Protein–Water Interactions in Ribonuclease A and Angiogenin: A Molecular Dynamics Study

B. S. Sanjeev and S. Vishveshwara*
Molecular Biophysics Unit, Indian Institute of Science, Bangalore, India

ABSTRACT **It is known that water molecules play an important role in the biological functioning of proteins. The members of the ribonuclease A (RNase A) family of proteins, which are sequentially and structurally similar, are known to carry out the obligatory function of cleaving RNA and individually perform other diverse biological functions. Our focus is on elucidating whether the sequence and structural similarity lead to common hydration patterns, what the common hydration sites are and what the differences are. Extensive molecular dynamics simulations followed by a detailed analysis of protein–water interactions have been carried out on two members of the ribonuclease A superfamily—RNase A and angiogenin. The water residence times are analyzed and their relationship with the characteristic properties of the protein polar atoms, such as their accessible surface area and mean hydration, is studied. The capacity of the polar atoms to form hydrogen bonds with water molecules and participate in protein–water networks are investigated. The locations of such networks are identified for both proteins. Proteins 2004;55:915–923.**

Key words: protein–water networks; invariant waters; maximum residence time

INTRODUCTION

Water plays an important role in the functioning of many enzymes and is crucial for their structural integrity.[1–7] A variety of experimental techniques, such as X-ray and neutron diffraction, NMR, Raman, IR spectroscopies, have been utilized[2–3,5–10] to gain insight into the nature of protein–water interactions. Although these techniques yield important information, each one has its own limitation. In order to obtain a composite picture of hydration pattern around proteins, as well as their energetics and dynamics, it is useful to combine information from these probes along with computer simulations. Computer simulations and analysis of the molecular dynamics trajectories[11–14] have become a powerful tool in studying the solvent behavior at atomic level at the protein–water and nucleic acid–water interfaces. While accurate potentials for folding are as yet unknown, a quantitative comparison of computer simulations and neutron diffraction experiments have shown that the force fields available at present are indeed capable of predicting the behavior and location of water molecules around proteins.[8]

Excellent reviews on the status of the field of protein and nucleic acid hydration are available in the literature.[12–13] Water molecules around a protein may be broadly classified[15] as a) strongly-bound internal water, b) water molecules interacting with the protein surfaces, and c) bulk water. In particular, a variety of measures such as solvent accessibility, radial distribution of water molecules, diffusion coefficient, transition probabilities, and water residence times, have been developed[11–15] which facilitate investigation of the structure, energetics, and dynamics of water in the hydration shell of proteins. The details of the protein–water interactions for distances ranging from 2.5 Å to 6 Å have been investigated.[16] The tightly bound water molecules (< 3 Å) exhibit anomalous diffusion properties[17] and it has been speculated that such water molecules may be related to mechanisms regulating the biological activity.[12] The diffusion model does not apply well to such tightly bound and buried water.[18] It has been suggested by Pettit and co-workers[19] that extensive dynamics simulation followed by an analysis at the atom–atom level is necessary for the identification of important sites for water interaction on protein molecule such as clefts and buried regions. Such interactions are dominated by protein–water hydrogen bonds and hydrogen-bonded networks and is one of the thrusts of this paper.

Detailed hydration studies based on simulations have been carried out on several proteins, which include BPTI,[20] crambin,[21] plastocynin,[16] myoglobin,[19] and β-amyloid peptide.[11] Some of these have focused on general characteristics while others have aimed at specific sites on proteins. Crystal structure analysis have been performed on proteins such as lysozyme,[6] aspartic protinases,[7] and RNase A,[5] as well as on a general set of proteins.[22] It has been suggested that the agreement between the results from the X-ray crystal structure analysis and those from simulations pertaining to the position of invariant water molecules which are tightly bound to the protein atom, are not in good agreement.[19] This has been attributed to the difference in their environment (crystal versus solvent).

Grant sponsor: the Department of Science and Technology (DST), India; Grant number: SP/50/D-108/98; Grant sponsor: the Department of Biotechnology (DBT), India, through Computational Genomics Initiative grant.

*Correspondence to: Prof. Saraswathi Vishveshwara, Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012, India. E-mail: sv@mbu.iisc.ernet.in

Received 10 July, 2003; Accepted 11 December 2003

Published online 1 April 2004 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.20114

The water residence times reported in the literature are less than 500 ps, based on simulations performed for the time scales of 1 ns or less. Indeed, to the best of our knowledge, there has been no evaluation available to indicate whether the time scale of simulation is long enough to extract the hydration properties of proteins in a reliable way. From our present simulations we find that the tightly bound water molecules which may exhibit anomalous diffusion properties, can be identified by their maximum residence times much greater than 1ns.

Our present study focuses on two proteins, RNase A and angiogenin. These proteins are sequentially and structurally similar and belong to the superfamily of RNase A. The members of the family carry out the obligatory function of cleaving RNA and also perform diverse functions. For instance, only angiogenin participates in the formation of new blood vessels (angiogenesis) and not the other members of the family. We have earlier studied this class of proteins from the points of view of dynamics,[23] ligand binding,[24] and mechanism of action.[25] This paper is devoted to the study of hydration patterns. Specific hydration sites on the protein at atomic level on RNase A and angiogenin are identified by performing long-time (5.4 ns) molecular dynamics simulations. These sites are characterized by their propensity to have hydrogen bonds with water, their solvent accessibility, the residence time of water in their vicinity and their environment (a network of hydrogen bonding connections with protein as well as with water molecules). Finally a comparison of the sites, which have potential to interact with water for an extended period (> 1 ns) is made between RNase A and angiogenin, in order to identify common sites of hydration and to explore the differences. Our studies could yield valuable information pertaining to the role of water in enzyme mechanisms.

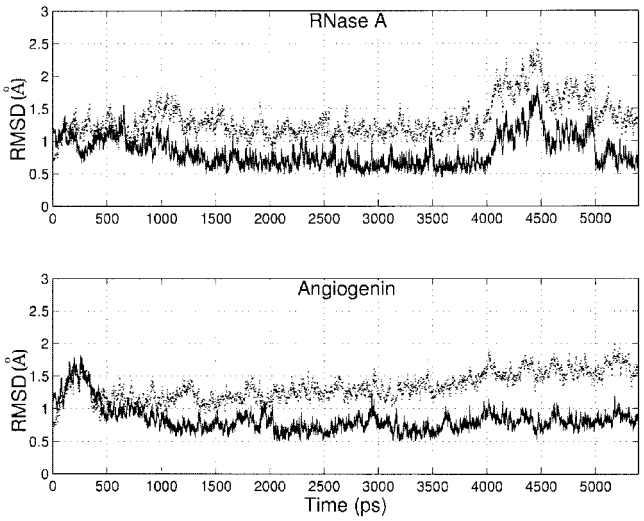

Fig. 1. C^{α} root-mean-square deviations (RMSD) of RNase A and angiogenin as a function of simulation time. The broken line represents the C^{α} RMSD with respect to the crystal structure and the continuous line with respect to the MD average structure.

METHODS

MD Simulations

Molecular Dynamics (MD) simulations were performed on the high-resolution (better than 1.8 Å) crystal structures of bovine pancreatic ribonuclease A (RNase A, PDB code - 7RSA)[26] and human Angiogenin (PDB code - 2ANG).[27] AMBER7[28] packages with parm98 parameters[29] were used. Explicit water molecules were used in the simulations, which were carried out for 5.4 ns in TIP3P water.[30] The solvation box was 8 Å along all the axes. This resulted in 5400 and 5715 water molecules in RNase A and Angiogenin systems respectively. Particle Mesh Ewald summation[31] was used for long range electrostatics and Van der Waals cut-off was 10 Å. Pressure and temperature relaxation was 0.5 ps^{-1}. The simulation was done under constant temperature and pressure conditions. The first 400 ps of simulation was the equilibration phase and the rest was used for analysis. The root-mean-square deviations (RMSD) of the simulations are shown in Figure 1. The C^{α} RMSD fluctuated around 0.7Å in all the systems during most of the simulation time. A 6-processor 1GHz Intel-based Linux cluster, a 4-processor Sun 420R, an IBM-SP3 and a 4-processor Origin300 (500MHz) were used for simulations. Depending on the machine and the system, a 10 ps simulation took typically between 37–80 minutes of CPU time.

Hydration Analysis

Maximum residence time

The first 400 ps of the simulation (starting from the native-state crystal structure) is considered as the equilibration phase and the next 5000 snapshots, which are stored at 1 ps interval from the MD trajectory are analyzed. The water molecules, which are within 3 Å from the protein polar atoms (oxygen and nitrogen) in each snapshot are identified. The water molecules are labeled so that one can readily extract the information on the number of times a given water molecule interacts with a selected protein atom. Water–water interactions are also kept track of and are used for benchmarking. The number of snapshots that a given water molecule is within a given polar atom is considered as the residence time of that water molecule on the polar atom (varying definitions for residence time are used in the literature[20,32]). We have identified all protein–water and water–water interactions with residence time greater than 500 (equivalent to or greater than 500 ps). The term maximum residence time (MRT) of an atom is defined as the maximum of number of times any particular water molecule has interacted with it. The mean hydration of an atom is defined as the average number of water molecules within the cut-off distance during the simulation.

Hydrogen bond analysis

Hydrogen bonds may be identified by means of a distance and angle constraints. Detailed statistical analyses have shown that a simple scheme to identify hydrogen bonds is to require that distance between the oxygen and nitrogen atoms of a protein with that of the oxygen of the

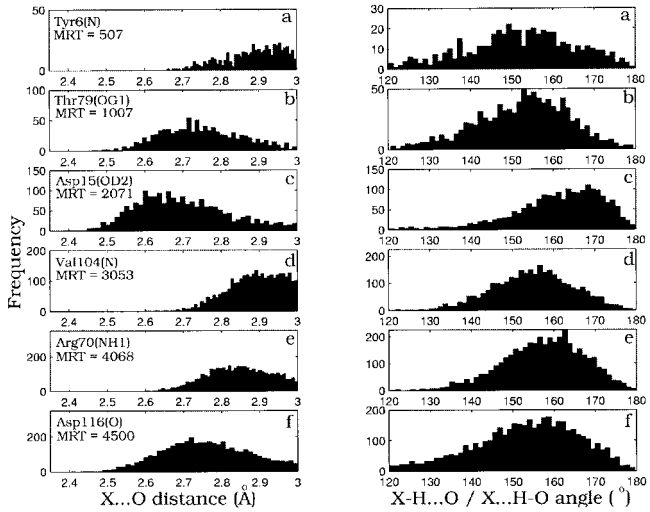

Fig. 2. Distribution of geometrical constraints imposed by the spatial–temporal (S–T) criterion for angiogenin. This criterion selects the water molecules, whose oxygen atoms(O) are within 3 Å from the protein polar atoms (X) with MRT ≥ 500. The left panels (a–f) depict the distribution of the distance (X. . .O) whereas the right panels (a–f) show the distribution of the angle [donor(X/O)–proton-acceptor(X/O)] involving the corresponding atoms. The protein atom type and MRT are indicated in the left panels. The distribution of angles show that the S–T criterion nicely captures the geometrical features of hydrogen bond. The distance distribution is consistent with earlier studies[22,33–34] which have shown that the hydrogen bond between a water molecule and a nitrogen atom is longer by about 0.07 Å compared with that of an oxygen atom.

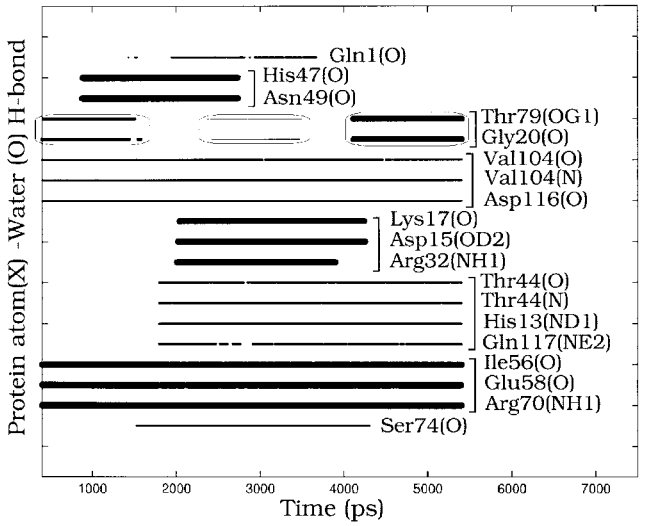

Fig. 3. Sketch of the temporal history of selected tagged water molecules with MRT ≥ 1000 in angiogenin. The horizontal lines indicate the duration of the interaction of the specific water molecule with a particular polar atom of the protein. Successive lines with the same thickness illustrate the interaction of the same water molecule with several protein atoms, leading to the formation of protein–water network (see Table I). Strikingly, the onset of several interactions occurs after 1.5 ns. This underscores the importance of long-time simulation for the purpose of identification of protein–water networks.

water molecule is around 2.88 Å and 2.95 Å respectively.[22,33–34] Furthermore, a good hydrogen bond should also satisfy the (donor–hydrogen-acceptor) angle to be greater than 120°. We have carried out extensive simulations to show that all these criteria can be encapsulated rather simply by just an effective spatial–temporal (S–T) criterion: We define a hydrogen bond to have formed when the distance between protein polar atom and the oxygen of the water molecule is less than 3 Å and the MRT of at least 500. The rationale is that unless both the distance and the angle constraints are well-satisfied, a water molecule does not reside for long intervals in the vicinity of a protein atom. More than 97% of the cases selected for interaction based on the S-T criterion automatically satisfy the angle cut-off criterion (also see Figure 2). As stated above, this result confirms the fact that the interaction of water molecule with a protein atom for an extended period can take place only through a good hydrogen bond. The bond length distribution is in accord with the average distances for a given type of atom, i.e., a peak for the oxygen atom is at a lower distance (2.6–2.8 Å) than that for the nitrogen atom (2.8–2.9Å).[22,33–34] The use of S–T criterion facilitates an efficient processing of a huge amount of data produced from long-time simulations.

Identification of water bridges, networks, and buried water

Clustering algorithm[35] was used to identify the connected atoms. When a given water molecule makes contact with two protein atoms simultaneously in more than 500 snapshots, then it is considered as a bridging water molecule (see Fig. 3). If several protein atoms with MRT of 1000 or more are interconnected through one or more tagged water molecules, the cluster of such atoms is called a network. If a given water molecule is in contact for more than 500 ps with one or more buried atoms (measured by its accessible surface area[36]), then it is denoted as a buried water molecule. Furthermore, the water–water interactions with MRT greater than 500 are also identified. All these water molecules interact strongly with protein atoms.

RESULTS AND DISCUSSION

Statistics of Protein–Water Interaction

There are 364 and 376 polar atoms in RNase A and angiogenin respectively. The mean hydration histogram presented in Figure 4 shows that in both cases, about 200 atoms have low hydration (< 0.5), about half of which do not come in contact with water. A fraction of polar atoms (< 30) are hydrated with a mean around 2–3 (high hydration) and about 150 polar atoms have a mean hydration between 0.5 to 1.5 (medium hydration). The mean hydration of all polar atoms put together is 0.64 for RNase A and 0.66 for angiogenin.

The capacity of a polar atom to retain a water molecule for a maximum length of time is given by the parameter maximum residence time (MRT), which is described in the Methods section. A majority of polar atoms in both the proteins have MRT less than 250 [Fig. 5(a)] and only a small fraction of them have MRT in excess of 1000. Since many atoms have MRT up to 250, it requires a long-time simulation to distinguish a special site with high MRT

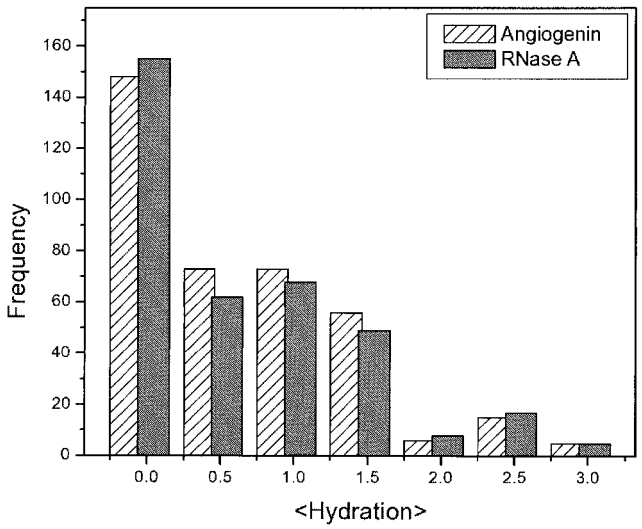

Fig. 4. The frequency distribution of the mean hydration of protein polar atoms of RNase A and angiogenin.

from the pool of polar atoms in the protein. Our focus has been on such special sites and hence our present 5.4 ns simulation is significant from this point of view. The plot of MRT as a function of sequence number is presented in Figure 5(b). The MRT of main-chain(mc)-oxygen, mc-nitrogen and the side chain atoms of a given residue are separately indicated in the plot. It is evident that the MRT of only a small fraction of atoms is greater than 500 ($>$ 10% of the simulation time or 0.5 ns) and only a few atoms are almost constantly bound to the same water molecule with MRT greater than 3000. Higher MRT values are generally seen for mc-oxygen than that of mc-nitrogen [Fig. 5(b)]. Our results are consistent with earlier crystal structure analysis.[22,33–34] Furthermore, our simulations (in progress) on another protein (eosinophil cationic protein) with slightly different starting structures have shown that our method of detecting protein sites with high MRT is indeed robust.

The location of the polar atom in the protein structure and its exposure to solvent is known to play a major role in its capacity to interact with water molecules. Several attempts have been made to predict the sites with high MRT.[13,15–16,20–22,37–38] We have carried out a systematic investigation of MRT as functions of the mean accessible surface area (<ASA>), that is the mean of accessible surface area (ASA) over the trajectory, as well as mean hydration of individual atoms obtained from the 5 ns trajectories. The results are plotted in Figures 6–8. As expected, the mean hydration is low (<0.5) for atoms with low mean ASA (<15Å^2) (see Fig. 6). However, for atoms with mean ASA in excess of 20Å^2, there is no correlation between mean ASA and mean hydration. The atoms with high mean hydration have low MRT (Fig. 7), which indicates that highly hydrated atoms frequently exchange water molecules with the bulk water. It is interesting to note that the atoms with low hydration (0.3 to 1.3) are the ones with high MRT. The relation between mean ASA and MRT is particularly striking (Fig. 8) in the region with

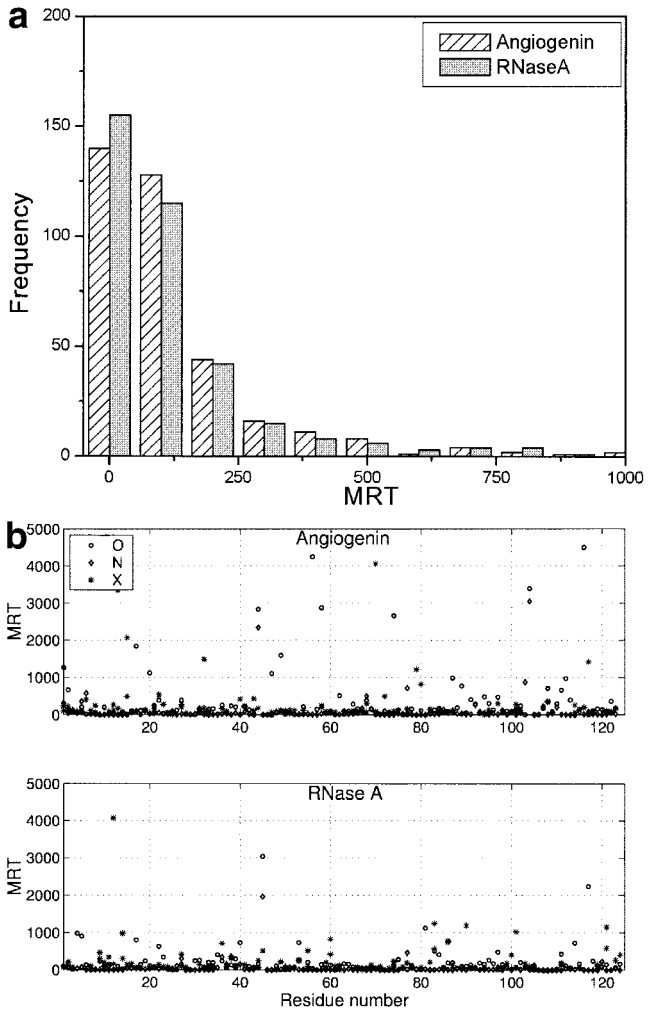

Fig. 5. **a**: The frequency distribution of the MRT of polar atoms in RNase A and angiogenin. There are 9 and 19 sites with MRT $\geq$ 1000 respectively in RNase A and angiogenin, which have been excluded from the plot. **b**: Plots of MRT as a function of residue number for the two proteins. The MRT values are separately plotted for the mainchain oxygen (**o**) , mainchain nitrogen (♦) and the side-chain polar atoms (*).

mean ASA less than 20Å^2. All those atoms with high MRT have low mean ASA and many of them have a mean ASA of close to 0Å^2. Furthermore, it should be noted that none of the atoms with mean ASA greater than 30Å^2 have MRT more than 500. Thus, highly exposed polar atoms are incapable of retaining a given water molecule for a long time and this suggests that invariant water molecules ought to be associated with slightly exposed protein polar atoms. Indeed, invariant water molecules could often be buried or internal water molecules. Our finding is consistent with earlier results on myoglobin,[19] where the sites with high residence time were identified to be present in cavities and clefts of the protein. The MRT is generally below 40 for water–water interactions, with the exception of a few water molecules which exhibit high MRT (>500). Such water molecules are invariably localized in the vicinity of protein atoms and are stabilized by their interactions.

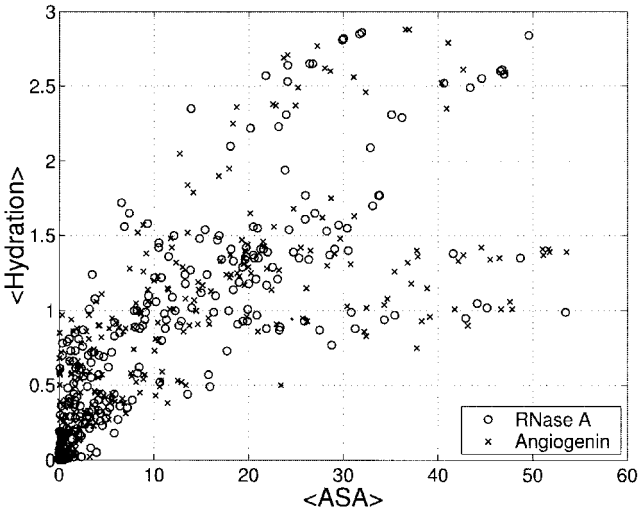

Fig. 6. A cross plot of <hydration> versus atomic <ASA> of polar atoms of RNase A and angiogenin.

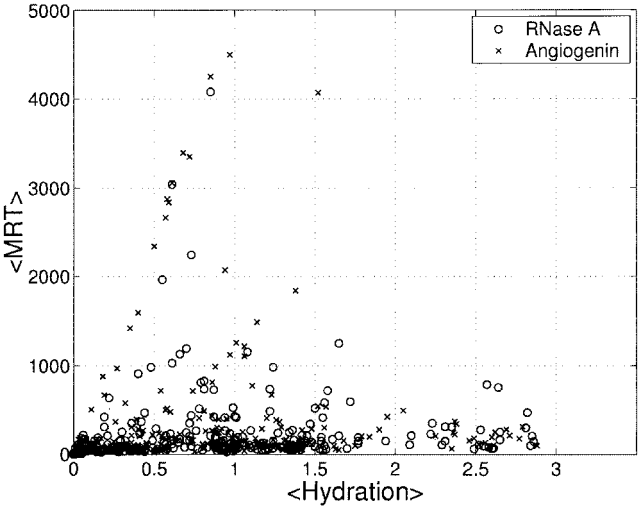

Fig. 7. A cross plot of <hydration> versus MRT of polar atoms of RNase A and angiogenin.

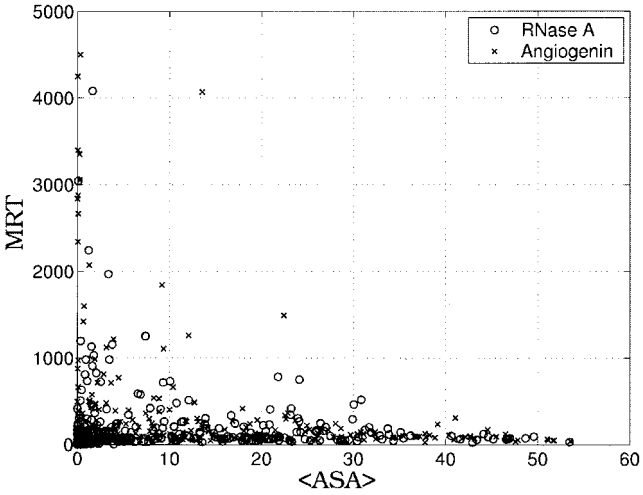

Fig. 8. A cross plot of atomic <ASA> versus MRT of polar atoms of RNase A and angiogenin.

Protein–Water Networks

The inter-connectedness of two or more protein atoms linked through the same water molecule(s) provides interesting new information. Such networks are well known to be contributors to the stability of biomolecules and are also implicated in catalytic reactions.[1–7,39] The networks are identified by a clustering algorithm.[35] Protein atoms with MRT of at least 500 which are interconnected through tagged water molecule(s) are identified as protein–water network clusters. The identified clusters with at least one protein with MRT over 1000 for RNase A and angiogenin are presented in Table I. The trajectories of the protein–water interactions for water molecules with MRT higher than 1000 (in simulation time it is invariably equal to or greater than 1ns because of a stringent cut-off criterion used for the detection of interaction) for angiogenin are shown in Figure 3. Two points can be noted from this figure and Table I: (1) The protein atoms in a given cluster with MRT over 1000 are generally connected through the same tagged water molecule in protein–water networks, and 2) in several cases, the formation of a complete network with the participation of a water molecule takes place after 1ns time of simulation, underscoring the need for long-time simulation to identify well equilibrated invariant water positions.

Ribonuclease A

Some of the significant networks given in Table I are depicted in Figure 9. His12(ND1), Lys41(NZ), and Thr45(N & O) are among the active-site residues of RNaseA superfamily and it is interesting to note that all of the three residues are involved in a connected interaction with a water molecule. The specific water molecule bound to His12 appeared within 100 ps of the equilibration period and retained its interaction throughout the simulation period These residues are further connected to Asp83(OD1) by an interaction through another water molecule. A network of special interest in RNase-A contains Ala4(O), Ala5(O), Gln55(NE2), and Pro117(O). The first helix takes two slightly different orientations during the simulation; one as was seen in the crystal structure [Figure 10(a)], and the other tilting towards the third helix [Figure 10(b)]. In the former orientation, Pro117(O) is bridged to Ala5(O) and in the latter, to Ala4(O) through a water molecule. The water molecule that connects to Ala5(O) is also bridged to Gln55(NE2) through another water molecule. This network, involving hydrogen bonding to water molecules, assumes importance as the entry point for the ligand into the active site region.

It is often believed that the exposed side chains and loops are unstructured. In this study, we can see examples of stabilized side-chain conformations and a loop adopting a rigid conformation through interaction with water molecules. An example of the stability of side-chain conformation is the interaction of water molecules between Cys40(O), Glu86(OE1 & OE2), Ser90(OG) (see Fig. 9), almost through-

TABLE I. The Protein-Water Network Clusters in RNase A and Angiogenin

System	Cl number[a]	Residues and atoms in the network[b]
RNase A	1	Ala4 (O) [982], Ala5 (O) [908], Gln55 (NE2) [523], Pro117 (O) [2242]
	2	His12 (ND1) [4080], Thr45 (N) [1966], Thr45 (O) [3042]
	3	Asp14 (OD2) [983], Thr17 (O) [810]
	4	Cys40 (O) [734], Glu86 (OE1) [755], Glu86 (OE2) [786], Ser90 (OG) [1193]
	5	Thr45 (OG1) [517], Asp83 (OD2) [1250]
	6	Ile81 (O) [1130], Gln101 (OE1) [1029]
Angiogenin	1	Gln1 (O) [1260], Asp2 (O) [667], Tyr6 (N) [577], Glu108 (O) [713], Leu111 (O) [664], Pro112 (O) [971]
	2	His47 (O) [1106], Asn49 (O) [1593]
	3[c]	Gly20 (O) [1121], Thr79 (OG1) [1216]
	4	Val104 (N) [3053], Val104 (O) [3394], Asp116 (O) [4500]
	5	Asp15 (OD2) [2071], Lys17 (O) [1841], Arg32 (NH1) [1489]
	6	His13 (ND1) [3349], Thr44 (N) [2342], Thr44 (O) [2837], Gln117 (NE2) [1418]
	7	Ile56 (O) [4250], Glu58 (O) [2875], Gly62 (O) [511], Arg70 (NH1) [4068]
	8	Ser74 (O) [2665], Val103 (N) [877]
	9	Ser87 (O) [989], Trp89 (O) [773]

[a]Cluster number.
[b]The number in square brackets is the MRT of the atom. In any given cluster, all the atoms with MRT in excess of 1000 are found to be connected by the same tagged water molecule.
[c]This cluster is seen with four different water molecules that bridge the two atoms at different times during the simulation.

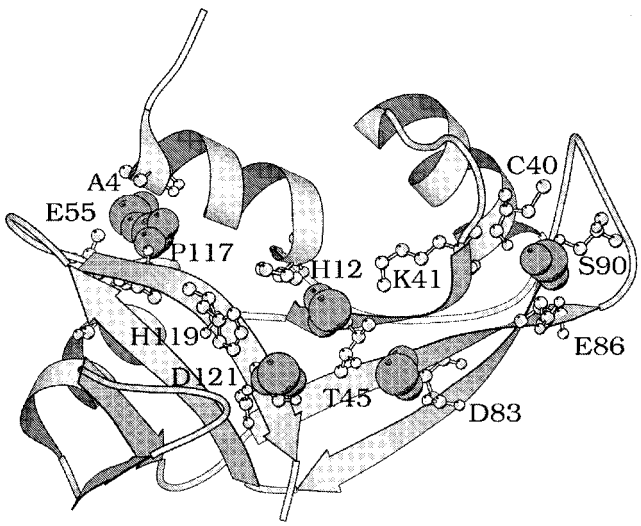

Fig. 9. Active sites and a few representative protein–water networks in RNase A. Protein residues are shown in ball and stick representation while the water molecules are shown in Van der Waals surface representation. All cartoon representations are made using MOLSCRIPT.[40]

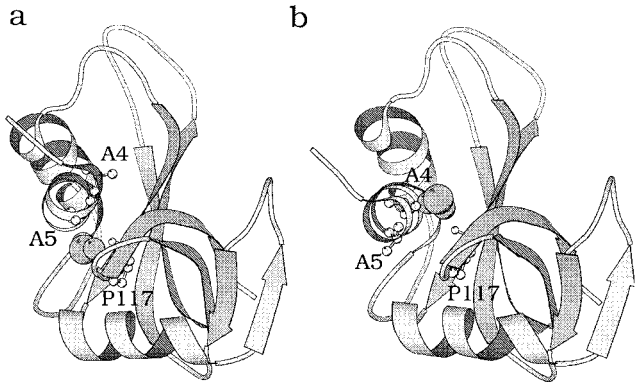

Fig. 10. Two different orientations of the N-terminus helix with the loop preceding the C-terminal strand in RNase A. A specific water molecule stabilizes the interaction of Pro117(O) with Ala5(O) in (**a**) and with Ala4(O) in (**b**).

out the simulation The side-chains of Ser90 and Glu86 are exposed and, consequently, the bound water molecules frequently exchange with the bulk water with MRT between 200–700. The stabilization of a loop conformation is witnessed by the water-mediated interaction of Asp53(O), belonging to the third helix preceding a loop and Ser77(N) which is a residue in a strand following the loop. Although this network is not as strong (MRT between 500–800) as some others, it represents a case where the two secondary structures are held together through water molecule.

A comparison with experimental results shows that a large number of residues (Ala4, Ala5, Cys40, Thr45, Asp53, Gln55, Asp83), which we have identified as important for interaction with water molecules have also been reported as being associated with invariant waters molecules.[5,23] Interestingly, the catalytic site residue His12, which we have detected as part of a network, does not appear as an invariant hydration site in X-ray structure-based analysis. It is not clear whether this is due to the effect of crystal packing or because the invariant positions for water molecules were derived from crystal structures comprising native as well as ligand-bound forms. Our present simulation is just on the native protein and the simulations of ligand-bound structures are in progress.

Angiogenin

Protein–water networks in angiogenin are presented in Table I and some of them are shown in Figure 11(a–d). Angiogenin, like RNase A, has a network with active site residues His13(ND1), Thr44(N & O) [Fig. 11(c)], connected through a tagged water molecule. Gln117(NE2) is also a part of this network. Lys40(NZ), which was a part of the corresponding network in RNase A, is connected through another water molecule with MRT of 421. The protein–

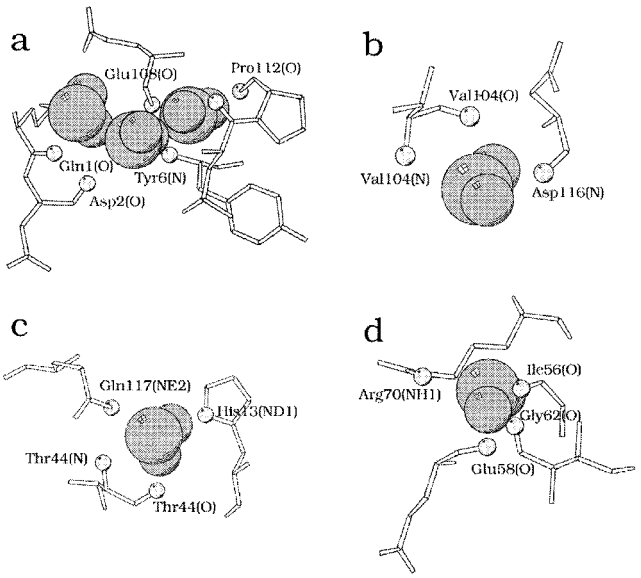

Fig. 11. Representative protein–water networks in angiogenin. The figures **a**, **b**, **c**, and **d** correspond to clusters 1, 4, 6, and 7 in Table I, respectively. Active site cluster, (c), is common with RNase A. The protein residues are shown in ball and stick representation while the water molecules are shown in Van der Waals surface representation. The polar atoms with MRT $\geq$ 1000 are represented by open spheres.

water network bridging the N-terminus helix and the loop before the C-terminal strand, detected in RNase A is also seen in angiogenin [Fig.. 11(a)]. Two to three water molecules are part of this network, which involves the residues Gln1(O), Asp2(O), Tyr6(N), Glu108(O), Leu111(O), and Pro112(O). Strikingly, the side-chain of Gln1 underwent a conformational change during the simulation and joined this network. A water bridge connecting Val104(O & N) and Asp116(O) [Fig. 11(b)] is seen almost throughout the simulation. This is placed near the active site, between His114 and the 3_{10}-helix at the C-terminus which is specific only to angiogenin in the RNase A superfamily. The network involving the residues of third helix [Ile56(O), Glu58(O)], loop [Gly62(O)] and the strand following the loop [Arg70(NH1)] with a water molecule is an interesting one [Fig. 11(d)]. This was pointed out to have a weak stabilizing influence in RNase A in the previous section. However, the corresponding loop is stabilized by a disulfide bond (Cys65—Cys72) in RNase A, which is absent in angiogenin.

Comparison of Invariant Water Positions Between Rnase A and Angiogenin

The protein atoms of RNase A and angiogenin with MRT higher than 1000 are listed in Table II along with their <ASA> and <hydration>. It can be seen that in general the atoms with high MRT have low ASA (near zero value both in MD average and in the crystal structure), and low <hydration>. This is in agreement with earlier studies on myoglobin, where the invariant water molecules were in buried and cleft locations. It can also be seen that many of these atoms are part of a network (Table I). The spatial positions of these atoms are depicted in Figure 12(a) and (b).

A comparison of the two proteins revealed that the catalytic residue His12 (His13), Thr45 (Thr44) and to some extent Lys41 (Lys40) are part of a network, hydrogen-bonding through a tagged water molecule. The interaction between the first helix and the loop before the C-terminal strand is also stabilized by water molecules in both the proteins. This region is crucial for the entry of the ligand into the active site of the proteins. The differences are in terms of the extent of network involving water molecules (angiogenin has more invariant sites for water with high MRT than RNase A) and also in other structural locations of the proteins. Thus, the interaction with water molecules in catalytically important regions are conserved in both proteins and differences are seen in the extent of network and in locations other than the RNA cleaving catalytic site, supporting the involvement of water molecules in various functions of the proteins of RNase A superfamily.

SUMMARY

The hydration patterns in RNase A and angiogenin have been investigated using long-time (5.4 ns) molecular dynamics simulations. The interactions between the protein polar atoms and water molecules have been analyzed at atomic level by tagging all the water molecules in the simulations.

The capacity of a protein polar atom to interact with a single tagged water molecule for the maximum length of time is evaluated and this parameter is denoted as MRT (maximum residence time). A majority of protein polar atoms exhibited MRT less than 250 (approximately equal to 250 ps in time), while a few of them interact with a specific water molecule for almost the entire simulation period (MRT $>$ 4500). Interestingly, the onset of some of the stable interactions can occur well into the simulation after about 1ns, underscoring the importance of long-time simulation to identify key protein–water interactions. We have shown that a simultaneous application of a spatial criterion (distance between the protein polar atom and oxygen of water molecule less than 3 Å) and a temporal criterion (the same water molecule remains attached to the protein atom for a period of 500 ps or more) leads to strong orientation constraints akin to that expected for hydrogen bonds. Our simple scheme aids in the analysis of hydrogen bonds from huge amounts of data produced in the long-time simulations.

Strikingly, the atoms with high MRT ($>$ 1000) have low accessible surface area and low mean hydration. Further, such polar atoms are associated with networks involving several protein atoms and one or more water molecules. This result is consistent with earlier results on myoglobin by Pettit and co-workers,[19] who concluded that water molecules with high residence times are associated with buried atoms in the cleft regions of the protein.

Protein–water networks have been identified by a clustering algorithm and their structural locations have been analyzed. Angiogenin is found to have more protein–water networks than RNase A. The catalytically important residues (His12, Thr45, Lys41 in RNase A and the equivalent

TABLE II. The Characteristics of Atoms with MRT > 1000 in RNase A and Angiogenin

System	Atom	⟨ASA⟩[a] (Å^2)	X-ASA[b] (Å^2)	⟨Hydration⟩	MRT
RNase A	His12 (ND1)	1.6	0.0	0.85	4080
	Thr45 (N)	3.3	1.4	0.55	1966
	Thr45 (O)	0.0	0.0	0.61	3042
	Ile81 (O)	1.5	1.8	0.66	1130
	Asp83 (OD2)	7.3	8.5	1.65	1250
	Ser90 (OG)	0.3	0.0	0.70	1193
	Gln101 (OE1)	1.7	37.7	0.61	1029
	Pro117 (O)	1.2	1.3	0.73	2242
	Asp121 (OD1)	3.7	6.0	1.08	1154
Angiogenin	Gln1 (O)	12.0	8.5	1.01	1260
	His13 (ND1)	0.2	0.0	0.72	3349
	Asp15 (OD2)	1.2	2.1	0.94	2071
	Lys17 (O)	9.1	17.3	1.38	1841
	Gly20 (O)	3.1	2.3	0.97	1121
	Arg32 (NH1)	22.4	26.3	1.14	1489
	Thr44 (N)	0.0	0.0	0.50	2342
	Thr44 (O)	0.0	0.0	0.59	2837
	His47 (O)	9.3	1.6	1.06	1106
	Asn49 (O)	0.7	0.0	0.40	1593
	Ile56 (O)	0.0	0.0	0.85	4250
	Glu58 (O)	0.0	0.0	0.58	2875
	Arg70 (NH1)	13.5	0.0	1.52	4068
	Ser74 (O)	0.0	0.0	0.57	2665
	Thr79 (OG1)	3.9	0.0	1.06	1216
	Val104 (N)	0.2	0.2	0.61	3053
	Val104 (O)	0.0	0.0	0.68	3394
	Asp116 (O)	0.3	0.0	0.97	4500
	Gln117 (NE2)	0.6	0.0	0.35	1418

[a]The mean accessible surface area.
[b]X-ASA: Accessible surface area in the crysytal structure. ⟨ASA⟩ is in agreement with X-ASA, with the exception of atoms Gln101(OE1) in RNase A and Arg70(NH1) in angiogenin.

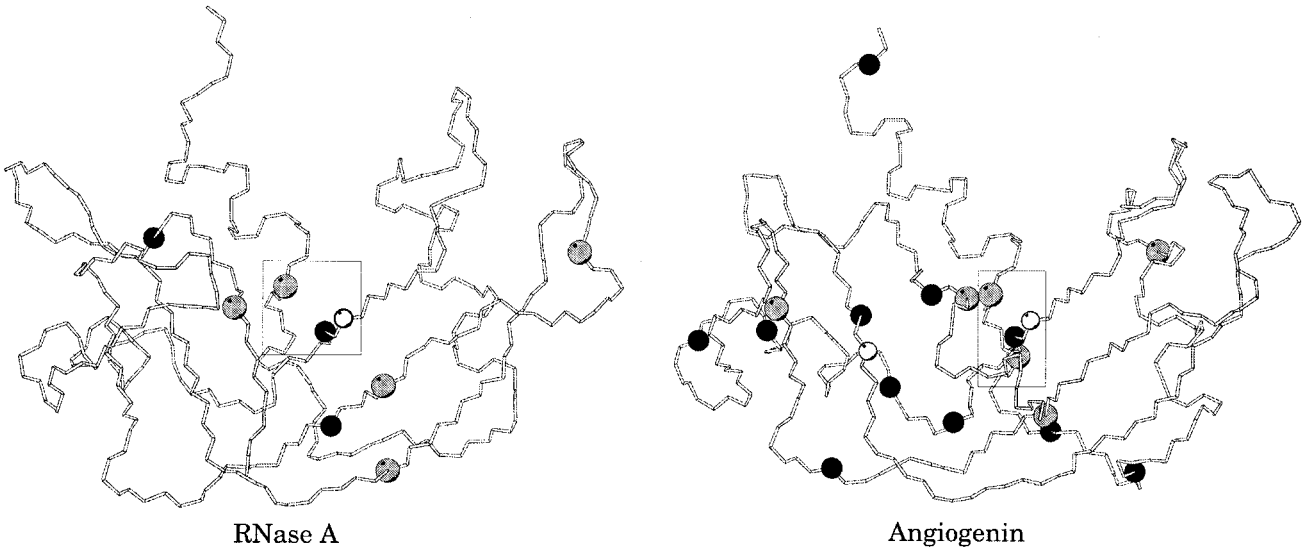

Fig. 12. The spatial locations of the residues with MRT ≥ 1000 are shown for (**a**) RNase A and (**b**) angiogenin on their backbone. Black, white, and grey spheres correspond to the main-chain oxygen atom, main-chain nitrogen atom and the side-chain atom respectively. The active-site residues in both proteins are enclosed in a box.

residues in angiogenin) form a strong protein–water network in both the proteins, suggesting its relevance to the catalytic property of cleaving RNA. Furthermore, the ligand-binding region—the N-terminus helix and the loop preceding the C-terminal strand—is found to be stabilized by water molecule(s) in both the proteins.

ACKNOWLEDGMENTS

Partial support for this project was received from the Department of Science and Technology (DST), India [No SP/50/D-108/98] and from the Department of Biotechnology (DBT), India, through Computational Genomics Initiative grant. Computational facilities from the Super Computer Education and Research (SERC), Indian Institute of Science and the NMITLI project of the Council for Scientific and Industrial Research (CSIR), India (which provided the Origin300 server) are acknowledged. One of us (B.S.S.) also thanks CSIR (India) for the fellowship.

REFERENCES

1. Westhof E, editor. Topics in molecular and structure biology: water and biological macromolecules. London: MacMillan Press Ltd.; 1993.
2. Levitt M, Park BH. Water: now you see it, now you don't. Structure 1993;1:223–226.
3. Zhang XJ, Matthews BW. Conservation of solvent-binding sites in 10 crystal forms of T4 lysozyme. Protein Sci 1994;3:1031–1039.
4. Simmerling CL, Elber R. Computer determination of peptide conformations in water: different roads to structure. Proc Natl Acad Sci USA 1995; 92:3190–3193.
5. Sadasivan C, Nagendra HG, Vijayan M. Plasticity, hydration and accessibility in ribonuclease A. The structure of a new crystal form and its low-humidity variant. Acta Crystallogr D Biol Crystallogr 1998;54:1343–1352.
6. Biswal BK, Sukumar N, Vijayan M. Hydration, mobility, and accessibility of lysozyme: structures of a pH 6.5 orthorhombic form and its low-humidity variant and a comparative study involving 20 crystallographically independent molecules. Acta Crystallogr D Biol Crystallogr 2000;56:1110–1119.
7. Prasad BV, Suguna K. Role of water molecules in the structure and function of aspartic proteinases. Acta Crystallogr D Biol Crystallogr 2002;58:250–259.
8. Tarek M, Tobias DJ. The dynamics of protein hydration water: a quantitative comparison of molecular dynamics simulations and neutron-scattering experiments. Biophys J 2000;79:3244–3257.
9. Chatake T, Ostermann A, Kurihara K, Parak FG, Niimura N. Hydration in proteins observed by high-resolution neutron crystallography. Proteins 2003;50:516–523.
10. Austin RH, Roberson MW, Mansky P. Far-infrared perturbation of reaction rates in myoglobin at low temperatures. Phys Rev Lett 1989; 62:1912–1915.
11. Massi F, Straub JE. Structural and dynamical analysis of the hydration of the Alzheimer's beta-amyloid peptide. J Comp Chem 2002;24:143–153.
12. Bizzarri AR, Cannistraro S. Molecular dynamics of water at the protein-solvent interface. J Phys Chem B 2002;106:6617–6633.
13. Makarov V, Pettitt BM. Solvation and hydration of proteins and nucleic acids: a theoretical view of simulation and experiment. Acc Chem Res 2002;35:376–384.
14. Kombo DC, Young MA, Beveridge DL. Molecular dynamics simulation accurately predicts the experimentally-observed distributions of the (C, N, O) protein atoms around water molecules and sodium ions. Proteins 2000;39:212–215.
15. Denisov VP, Halle B. Protein hydration dynamics in aqueous solution. Faraday discuss 1996;103:227–244.
16. Rocchi C, Bizzarri AR, Cannistraro S. Water residence times around copper plastocyanin: a molecular dynamics simulation approach. Chem Phys 1997;214:261–276.
17. Bizzarri AR, Cannistraro S. Molecular dynamics simulation evidence of anamolous diffusion of protein hydration water. Phys Rev E 1996; 53:R3040–R3043.
18. Bizzarri AR, Cannistraro S. Anamolous and anisotropic diffusion of plastocyanin hydration water. Europhys Lett 1997;37:201–206.
19. Makarov VA, Andrews BK, Smith PE, Pettit BM. Residence times of water molecules in the hydration sites of myoglobin. Biophys J 2000; 79:2966–2974.
20. Muegge I, Knapp EW. Residence times and lateral diffusion of water at protein surfaces: application to BPTI. J Phys Chem 1995;99:1371–1374.
21. Garcia AE, Stiller L. Computation of the mean residence time of water in the hydration shells of biomolecules. J Comp Chem 1993;14:1396–1406.
22. Roe SM, Teeter MM. Patterns for prediction of hydration around polar residues in proteins. J Mol Biol 1993;229:419–427.
23. Madhusudhan MS, Vishveshwara S. Deducing hydration sites of a protein from molecular dynamics simulations. J Biomol Struct Dyn 2001;19:105–114.
24. Madhusudhan MS, Sanjeev BS, Vishveshwara S. Computer modeling and molecular dynamics simulations of ligand bound complexes of bovine angiogenin: dinucleotide topology at the active site of RNase a family proteins. Proteins 2001;45:30–39.
25. Vishveshwara S, Madhusudhan MS, Maizel Jr JV. Short-strong hydrogen bonds and a low barrier transition state for the proton transfer reaction in RNase A catalysis: a quantum chemical study. Biophys Chem 2001;89:105–117.
26. Wlodawer A, Svensson LA, Sjolin L, Gilliland GL. Structure of phosphate-free ribonuclease A refined at 1.26 Å. Biochemistry 1988;27:2705–2717.
27. Leonidas DD, Shapiro R, Allen SC, Subbarao GV, Veluraja K, Acharya KR. Refined crystal structures of native human angiogenin and two active site variants: implications for the unique functional properties of an enzyme involved in neovascularisation during tumour growth. J Mol Biol 1999;285:1209–1233.
28. AMBER, Version 7, University of California, San Francisco; 2002
29. Cheatham III TE, Cieplak P, Kollman PA. A modified version of the Cornell et al. force field with improved sugar pucker phases and helical repeat. J Biomol Struct Dyn 2002;16:845–861
30. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. J Chem Phys 1983;79:926–935.
31. Darden TA, York DM, Pedersen LG. Particle mesh Ewald: An N. log (N) Method for Ewald sums in large systems. J Chem Phys 1993;98:10089–10092.
32. Schoenborn BP, Garcia A, Knott R. Hydration in protein crystallography. Prog. Biophys Mol Biol 1995;64:105–119.
33. Blake CCF, Pulford WCA, Artymuik PJ. X-ray study of water in crystals of lysozyme. J Mol Biol 1983;167:693–723.
34. Baker EN, Hubbard RE. Hydrogen bonding in globular proteins. Prog Biophys Mol Biol 1984;44:97–179.
35. Narsingh D. Graph theory with applications to engineering and computer science. New York: Prentice-Hall; 1974. p 274–276.
36. NACCESS, Version 2.1.1, Department of Biochemistry and Molecular Biology, University College London; 1996.
37. Brunne RM, Liepish E, Otting G, Wuthrich K, van Gunsteren WF. Hydration of proteins: a comparison of residence times of water molecules solvating the bovine pancreatic trysin inhibitor with theoretical model calculations. J Mol Biol 1993;231:1040–1048.
38. Kovacs H, Mark AE, van Gunsteren WF. Solvent structure at a hydrophobic protein surface. Proteins 1997;27:395–404.
39. Rupley JA, Careri G. Protein hydration and function. Adv Protein Chem 1991;41:37–172.
40. Kraulis PJ. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. J Appl Crystallogr 1991;24: 946–950.