain/Main-Chain and Side-Chain/Side-Chain H-Bonds Stabilising the BS-RNase Interfaces Between the Protein Chains* A *and* B (see *Figs. 1* and *9*)

Donor-Atom	Chain	Acceptor-Atom	Chain
His^{12}-ND1	*A*	Thr^{45}-O	*B*
Ser^{15}-OG	*A*	Glu^{49}-O	*B*
Tyr^{25}-OH	*A*	Asp^{14}-OD2	*B*
Arg^{33}-NH1	*A*	Arg^{10}-O	*B*
Arg^{33}-NH2	*A*	Met^{13}-O	*B*
Asn^{44}-ND2	*A*	Gln^{11}-O	*B*
His^{48}-ND1	*A*	Asp^{14}-O	*B*
Arg^{10}-NH1	*B*	Arg^{33}-O	*A*
His^{12}-ND1	*B*	Thr^{45}-O	*A*
Ser^{15}-OG	*B*	Glu^{49}-O	*A*
Tyr^{25}-OH	*B*	Asp^{14}-OD2	*A*
Arg^{33}-NH1	*B*	Arg^{10}-O	*A*
Arg^{33}-NH1	*B*	Met^{13}-O	*A*
Arg^{33}-NH2	*B*	Met^{13}-O	*A*
Asn^{44}-ND2	*B*	Gln^{11}-O	*A*
His^{48}-ND1	*B*	Asp^{14}-O	*A*

Table 2. *Protein Side-Chain H-Bond Clusters in BS-RNase*

Cluster No.	Residues involved [a])
1a	Glu^{2} (*A*), Arg^{10} (*A*), Met^{13} (*A*), Arg^{33} (*B*)
1b	Arg^{33} (*A*), Glu^{2} (*B*), Arg^{10} (*B*), Met^{13} (*B*)
2a	Asp^{14} (*A*), Tyr^{25} (*B*), His^{48} (*B*), Thr^{82} (*B*)
2b	Tyr^{25} (*A*), His^{48} (*A*), Arg^{80} (*A*), Thr^{82} (*A*), Asp^{14} (*B*), Asn^{17} (*B*)
3a	Lys^{66} (*B*), Asn^{67} (*B*), Gln^{69} (*B*), Phe^{120} (*B*)
3b	Lys^{66} (*A*), Asp^{121} (*A*)
3c	Asn^{67} (*A*), Gln^{69} (*A*)
4	Ser^{75} (*A*), Ser^{77} (*A*), His^{105} (*A*), Ile^{106} (*A*)

[a]) Protein chains (*A* or *B*) specified in parentheses.

main-chain C=O O-atoms, in general, display higher *MRT* values than either the main-chain N-atoms or the side-chain atoms, a trend that was also observed in other members of the superfamily [11]. A comparison of *MRT* values of polar atoms in RNase A and BS-RNase (*Table 3*) shows that BS-RNase has a larger number of side chains with high *MRT* values than RNase A does. The side-chain atoms with *MRT* greater than 1000 ps

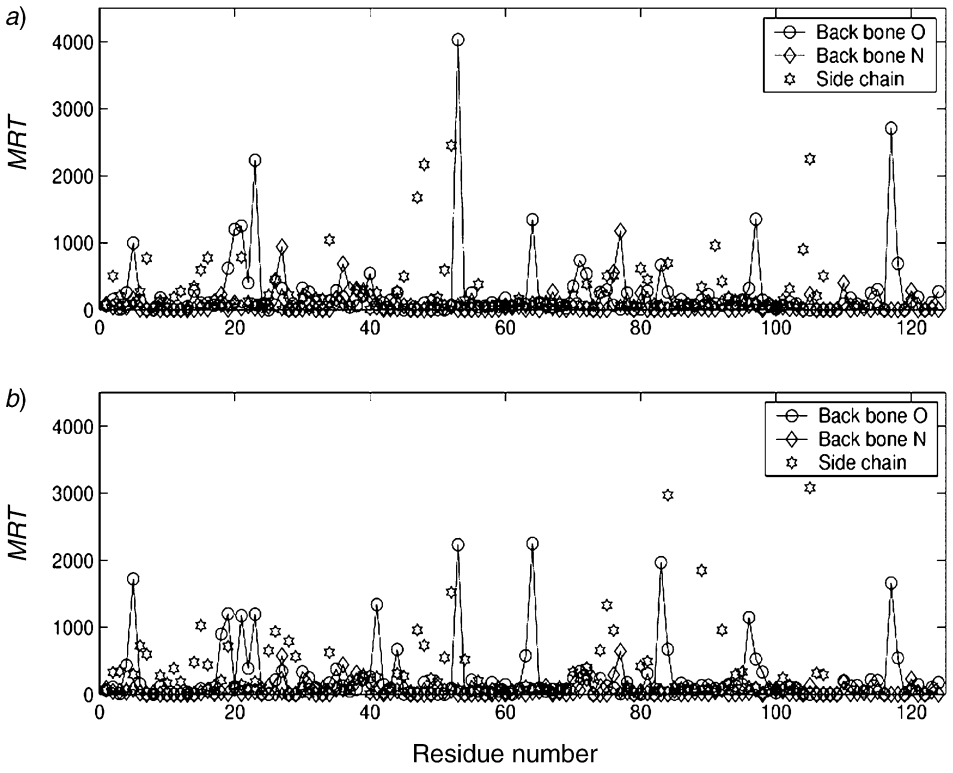

Fig. 5. *Calculated residue-wise* maximum residence time (*MRT*) *of* a) *chain* A, *and* b) *chain* B *of BS-RNase.* The polar protein atoms are backbone O-, backbone N-, and side-chain atoms.

Table 3. *Calculated Parameters for the Characterisation of BS-RNase Atoms with* Maximum Residence Times *Higher than 1000 ps.* The parameters *hydration*, *maximum residence time* (*MRT*), and *depth* are given as mean values. The term *ASA* stands for *accessible surface area.*

Protein atom	Chain	*Hydration*	*MRT* [ps]	*ASA* [Å^2]	X-*ASA* [Å^2][a])	*Depth* [Å]
Asp^{14}-OD1	*B*	2.17	1025	10.692	13.408	1.400
Thr^{36}-OG1	*A*	1.64	1042	0.824	13.852	1.623
Ala^{96}-O	*B*	0.31	1145	0.000	0.000	3.422
Ser^{77}-N	*A*	0.37	1173	3.006	0.698	1.942
Ser^{21}-O	*B*	0.53	1173	0.000	5.421	2.330
Ser^{23}-O	*B*	0.45	1193	0.000	4.77	1.56
Ser^{20}-O	*A*	1.09	1199	13.951	6.317	1.563
Pro^{19}-O	*B*	1.03	1201	3.282	5.575	2.405
Ser^{21}-O	*A*	0.73	1250	3.172	2.722	2.970
Asp^{53}-O	*A*	0.87	4026	0.330	1.450	2.137
Lys^{41}-O	*B*	0.46	1337	0.921	0.000	2.214
Thr^{64}-O	*A*	1.21	1343	10.538	14.689	1.400
Tyr^{97}-O	*A*	1.51	1348	25.661	7.962	1.400
Gln^{60}-NE2	*B*	1.11	1518	15.033	11.575	1.83
Pro^{117}-O	*B*	0.82	1660	2.060	1.067	2.204
Asp^{53}-OD1	*A*	2.74	1672	19.593	31.086	1.400
Ala^{5}-O	*B*	0.61	1719	0.146	0.548	1.699
Tyr^{97}-OH	*B*	0.79	1845	0.642	0.000	2.409
Asp^{83}-O	*B*	0.67	1966	1.085	0.112	2.094
Asp^{53}-OD2	*A*	2.73	2164	38.365	24.749	1.400
Asp^{53}-O	*A*	0.87	2230	0.000	1.517	2.051
Ser^{23}-O	*A*	1.48	2231	24.969	6.887	1.401
Thr^{64}-O	*B*	1.24	2247	9.588	2.534	1.404
Asp^{121}-OD2	*A*	1.02	2248	0.173	2.316	1.542
Gln^{60}-NE2	*A*	1.07	2448	25.025	19.701	1.813
Pro^{117}-O[b])	*A*	0.89	2707	1.826	1.329	2.022
Ser^{90}-OG[b])	*B*	0.91	2966	0.000	0.000	1.961
Asp^{121}-OD2[b])	*B*	0.96	3079	1.870	0.000	1.631
Asp^{83}-OD1[b])	*B*	1.81	1326	3.904	3.911	1.407

[a]) *ASA* Values derived from the X-ray crystal structure of BS-RNase [9]. [b]) In RNase A, the *MRT* values of these atoms were also >1000 ps. In addition, the atoms His^{12}-ND, Thr^{45}-N, Thr^{45}-O, Ile^{81}-O, and Glu^{101}-OE1 of RNase A displayed *MRT* values of 4080, 1966, 3042, 1130, and 1029 ps, resp. [11].

are part of interchain clusters, as will be discussed below in the context of protein/water networks. A number of atoms at the interface region (belonging to the residues Asn^{17} to Leu^{28}), consisting of a loop and the N-terminal part of the second helix of both chains, exhibit high *MRT* values. It should be noted that the C-terminal part of this helix from both chains are restricted by interchain disulfide bonds. Furthermore, the main-chain O-atoms of Pro^{117} of both chains displayed high *MRT* values. Interestingly, this residue interacts with the first helix of the swapped chain through a H_2O molecule. Such an interaction not only stabilises the dimer, but also is of functional significance because of its proximity to the ligand-binding region.

2.3.2. *Correlation between* Maximum Residence Time *and the Properties of Polar Atoms.* The properties of polar atoms in relation to their capacity to retain a single H_2O molecule for an extended period of time (high *MRT* values) were investigated. The *accessible surface area* (*ASA*), the *hydration* number, and the *depth* of each polar atom

were evaluated from the coordinate snapshots stored during the simulation, and these values were averaged. The distribution ranges of *ASA*, *hydration*, and *depth* are presented as histograms in *Figs. 6–8*. Separate histograms were drawn for main-chain O-, main-chain N-, side-chain O-, side-chain N-, and the total polar atoms, respectively.

The *ASA* histogram (*Fig. 6*) shows that a large number of polar atoms are completely buried, with an *ASA* value close to 0.0 Å^2. The *hydration* maxima (*Fig. 7*) and the *depth* maxima (*Fig. 8*) were at 0.5–1.7 and 1.5–2.0 Å, respectively. However, there were remarkable differences in the distribution of main-chain O-, main-chain N-, and the side-chain O-, or side-chain N-atoms as a function of these parameters. For instance, the *ASA* values of most of the main-chain N-atoms were less than 10–20 Å^2, and those of main-chain O-atoms were below 20–30 Å^2. However, the distribution of polar side-chain atoms with *ASA* values below 5–10 Å^2 was negligible, and a considerable distribution was found above 40–50 Å^2. These distributions are consistent with the property of main-chain atoms being mainly buried, and with the side-chain atoms being considerably exposed.

The *hydration* profile (*Fig. 7*) was also consistent with the exposed nature of the side chains, a significant number of them being hydrated by an average of 1.7–2.7 H_2O molecules, although a majority of atoms were hydrated with an average of 0.5–1.7 H_2O

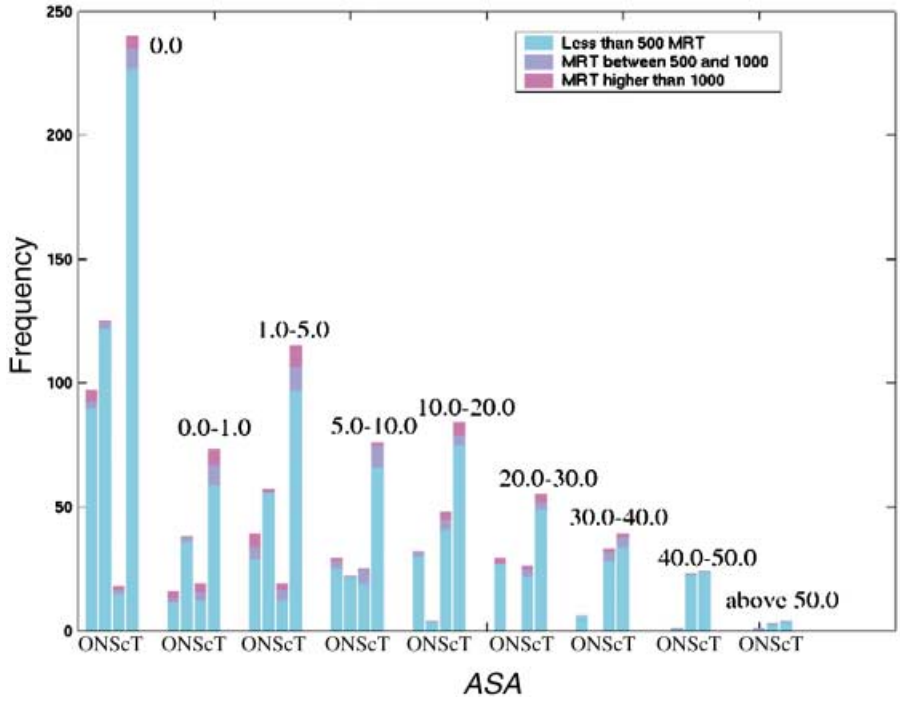

Fig. 6. *Histogram of calculated* maximum residence time (*MRT*) *as a function of averaged* accessible surface area (*ASA*; in Å^2) *of BS-RNase*. Histograms are drawn separately for main-chain O-, main-chain N-, and side-chain (Sc) atoms, as well for the total (T) structure. The averaged *ASA* values (in Å^2) are marked above the histogram. The frequencies are shown for atoms with *MRT* values below 500 (green), between 500 and 1000 (blue), and above 1000 (red).

molecules. Interesting differences, however, can be seen in the *hydration* number of both main-chain O- and N-atoms. The peak for main-chain N-atoms was in the region 0.0–0.5, just an insignificant number being above 0.5. The *hydration*-number peak for the main-chain O-atoms was in the range of 0.5–1.7.

The *depth* property (*Fig. 8*) provided additional information complementary to that of *ASA* and *hydration* properties. A majority of main-chain N-atoms were in the *depth* region of 1.5 to 2.0 Å, whereas the main-chain O-atoms spanned the entire *depth* range of 1.4 to 4.0 Å (*depth* values greater than 4 Å were also found in a few cases). The *depths* of side-chain N- and O-atoms were most prominent in the range of 1.5–2.0 Å, and small distributions were found in the 1.4- and 2.5- to 3.0-Å regions. As expected, no buried side-chain atoms were observed with a *depth* value of 4 Å or greater.

The distribution of atoms with *MRT* values between 500 and 1000 ps (blue) and those lying above 1000 ps (red) are marked on each of the bar chart of *Figs. 6–8*. The properties of the atoms with *MRT* values over 1000 ps are also presented in *Table 3*. The main-chain O-atoms with *MRT* values over 1 ns were found mainly in the low-*ASA* region (0–1 and 1–5 Å^2), and the side-chain atoms with high *MRT* values were predominantly located in the high-*ASA* region (10–20 and 20–30 Å^2). The percentage of main-chain N-atoms with a high *MRT* value was insignificant in all regions of *ASA* (*Fig. 6*). The correlation between atoms with high *MRT* values as a function of *hydration* shows that a majority of them have an average *hydration* number of 0.5–1.7

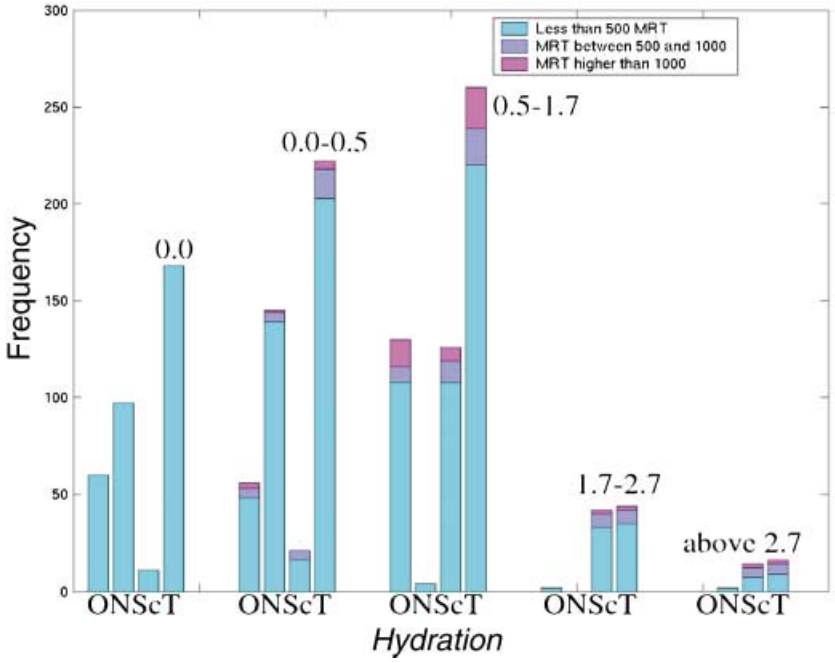

Fig. 7. *Histogram of calculated* maximum residence time (*MRT*) *as a function of the averaged* hydration *parameter for atoms of BS-RNase.* The details of the atoms and frequencies are the same as in *Fig. 6*.

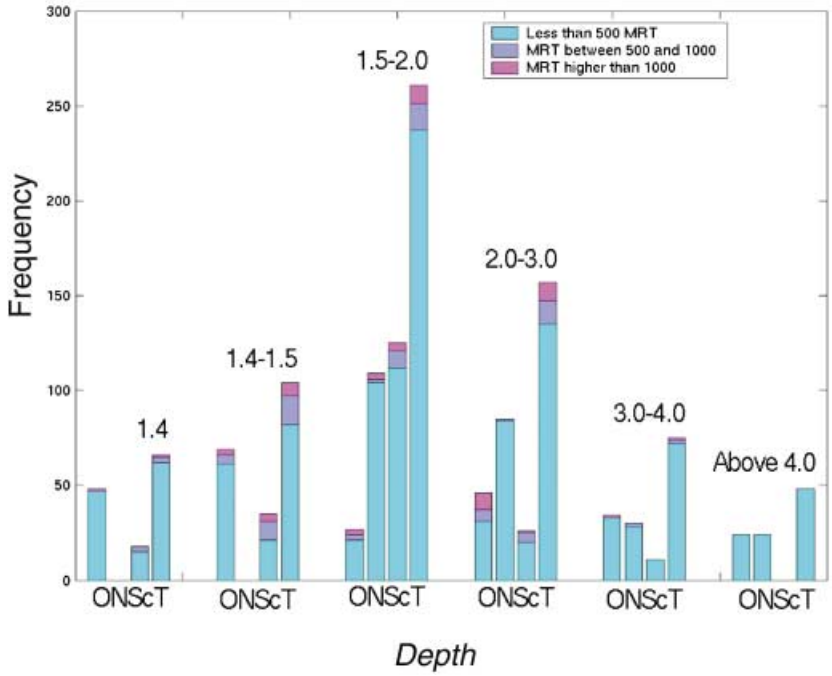

Fig. 8. *Histogram of* maximum residence time (*MRT*) *as a function of the averaged* depth *parameter* (in Å), *calculated at each snapshot from the MSMS* [12] *surface, for atoms of BS-RNase.* The details of the atoms and frequencies are the same as in *Fig. 6*.

(*Fig. 7*). A small percentage of both the main-chain O- and N-atoms posses high *MRT* values in the low hydration range (0.0–0.5). A majority of atoms with high *MRT* values are in the *depth* range of 1.4–3.0 Å (*Fig. 8*).

In summary, most of the main-chain N-atoms have a *hydration* number below 0.5, and their capacity for high *MRT* values is low. A high proportion of the main-chain O-atoms and the polar side-chain atoms display *hydration* numbers in the range of 0.5–1.7, and the atoms with high *MRT* values come from this region. Strikingly, some of the main-chain O-atoms with *ASA* values near 0.0 Å^2 also exhibit high *MRT* values. The atoms with high *MRT* values are distributed in the *ASA* range of 0–30 Å^2, and in the *depth* range of 1.4 to 3 Å. So, the local geometry of backbone O-atoms and the accessibility of the site in terms of *depth* from the *solvent accessible surface* may be the two most important factors in determining structurally important H_2O molecules.

2.3.3. *Protein/Water Networks.* Protein/water networks are often involved in catalytic reactions and are known to contribute to the stability of proteins [13][14]. Buried H_2O molecules are found in the cavities and interfaces of proteins [15]. Our studies on the family of ribonuclease-A proteins such as RNase A and angiogenin [11] have shown that the atoms with high *MRT* values are generally associated with protein/water networks. In the present study, such networks have been identified in BS-RNase by a clustering procedure. Protein atoms with *MRT* values of 500 ps or higher, which are connected through H_2O molecules, were identified as protein/water networks. The

identified networks with at least one of the atoms having an *MRT* value greater than 1000 ps (with the exception of the network *5A*, which is shown for comparison with *5B*) are listed in *Table 4*.

Table 4. Maximum Residence Time (*MRT*) *Values of Amino Acid Residues Involved in Protein/Water-Network Clusters of BS-RNase. MRT* Values expressed in ps; amino acid residues given in the one-letter code; protein chains (where required) given in parentheses.

Cluster	Chain-*A* atoms	*MRT*	Cluster	Chain-*B* atoms	*MRT*	Cluster	Chain-*A* and -*B* atoms	*MRT*
1A	N^{27}-N	937	*1B*	N^{27}-N	573	*1C*	A^{5}-O (*A*)	995
	Y^{97}-O	1348		Y^{97}-O	524		P^{117}-O (*B*)	1660
	T^{99}-OG1	958		S^{23}-O	1193		E^{9}-OE2 (*A*)	766
2A	T^{36}-OG1	1042	*2B*	T^{36}-OG1	622	*2C*	A^{5}-O (*B*)	1719
	T^{36}-N	683		N^{27}-OD1	791		P^{117}-O (*A*)	2707
				Y^{97}-OH	1845		E^{9}-OE1 (*B*)	725
							E^{9}-OE2 (*B*)	597
3A	D^{53}-OD1	1678	*3B*	D^{53}-OD1	957	*3C*	P^{19}-O (*A*)	620
	D^{53}-OD2	2164		D^{53}-OD2	728		S^{20}-O (*A*)	1199
	D^{53}-O	4026		D^{53}-O	2230		S^{21}-O (*A*)	1250
	Q^{60}-OE1	590		Q^{60}-OE1	547		D^{14}-OD1 (*B*)	1025
	Q^{60}-NE2	2448		Q^{60}-NE2	1518		N^{17}-ND2 (*B*)	714
	K^{76}-N	559		S^{77}-N	639		S^{18}-O (*B*)	896
							P^{19}-O (*B*)	1201
4A	T^{64}-O	1343	*4B*	T^{64}-O	2247		S^{21}-O (*B*)	1173
	D^{121}-OD1	2248		D^{121}-OD2	3079		N^{24}-ND2 (*B*)	934
5A	D^{83}-OD1	502	*5B*	D^{83}-OD1	1326			
	D^{83}-OD2	503		D^{83}-OD2	949			
	D^{83}-O	667		D^{83}-O	1966			
	T^{100}-OG1	422		T^{100}-OG1	959			
				N^{44}-O	667			
				C^{84}-O	671			
			6B	K^{41}-O	1337			
				S^{90}-OG	2966			

In BS-RNase, the presence of both intra- and interchain networks could be found. Interestingly, the networks of the intrachain type were similarly located in both chains. For instance, the networks *1A* and *1B* connected the residues Asn^{27} and Tyr^{97}, and the networks *3A* and *3B* connected the third helix with the nearby β-strand. Some differences in the protein/water networks between the two chains were observed. For instance, a network equivalent of *6B* was not found in chain *A*, and the composition of the residues in clusters *1–5* in the two chains were slightly different. The interchain protein/water networks (*1C*, *2C*, and *3C*) given in *Table 4* are also depicted schematically in *Fig. 9*, which was created with the VMD tool [16]. It is impressive to see a strong interchain protein/water network (*3C*) at the interface. This network involved a number of residues in the neighbourhood of the crucial mutation (residues Asn^{17}, Pro^{19}, and Ser^{20}) between the dimeric BS-RNase and the monomeric protein, RNase A. Apart from this large interface network, the dimer is also stabilised by the interaction of the N-terminal end of the swapped first helix with the loop preceding the C-terminal strand (clusters *1C* and *2C*). Specifically, the main-chain O-atoms of Ala^{5} and Pro^{117} were involved in this network. A corresponding network was also observed in monomeric

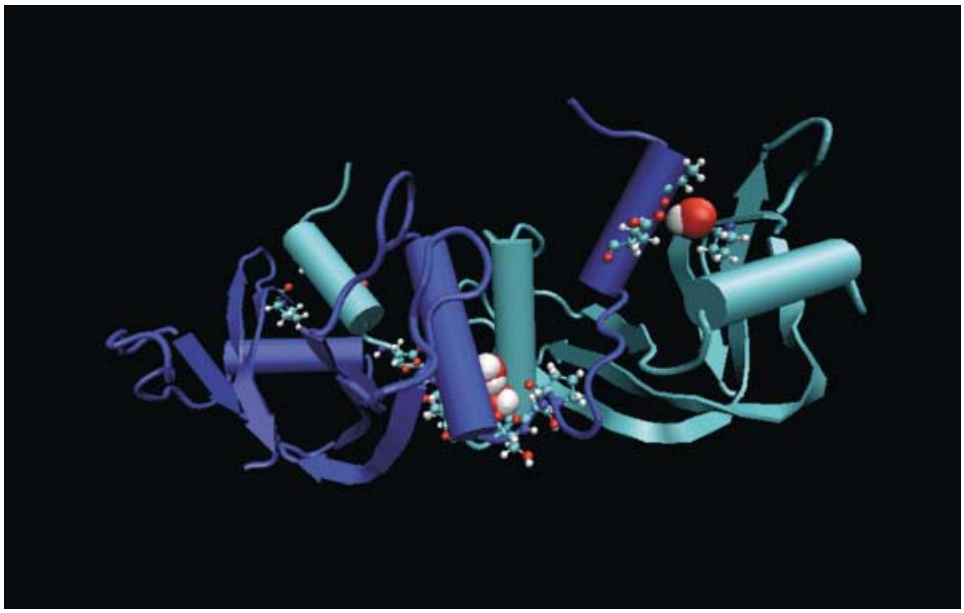

Fig. 9. *Representation of three interchain protein/water networks in BS-RNase.* For structural details, see *Table 4*. The protein is pictured in cartoon representation, the residues participating in the networks are shown in *CPK* representation (small), and the H_2O molecules are drawn in *Van der Waals* representation (large).

RNase A. It is interesting to note the conservation of such protein/water networks even among the swapped units in the dimer of BS-RNase. However, a significant difference, as compared to other members of monomeric members of the RNase-A superfamily, was the absence of protein/water-network-involving active-site His^{12} in BS-RNase. The atoms with high-*MRT* values of RNase A (*Table 3*), *e.g.*, His^{12}-ND, Thr^{45}-N, and Thr^{45}-O, are all part of a water network [11].

A comparison of the networks (*Table 4*) and the properties of the atoms with high *MRT* values (*Table 3*) shows that most of these atoms are parts of protein/water networks in BS-RNase. Furthermore, the main-chain O-atoms and the side-chain polar atoms dominate the high *MRT* list, with poor representation of the main-chain N-atoms. These trends were also observed in both RNase A and angiogenin [11]. The *hydration* and *depth* values of the atoms with high *MRT* values are greater for the side-chain atoms than for the main-chain O-atoms.

2.4. *Essential Dynamics.* In the cross-correlation map shown in *Fig. 10*, a strong intrachain secondary-structure stabilisation is evident. The points along the diagonal represent helices, and the ones perpendicular to the diagonal arise from antiparallel β-sheets. Stray points were mainly due to the interactions of side-chain/main-chain atoms. Furthermore, the dimer interface showed correlation in a few regions. Notable correlations were observed between residues Arg^{10} to Cys^{40} of both chains *A* and *B*, corresponding to the residues of the end of the first helix to those of the end of the second helix in both chains. Correlations were also found between Ser^{50} to Arg^{80} of both chains *A* and *B*, as well as between Ser^{50} to Arg^{80} of chain *A*, and Ile^{106} to Val^{124} of chain

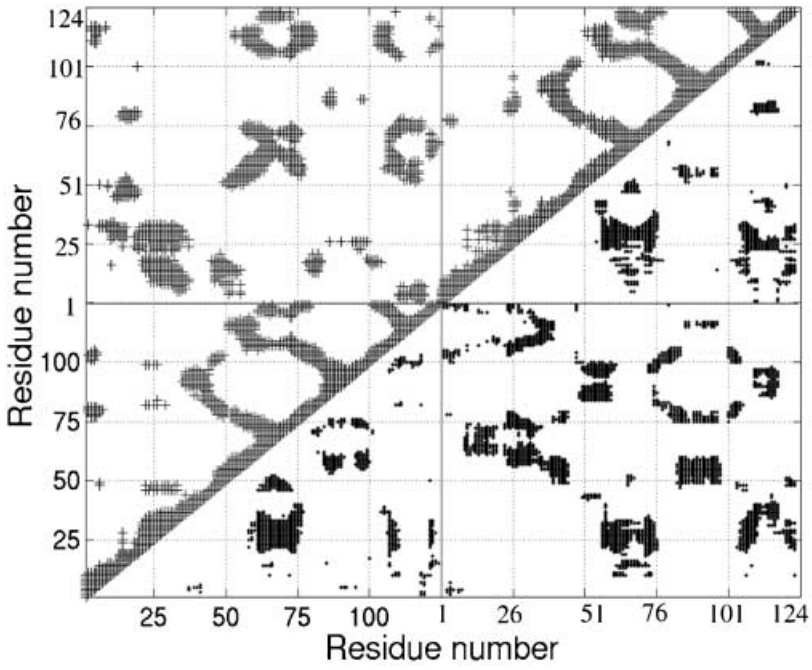

Fig. 10. *Dynamic cross-correlation map of BS-RNase.* The triangular portion above the diagonal shows correlation ($r > +0.35$), that below the diagonal shows anti-correlation ($r < -0.35$). Residue numbering of the two chains as in *Fig. 4*

B. Among these regions of correlated motion, the one involving the dimer interface, *i.e.*, in between the residues Arg^{10} to Cys^{40} of both chains, is of prime importance in holding the two chains together.

Next, we studied the modes of BS-RNase important for both regional stability and dynamics. The first two modes showed almost equal mobility uniformly throughout the protein, whereas the third showed noticeable differences in flexibility in different chains (*Fig. 11*). In chains *A* and *B*, the regions preceding and following the second helix, *i.e.*, Ser^{20} and Cys^{40}, respectively, were found to be very flexible. Since Cys^{40} is covalently linked to Cys^{95} *via* a disulfide bond, Cys^{95} also showed a high flexibility in chain *B*, but not in chain *A*. The second helix interface showed lesser mobility, which probably gives rise to the calculated stability of the dimer. The third mode captured the correlated motion of loop regions, which are on either side of the interface helices in the two chains.

2.5. *Determinants of Dimer Stability.* The energetic stabilisation of the dimeric form of BS-RNase is putatively important for its physiological action in intracellular fluids [5][6]. It is instructive to examine the determinants of the dimer stability, since the sequence of BS-RNase differs from that of the monomeric pancreatic RNase A only by *ca*. 20 amino acid residues. Nature selected the mutations [$Lys^{31} \rightarrow Cys^{31}$] and [$Ser^{32} \rightarrow Cys^{32}$], allowing for the formation of two disulfide bonds between the two chains *A* and

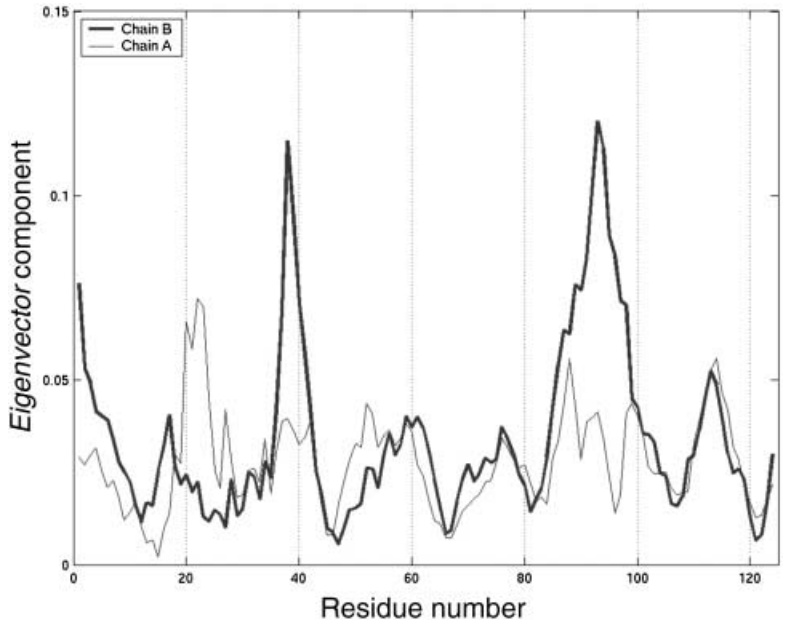

Fig. 11. *Residue-wise* Eigenvector *component corresponding to the third top mode of motion of BS-RNase chains* A *and* B (obtained from the essential dynamics of C_α-atoms)

B. In addition, the mutation [$Ala^{19} \rightarrow Pro^{19}$] seems to facilitate the swapping of the N-terminal helix domain [5].

According to the present MD simulation and analysis of water networks, it is evident that the dimer is further stabilised by interchain H-bonds (*Tables 1* and *2*) and by protein/water networks (*Table 4*). A protein/water network was found at the interface of the two chains, complementing the dimer stability conferred by the above-mentioned disulfide bridges (see *Fig. 9*). Strikingly, the residues involved in the interface network (Pro^{19} and Ser^{20}) are the ones that are mutated from RNase A. Several polar atoms with *MRT* values greater than 1 ns can be found in this region (*Table 3*). These findings bring out the importance of protein/wate -networks in dimer stabilisation. Several H-bonds and a network of H-bonds stabilise the position of the swapped first helix. Protein/water networks (*Table 4*, clusters *1C* and *2C*) are also involved in stabilising the first helix. These two networks are remarkably similar to the one observed in the simulation of RNase A [11]. However, the protein/water network connecting the active-site residues His^{12} and Thr^{45} in RNase A are, conspicuously, absent in BS-RNase, although the protein retains the direct H-bonded network between protein atoms, as discussed earlier. It would be interesting to examine, whether this difference in the protein/water-network pattern between monomeric RNase A and dimeric BS-RNase is correlated with substrate specificity.

3. Conclusions. – Bovine seminal ribonuclease (BS-RNase) is a domain-swapped dimer, whose dimerization is facilitated by interchain disulfide bonds and a flexible

hinge region. In the present study, we have investigated the dynamics of this protein. The dimer structure was found to be very stable in aqueous solution over a molecular-dynamics (MD) simulation time of 5.5 ns. Furthermore, the simulation brought out additional factors contributing to the observed dimer stability. Interchain interactions were found to be stabilised by dynamically stable (> 90% of simulation time) H-bonds, including the active-site residues His^{12} and Thr^{45}. The hydration sites (atoms) on the protein were identified by the so-called *maximum residence time* (*MRT*) parameter, atoms with high *MRT* values being characterised in terms of properties such as average *accessible surface area* (*ASA*), *hydration*, and *depth* of the polar atoms. A comparison of the hydration sites between RNase A and BS-RNase indicates that there are more hydration sites on BS-RNase. Interestingly, many of these additional hydration sites appear at the protein/protein interface in the region Asn^{17} to Leu^{28}. This emphasizes the role of H_2O molecules in the stabilisation of the domain-swapped dimer of BS-RNase.

Protein/water networks were analysed by the clustering method. The inherent capacity of the protein to engage in such networks has emerged from the observation that several intrachain protein/water networks are common to both the chains. Furthermore, networks involving H_2O molecules and interchain residues were also found. The involvement of the BS-RNase specific residues Asn^{17}, Pro^{19}, and Ser^{20} in these networks emphasizes the importance of protein/water networks in the stabilisation of the dimer.

Support from the *Computational Genomics* project sponsored by the *Department of Biotechnology*, India, is gratefully acknowledged.

Experimental Part

Molecular-Dynamics Simulation. The molecular-dynamics (MD) simulation was performed based on the high-resolution (1.9 Å) crystal structure of BS-RNase (11BG) [9]. The AMBER-7 package [17] and the parm99 parameters [18] were used. The simulation was carried out for a time of 5.5 ns in TIP3P water [19]. The solvation box was 8 Å from the farthest atom along any axis, which resulted in 11723 H_2O molecules. Particle-mesh *Ewald* summation [20] was used for long-range electrostatics. The simulation was performed under constant pressure and temp. (300 K). The *Van der Waals* cut-off was set at 10 Å, and the pressure and temp. relaxations were set at 0.5 – 1 ps. The first 500 ps of simulation was considered as the equilibration phase, with 30 ps of restrained dynamics, and the coordinate sets (snapshots) from the remaining 5 ns were used for analysis. These snapshots were stored at intervals of 1 ps. A *Silicon-Graphics Origin300* server (500 MHz, 4-processor) was used for simulation. Real-time simulation of 10 ps required *ca*. 78 min computing time.

Hydrogen-Bond Analysis. Intraprotein H-bonds between the polar N- and O-atoms of the protein were divided into two types: *1*) main-chain H-bonds (McHB), *i.e.*, H-bonds in which both donor (D) and acceptor (A) atoms are from the main chain, and *2*) side-chain H-bonds (ScHB), which included at least one side-chain atom. A standard-distance criterion of 3.5 Å or less between the donor and the acceptor, and a [D–H–A]-angle criterion of less than 60° deviation from linearity, were used to define a H-bond. A H-bond was considered as stable, when present for at least 50% of the simulation time (*i.e.*, observed in at least 2500 of the 5000 snapshots). A systematic identification of connected networks of ScHB was carried out by clustering the connected atoms. ScHBs Present in at least 90% of the coordinate sets were considered for clustering. Clusters with two or more residues were identified from the simulation.

Hydration Analysis. This analysis was focussed on the interaction of H_2O molecules with all the polar atoms of the protein. The analysis was carried out on all 5000 snapshots of the coordinates obtained from the simulation. H_2O Molecules were labelled to keep track of the number of times a given molecule interacted with a selected protein atom. The H_2O molecules located within a range of 3.0 Å from the polar atoms of the protein were identified. Likewise, this procedure was carried out in the case of H_2O/H_2O interaction. The number of snapshots in which a tagged H_2O molecule was within 3.0 Å distance of a given polar atom was considered as the

residence time of that H_2O molecule on the polar atom. All protein/water and water/water interactions with residence times greater than 500 (equivalent to or greater than 500 ps), were identified. The parameter *maximum residence time* (*MRT*) of an atom was defined as the maximum of the residence times of all H_2O molecules interacting with that particular atom. Also note that the mean *hydration* parameter of a protein atom was defined as the average number of H_2O molecules within the cut-off distance of 3.0 Å during the simulation.

A relevant H-bond between a protein atom and a H_2O molecule was characterised by a donor/acceptor (D/A) distance of less than 3 Å and a [D–H–A] angle of 120–180°. We have shown previously [11] that a spatial-temporal (S-T) criterion, namely the distance constraint of 3.0 Å obeyed by the interaction between a tagged H_2O molecule and a polar protein atom with a residence time greater than 500 ps, can accurately capture the angle distribution. The interaction of a H_2O molecule with a protein atom for an extended period of time can take place only *via* a *strong* H-bond. Thus, the protein atoms with high *MRT* values are found to interact with a H_2O molecule through one or more H-bonds.

A clustering algorithm was used to identify H_2O bridges, networks, and buried H_2O molecules interconnected through H-bonds. When a tagged H_2O molecule came in contact with two protein atoms simultaneously in more than 500 snapshots, then it was considered as a bridging H_2O molecule. If several protein atoms shared one or more H_2O molecules for at least 500 ps, the atoms were grouped as a cluster. The *accessible surface area* (*ASA*), which indicates the extent to which a protein atom is buried, was calculated by means of MSMS [12].

Depth Analysis. The *solvent accessible surface area* (*ASA*) was calculated with MSMS [12], using a 1.4-Å probe radius. The output of the program was a set of evenly distributed points, with a density of five points per $Å^2$. The *depth* parameter for protein atoms was defined as the distance of the atom from the nearest surface point.

Essential-Dynamics Analysis and Dynamic Cross-Correlation Map. The correlation between atomic motions in the protein were extracted from the coordinates obtained from MD simulations [21–25]. An established method to identify 'essential' subspace from protein dynamics has been provided by *Amadei et al.* [26]. The method, as applied to the present work, has been discussed by *Sanjeev* and *Vishveshwara* [27].

The 5000 snapshots obtained from the MD trajectory during the production run were superimposed with respect to the average MD structure, using all C_α-atoms. A covariance matrix was generated from the position coordinates of C_α-atoms.

The value of the correlation coefficient (r) between the instantaneous atomic motions (from the mean position) can vary, theoretically, from -1 (fully anticorrelated) to $+1$ (fully correlated). A cut-off of $r = 0.35$ was chosen to be the minimum magnitude of correlation considered for interpretation. The data was presented in the form of a cross-correlation map [27], as shown in *Fig. 10*.

REFERENCES

[1] Y. Liu, D. Eisenberg, *Protein Sci.* **2002**, *11*, 1285.
[2] M. J. Bennett, M. P. Schlunegger, D. Eisenberg, *Protein Sci.* **1995**, 4, 2455.
[3] J. Heringa, W. R. Taylor, *Curr. Opin. Struct. Biol.* **1997**, *7*, 416.
[4] Y. Liu, P. J. Hart, M. P. Schlunegger, D. Eisenberg, *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 3437.
[5] G. D'Alessio, *Eur. J. Biochem.* **1999**, *266*, 699.
[6] 'Ribonucleases: Structures and Functions', Eds. G. D'Allesio, J. F. Riordan, Academic Press, N.Y., 1997.
[7] R. T. Raines, *Chem. Rev.* **1998**, *98*, 1045.
[8] F. Sica, A. Di Fiore, A. Zagari, L. Mazarella, *Proteins: Struct., Funct., Genet.* **2003**, *52*, 263.
[9] L. Vitagliano, S. Adinolf, F. Sica, A. Merlino, A. Zagari, L. Mazzarella, *J. Mol. Biol.* **1999**, *293*, 569.
[10] P. J. Kraulis, *J. Appl. Crystallogr.* **1991**, *24*, 946.
[11] B. S. Sanjeev, S. Vishveshwara, *Proteins: Struct., Funct., Genet.*, in press.
[12] M. F. Sanner, 'MSMS', Ph.D. Thesis, UHA, France, 1992 (see also http://www.scripps.edu/pub/olson-web/people/sanner/html/msms_home.html).
[13] 'Topics in Molecular and Structural Biology: Water and Biological Macromolecules', Ed. E. Westoff, MacMillan, London, 1993.
[14] B. V. Prasad, K. Suguna, *Acta Crystallogr., Sect. D* **2002**, *58*, 250.
[15] V. P. Denisov, B. Halle, *Faraday Discussions* **1996**, *103*, 227.
[16] W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graphics* **1996**, *44*, 33.
[17] AMBER, Version 7, University of California, San Francisco, 2002.

[18] T. E. Chetham III, P. Cieplak, P. A. Kollman, *J. Biomol. Struct. Dyn.* **2002**, *16*, 845.
[19] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926.
[20] T. A. Darden, D. M. York, L. G. Pederson, *J. Chem. Phys.* **1993**, *98*, 10089.
[21] T. Ichiye, M. Karplus, *Proteins: Struct., Funct., Genet.* **1993**, *11*, 205.
[22] T. Horiuchi, N. Go, *Proteins* **1991**, *10*, 106.
[23] M. M. Teeter, A. D. Case, *J. Phys. Chem.* **1990**, *94*, 8091.
[24] D. Parahia, R. M. Levy, M. Karplus, *Biopolymers* **1990**, *29*, 645.
[25] A. Kitao, F. Hirata, N. Go, *Chem. Phys.* **1991**, *158*, 447.
[26] A. Amadei, A. B. Linssen, H. J. Berendsen, *Proteins: Struct., Funct., Genet.* **1993**, *17*, 412.
[27] B. S. Sanjeev, S. Vishveshwara, *Eur. Phys. J., Sect. D.* **2002**, *20*, 601.

Received January 13, 2004