

# Computer Modeling and Molecular Dynamics Simulations of Ligand Bound Complexes of Bovine Angiogenin: Dinucleotide Topology at the Active Site of RNase A Family Proteins

M.S. Madhusudhan, B.S. Sanjeev, and S. Vishveshwara\*

*Molecular Biophysics Unit, Indian Institute of Science, Bangalore, India*

**ABSTRACT** We have undertaken the modeling of substrate-bound structures of angiogenin. In our recent study, we modeled the dinucleotide ligand binding to human angiogenin. In the present study, the substrates CpG, UpG, and CpA were docked onto bovine angiogenin. This was achieved by overcoming the problem of an obstruction to the B1 site by the C-terminus and identifying residues that bind to the second base. The modeled complexes retain biochemically important interactions. The docked models were subjected to 1 ns of molecular dynamics, and structures from the simulation were refined by using simulated annealing. Our models explained the enzyme's specificity for both B1 and B2 bases as observed experimentally. The nature of binding of the dinucleotide substrate was compared with that of the mononucleotide product. The models of these complexes were also compared with those obtained earlier with human angiogenin. On the basis of the simulations and annealed structures, we came up with a consensus topology of dinucleotide ligands that binds to human and bovine angiogenins. This dinucleotide conformation can serve as a starting model for ligand-bound complex structures for RNase A family of proteins. We demonstrated this capability by generating the complex structure of CpA bound to eosinophil-derived neurotoxin (EDN) by fitting the consensus topology of CpA to the crystal structure of native EDN. *Proteins* 2001;45:30–39. © 2001 Wiley-Liss, Inc.

**Key words:** bovine angiogenin-substrate (CpG/UpG/CpA) complex; docking; molecular dynamics (MD); consensus conformation; substrate specificity; simulated annealing

## INTRODUCTION

A number of proteins belonging to the RNase A superfamily perform diverse biological functions. Despite the diverse nature of their functions they are structurally and sequentially similar to RNase A, conserving the RNase active site. The cleavage of RNA is a function they share in common, and this is crucial for their biological activity.<sup>1</sup> To understand the similarities and diversities in their func-

tions, we have been investigating the nature of ligand binding in these proteins<sup>2–7</sup> for ribonuclease activity at a molecular level. Angiogenin is an interesting member of the RNase A family. It is involved in angiogenesis and shares 33% sequence identity with bovine pancreatic ribonuclease.<sup>8</sup> X-ray structures of angiogenin are available only in the native, uncomplexed form from human and bovine sources.<sup>9–11</sup> Bovine angiogenin is 64% identical in sequence to its human counterpart.<sup>12</sup> The structures of the two angiogenins are very similar, and they both take up the RNase A fold. Both proteins have the same function, which is to induce the formation of blood vessels. Bovine angiogenin also has ribonucleolytic activity but is reportedly enzymatically less active than human angiogenin.<sup>9</sup> Subtle differences in the sequence and architecture tell them apart. In our previous study,<sup>7</sup> we focused on the dinucleotide substrate binding to human angiogenin. Here we discuss our efforts on bovine angiogenin and compare the nature of binding in the two systems.

One major difference in the structures of bovine and human angiogenins is in the C-terminus region. In the bovine protein, this region does not adhere to any regular secondary structure<sup>9</sup> and is unstructured, whereas in the human protein it is a 3–10 helix.<sup>10</sup> The loop connecting strands 2 and 3, comprising residues 67–69 has been postulated as important for angiogenin to recognize receptors on cell surfaces.<sup>8</sup> The amino acid sequence of this region is not conserved between the human and bovine systems. The sequence is R67, G68, and D69 in bovine and R66, E67, and N68 in human angiogenin. In the human case, our recent studies on docking of substrates<sup>7</sup> have shown that this region interacts with the second base of the substrate.

In this study, we docked dinucleotide ligands (substrates) onto bovine angiogenin. The substrates docked are CpG, UpG, and CpA, shown schematically in Figures 1(a–c). It is known that bovine angiogenin prefers a pyrimidine in the first base position, with a preference of cytosine to uracil among the pyrimidines. The second base specificity is for a purine, and guanine is preferred over

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

Received 11 February 2000; Accepted 15 May 2001

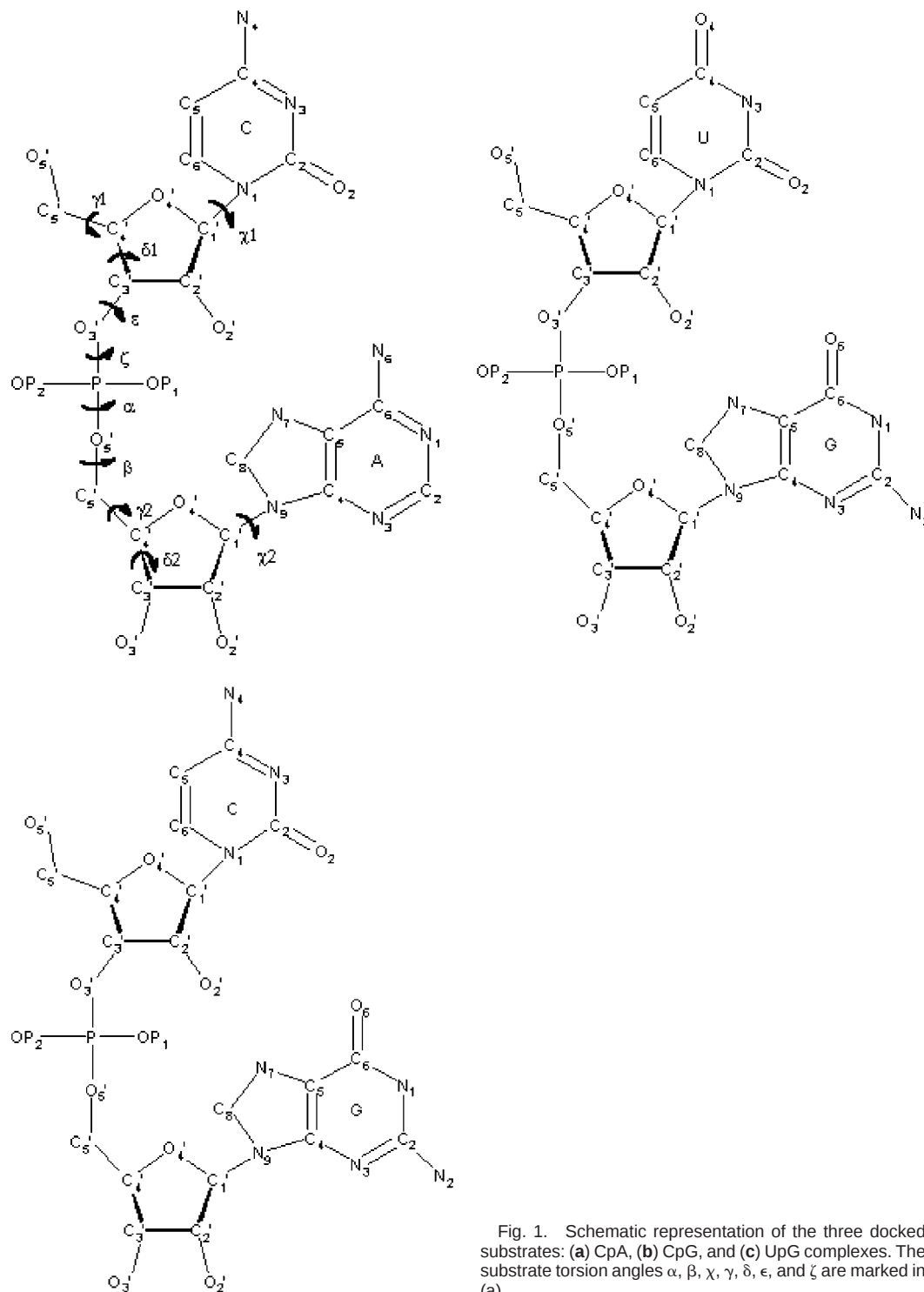


Fig. 1. Schematic representation of the three docked substrates: (a) CpA, (b) CpG, and (c) UpG complexes. The substrate torsion angles  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$ ,  $\epsilon$ , and  $\zeta$  are marked in (a).

adenine,<sup>9</sup> unlike in human angiogenin where adenine is preferred over guanine.<sup>10</sup> The residues that interact with the first base are conserved in the human and bovine systems, whereas those that interact with the second base are not. In fact, this is also true in the other proteins in RNase A family. In this modeling study, we predict the residues of bovine angiogenin that interact with the sec-

ond base. We submitted our docked structures to molecular dynamics (MD) simulations for 1 ns each. Snapshots from the MD simulation are then taken up for refinement by simulated annealing. From an analysis of the docked structures and the MD simulation trajectories we investigate details of protein-substrate interactions and examine the basis for bovine angiogenin's substrate specificity. The

**TABLE I. Number of Atoms and Physical Dimensions of the Systems in Which MD Simulations Were Conducted**

| System                  | Total no. of atoms<br>(no. of waters) | Box size ( $\text{\AA}^3$ )       |
|-------------------------|---------------------------------------|-----------------------------------|
| Bovine angiogenin       | 11,741 (3243)                         | $56.42 \times 55.80 \times 43.10$ |
| Bovine angiogenin + CpG | 11,702 (3029)                         | $58.48 \times 52.82 \times 42.68$ |
| Bovine angiogenin + CpA | 11,731 (3219)                         | $58.22 \times 53.30 \times 43.49$ |
| Bovine angiogenin + UpG | 11,770 (3232)                         | $57.97 \times 53.17 \times 43.42$ |

results are also compared with the docked structures of substrates with human angiogenin.<sup>7</sup> A general, consensus topology of dinucleotide binding to RNase A family of proteins has been obtained from the simulations and annealed structures. Such a conformation is shown to fit reasonably well into the active site of eosinophil-derived neurotoxin (EDN), a protein belonging to RNase A family for which ligand bound structures are not yet available.

## MATERIALS AND METHODS

The substrates (CpA, CpG, and UpG) docked onto bovine angiogenin are shown in Figures 1(a–c). The torsion angles of the substrate that were varied during the docking procedure are marked in Figure 1(a). The starting structure for the native bovine angiogenin was a crystal structure<sup>9</sup> (PDB code 1AGI), and the starting template for docking the dinucleotide substrate was that of the modeled bovine angiogenin-monomonucleotide complex.<sup>5</sup> The second nucleotide was built onto the template of the existing first by using a grid search method similar to that described in the previous study.<sup>7</sup> Steric-free models possessing the important protein–ligand interactions were chosen for all three ligand-bound complexes of bovine angiogenin. The selected complexes were subjected to energy minimization and 1 ns of MD. The sizes of the systems in this study are listed in Table I. Random snapshots of the simulation were taken up for refinement by simulated annealing. All simulations were performed by using the AMBER 4.1 suite of programs.<sup>13</sup> The detailed protocol for the MD and simulated annealing have been discussed in detail in the previous study on human angiogenin-substrate modeling.<sup>7</sup>

## RESULTS

### Docking of Substrates

As in the case of docking dinucleotide structures onto human angiogenin,<sup>7</sup> in the bovine case also there are two major hurdles. The first one is the obstruction of the B1-binding site by the side chain of E118, and the second is the identification of residues that interact with the second nucleotide base B2. We have addressed the first question in depth in earlier studies.<sup>5,6</sup> The entire C-terminus region was shifted by changing the  $\psi_{117}$  torsion angle. This resulted in a movement of the C $\alpha$  of E118 by 3.27 Å from its original position in the crystal structure. This change facilitates the docking of the first pyrimidine base onto the

B1- and P1-binding sites. The docking of both uracil and cytosine in the first base position are investigated to understand the origin of substrate specificity of bovine angiogenin. The position of the mononucleotides in the active site from our earlier investigations<sup>5</sup> formed the basis for building the dinucleotide substrates. Some of the protein–ligand interactions are presented in Figure 2. The important interactions of the base B1 with T45, K41 with ribose R1, and H14 with P1 are retained. The second base of the dinucleotide is built up, and its vicinity is scanned for polar groups with which it can possibly interact. Two of the residues that constitute the loop between strands 2 and 3, R67 and D69, are within hydrogen bond-interacting distance of the second nucleotide. A thorough grid search is conducted in torsion angle space of the ligand and protein side chains to find optimum docking conformations. Of the various steric-free conformations of the ligand, we have chosen the complexes in which B2 interacts with residues R67 and D69 through hydrogen bonds for further investigations by MD studies. In this model, we have optimal Van der Waals interactions as well as a maximum number of hydrogen bonds. The side chain of F116 stacks with B1, as in the case of the RNase A-UpA complex, whereas the side chain of H115 stacks with the second base B2. This model does not have the second nucleotide base interacting with either E109 or R6. It has been shown experimentally in the case of human angiogenin that the equivalent residues of E109 and R6 (E108 and R5) probably do not participate in interactions with the dinucleotide base.<sup>14</sup> However, it has been suggested that R5 bind to P2.

### MD Simulations of the Complexes

One representative structure of each of the modeled protein-substrate complexes (CpG, CpA, and UpG) was taken up for minimization followed by a 1-ns MD simulation run. The simulations gave rise to stable trajectories, indicating that the system has equilibrated as seen in Figure 3. Also notable is the fact that the root-mean-square deviation (RMSD) trajectory of the CpG complex is always lower in value than those of the other two complexes. Figure 4 shows a comparison of the residue-wise RMS fluctuation of the complexes of three substrates, CpG, CpA, and UpG with each other and with that of the native bovine angiogenin simulation. Note that the fluctuations in the CpG-bound complex is the least compared with the other structures. The residues of the protein in the UpG structure seemingly have larger fluctuations. These observations that were also made in the case of mononucleotide-bound complexes<sup>5</sup> and the dinucleotide complexes of human angiogenin<sup>7</sup> are in agreement with the experimentally observed nucleotide preference of bovine angiogenin. Low RMS fluctuations are indicative of tighter bound complexes.

We have analyzed the MD trajectories for stable hydrogen bonds between the ligand and the protein. These hydrogen bonds have been listed in Table II and have also been used as input for annealing the complex structures. Some of the protein-CpG interactions are graphically

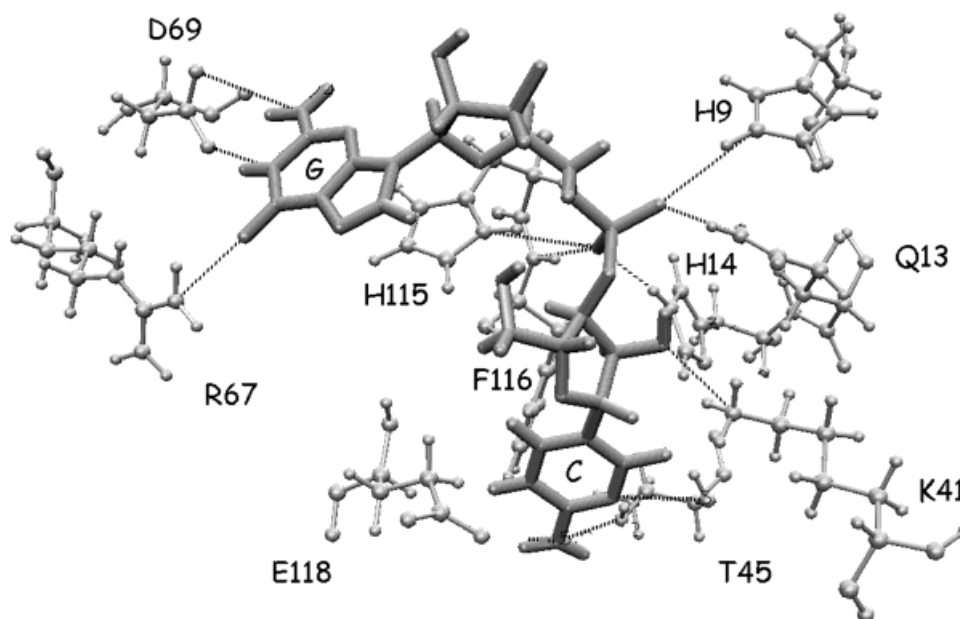


Fig. 2. Residues of bovine angiogenin (in ball and stick representation<sup>17</sup>) that interact with the ligand (in stick representation) CpG obtained by modeling and molecular dynamics. The dotted lines indicate the protein-ligand hydrogen bonds.

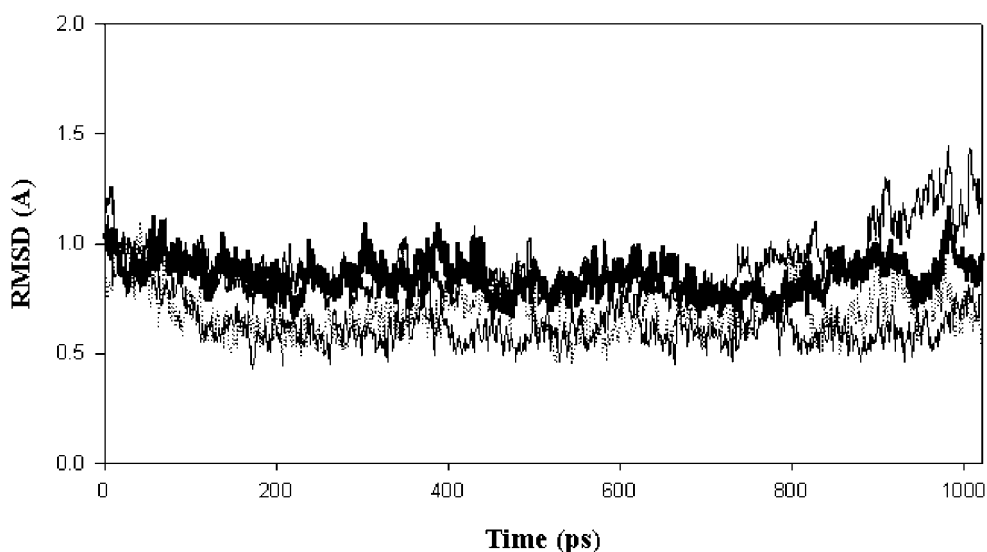


Fig. 3. RMSD trajectory of the three substrates-bound angiogenin complexes. The RMSD of CpG is represented in thin line and is the bottom most line in the plot. The CpA RMSD trajectory is shown in thick line and the UpG trajectory in thin. The trajectory of the native angiogenin<sup>18</sup> is shown as a dotted line.

presented in Figure 2. In the MD of all three complexed structures, hydrogen bonds between the catalytic residues H14, H115, and K41 and the substrate are stable throughout the simulation. H14 interacts with the phosphate oxygen, and the 2' OH group of the ribose R1. K41, which assists in proton transfer from the 2' OH group to H14 exhibits a stable hydrogen bond with the 2' OH and the O2 atoms of the first nucleotide. H115, which donates a proton to the phosphate oxygen during catalysis is constantly hydrogen bonded with the phosphate oxygens. Other stable interactions are between the phosphate oxygens and the

side chain of H9 and the main-chain amide nitrogen of F116. The O $\gamma$ 1 and amide nitrogen of T45 are bonded to the ligand in the three complexes.

However, significant differences exist between the hydrogen-bonding patterns seen in the complexes of three docked substrates. In the case of CpA there is an additional interaction of the N4 atom of the first base with the side-chain O $\epsilon$  atoms of the E118. This is an interesting interaction, because E118 was the residue that obstructed binding of the first nucleotide base. This interaction is absent in the case of the UpG complex and occurs for only a

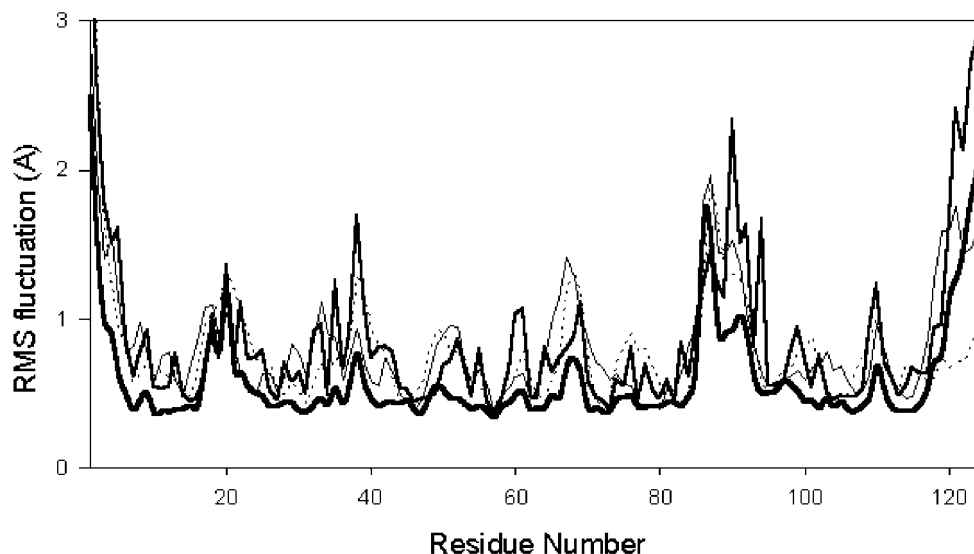


Fig. 4. RMS fluctuations of the residues of bovine angiogenin in the three substrate-protein complexes, CpG (very thick line), CpA (thick line), UpG (dotted line), and the native uncomplexed structure (thin line).

**TABLE II. Hydrogen Bonds Formed Between Bovine Angiogenin-Substrate Complexes During MD Simulations and Used as Distance Restraints for Simulated Annealing**

| Donor acceptor pairs           | CpG complex | CpA complex  | UpG complex |
|--------------------------------|-------------|--------------|-------------|
| 1 Ne2 H9–O2P Py <sup>a</sup>   | ✓           | ✓            | ✓           |
| 2 Ne2 Q13–O2P Py               | ✓           | ✓            | ✓           |
| 3 Ne2 H14–O2' Py               | ✓           | ×            | ✓           |
| 4 Ne2 H14–O1P Py               | ✓           | ✓            | ✓           |
| 5 Nζ K41–O2 Py                 | ✓           | ✓            | ✓           |
| 6 Nζ K41–O2' Py                | ✓           | ✓            | ✓           |
| 7 Oγ1 T45–N4 Py                | ✓           | ✓            | ×           |
| 8 N T45–N3 Py                  | ✓           | ✓            | ×           |
| 9 Oγ1 T45–N3 Py                | ×           | ×            | ✓           |
| 10 N T45–O2 Py                 | ×           | ×            | ×           |
| 11 NH2 R 67–O6 Pu <sup>b</sup> | ✓           | ×            | ×           |
| 12 Oδ1/Oδ2 D69–N1 Pu           | ✓           | Od2 D69–N6 A | ✓           |
| 13 Oδ1/Oδ2 D69–N2 Pu           | ✓           | ×            | ✓           |
| 14 Nδ1 H 115–O5' Pu            | ✓           | ✓            | ✓           |
| 15 Ne2 H115–O5' Pu             | ✓           | ✓            | ✓           |
| 16 Nδ1 H115–O1P Py             | ✓           | ✓            | ✓           |
| 17 N F116–O1P Py               | ✓           | ✓            | ✓           |
| 18 Oε1/Oε2 E118–N4 Py          | ×           | ✓            | ×           |

<sup>a</sup>Py, pyrimidine base (cytosine or uracil).

<sup>b</sup>Pu, purine base (guanine or adenine).

very small portion of the simulated time in the CpG complex. It was observed earlier in the MD simulations of complexes of bovine angiogenin with mononucleotides,<sup>6</sup> that this hydrogen bond was formed in the cytosine monophosphate complex and not in the uracil monophosphate complex. Intermittently, the side-chain atoms R43 also forms hydrogen bonds with the atoms of the first base. This is a common feature of all three simulations. The instability of this hydrogen bond leads us to conclude that it could be relatively unimportant to the complex.

Comparing the hydrogen bonds in the CpG and UpG complexes, we notice that the Oγ1 atom of T45 binds to the N4 atom of cytosine and to N3 of uracil. Similarly, the N of T45 binds to the N3 of cytosine but has no corresponding interaction with uracil in the UpG complex, although N of T45 interacted with O2 of uracil at the beginning of the MD simulation. Furthermore, the residue T80 (equivalent residue in bovine angiogenin is I81) in human angiogenin interacts with T44, helping in proper orientation of T44 to interact with the pyrimidine base. The equivalent residue is missing in bovine angiogenin. Such differences could contribute to the first base specificity of angiogenins.

The second base specificity for guanine over adenine is also borne out in the MD simulations of our model. In the CpG complex, there are stable hydrogen bonds with of the guanine base (N1, N2, and O6 atoms) with the side chains of residues R67 and D69. In the CpA complex, only the side chain of D69 interacts with the N6 atom of adenine. UpG retains its interactions of N1 and N2 with D69. However, the O6 interaction with R67, as seen in the case of the CpG substrate, is not seen in this case. Perhaps this is due to the effect of a subtle change in the positioning of B1.

### Simulated Annealing

Snapshots from the MD simulations of the modeled substrate-protein complexes are taken and refined by simulated annealing. We have analyzed 10 annealed structures of each of the three substrate-bound bovine angiogenin complexes. Table II lists the atom pairs whose distance was restrained to within a hydrogen bond-forming range during the annealing procedure. The Cα traces of these models, along with the substrates, are shown in Figures 5(a–c).

The conformation adopted by the substrates in the annealed structures is analyzed in the torsion angles represented in Figure 1(a) and the phase angle of ribose rings. For each complex, the values from the 10 annealed



structures are averaged, and the spread is noted as standard deviation. The mean and the standard deviation values of these substrate conformational parameters in three complexes are listed in Table III. The protein-ligand interaction energies are also reported in this table. Furthermore, the results obtained by similar studies on human angiogenin are also listed in this table for comparison and consolidation. It can be seen from the table that the pucker of ribose, R1(p1) is predominantly  $C'_4$ -exo [ $52^\circ$  to  $65^\circ$ ]. The pucker of ribose R2(p2) varies between  $C'_2$ -exo,  $C'_1$ -endo and  $C'_3$ -endo [ $-38^\circ$  to  $18^\circ$ ]. This indicates a fairly rigid constraint on ribose R1 and flexibility in ribose R2 conformation. This feature is also evident from the related torsion angles  $\delta 1$  and  $\delta 2$  where  $\delta 2$  shows a larger standard deviation than  $\delta 1$ . The orientation of the bases, as shown by  $\chi^1$  and  $\chi^2$ , exhibits distinct preferences, with B1 taking values from  $96^\circ$  to  $-150^\circ$  and B2 from  $88^\circ$  to  $-123^\circ$ . The parameters  $\alpha$  and  $\zeta$  together represent the torsion angles involved in anomeric effect.<sup>15</sup> It is interesting to note that most of the conformations adopted are in the range of (g,g) to (g,t), which had been recognized to be the energetically preferred conformation from the anomeric effect point of view. This leads to an interesting possibility that the substrate specificity is not only in the terms of the bases but also in their mutual orientation with well-defined conformation of the connecting phosphate unit. This explains why only the specific regions of RNA are conformationally amenable for cleavage by RNases. The access of torsion angles  $\beta$ ,  $\gamma 2$  and  $\epsilon$  in a limited range is probably required to maintain the overall topology of the dinucleotide substrates.

The protein substrate hydrogen bond interactions that were discussed from the MD trajectories also hold for the annealed structures. In the models of the UpG complex, there is no stable hydrogen bond formed between the O2 atom of the first base with N of T45, as there is in the case of human angiogenin-UpA complex. The only stable interaction with the first base is that of the N3 atom with O $\gamma 1$  of T45. In the CpG and CpA complexes, the O $\gamma 1$  and backbone N atom of T45 of bovine angiogenin interact with the N4 and N3 atoms of the first base. The cytosine base has two interactions, whereas the uracil has only one. The interaction energies (Table III) show that the CpG complex is marginally more stable than the CpA complex, and the two of them are more stable than the UpG complex. The magnitude of this enthalpy component of energy is similar to the human angiogenin case (given in Table III) where the replacement of the first nucleotide pyrimidine base of cytosine by uracil also led to about 60 kcal/mol difference in favor of cytosine. Based on these interactions and energetics, we infer that the substrate specificity of bovine angiogenin for the first base stems from minor differences in interactions of B1. This subtle shifting of the first base results in a significant difference in overall interaction energy.

Fig. 5. The  $C\alpha$  trace of the protein in the (a) CpG, (b) CpA, and (c) UpG complexes. The substrates are shown in stick representation.<sup>17</sup>

**TABLE III. Conformational Parameters [Fig. 1(a)] Obtained by Averaging Simulated Annealed Structures [Mean (Standard Deviation)] of the Complexes of Dinucleotides With Bovine and Human Angiogenin Values<sup>†</sup>**

|                 | Bovine          |                 |                 | Human                       |                            |
|-----------------|-----------------|-----------------|-----------------|-----------------------------|----------------------------|
|                 | CpA             | CpG             | UpG             | CpA                         | UpA                        |
| P1 <sup>a</sup> | 53.79 (13.12)   | 65.17 (6.51)    | 60.68 (10.09)   | 56.24 (2.10)                | 51.87 (16.14)              |
| P2 <sup>b</sup> | -37.58 (48.66)  | 18.42 (54.69)   | 18.08 (46.66)   | -3.59 (42.99)               | 1.36 (56.04)               |
| $\chi_1$        | 96.61 (12.63)   | 132.53 (22.28)  | 106.34 (35.12)  | -150.24 (5.81)              | -170.40 (14.96)            |
| $\chi_2$        | -88.01 (19.07)  | -113.67 (10.62) | -116.90 (11.32) | -92.61 (19.01)              | -123.21 (56.50)            |
| $\alpha$        | 43.91 (37.40)   | 63.12 (7.64)    | 70.38 (51.31)   | 76.43 (23.95)               | 84.01 (26.07)              |
| $\beta$         | 177.21 (10.18)  | 169.53 (8.97)   | 168.88 (28.78)  | 159.93 (30.08)              | 127.74 (44.77)             |
| $\gamma_2$      | 157.31 (48.91)  | 175.24 (8.09)   | 157.43 (48.11)  | 163.73 (39.09)              | 173.57 (47.93)             |
| $\delta_1$      | 80.91 (5.00)    | 90.36 (6.15)    | 78.72 (5.26)    | 79.77 (7.40)                | 86.77 (5.93)               |
| $\delta_2$      | 108.04 (23.59)  | 89.83 (23.35)   | 93.33 (28.86)   | 118.14 (30.73)              | 111.63 (30.67)             |
| $\epsilon$      | -179.18 (8.93)  | -159.63 (6.09)  | -166.94 (16.51) | -160.22 (21.31)             | -136.63 (28.11)            |
| $\zeta$         | 127.19 (14.38)  | 98.14 (9.56)    | 107.17 (24.65)  | 100.51 (12.30) <sup>c</sup> | 82.27 (24.32) <sup>c</sup> |
| IE <sup>d</sup> | -369.17 (17.89) | -378.99 (14.44) | -332.48 (26.02) | -459.43 (13.69)             | -399.99 (20.78)            |

<sup>†</sup>Values reported in Ref. 7 (Table IVa and IVb).<sup>a</sup>Phase angle of ribose R1.<sup>b</sup>Phase angle of ribose R2.<sup>c</sup>The  $\zeta$  values were wrongly reported in Ref. 7.<sup>d</sup>Protein ligand interaction energy (kcal/mol).

In our models, we see that the binding of the second base is  $\approx 10$  kcal/mol in favor of guanine over adenine. This difference in energy can be partially accounted for by the additional interaction of R67 with B2 in the case of guanine. In both CpG and CpA complexes, the O $\delta$  atoms of D69 participate in hydrogen bonds with the N1 and N2 atoms of guanine and the N6 atom of adenine. In the CpG complex, the NH2 atom of R67 interacts with the O6 atom of guanine. Our models differentiate between adenine and guanine for the second base, according to experimental evidence.

In Table III we can see that generally the values of the torsion angles of the substrates in the CpG complex fluctuate to a lesser extent than the other two complexes. The substrates CpG, CpA, and UpG from the annealed complexes are superimposed in Figures 6(a–c). It can be seen that the fluctuations of B2 in UpG > CpA > CpG. The fluctuations in UpG start from the R1 position. This hierarchy of fluctuations is also indicative of substrate preference of bovine angiogenin.

## DISCUSSION

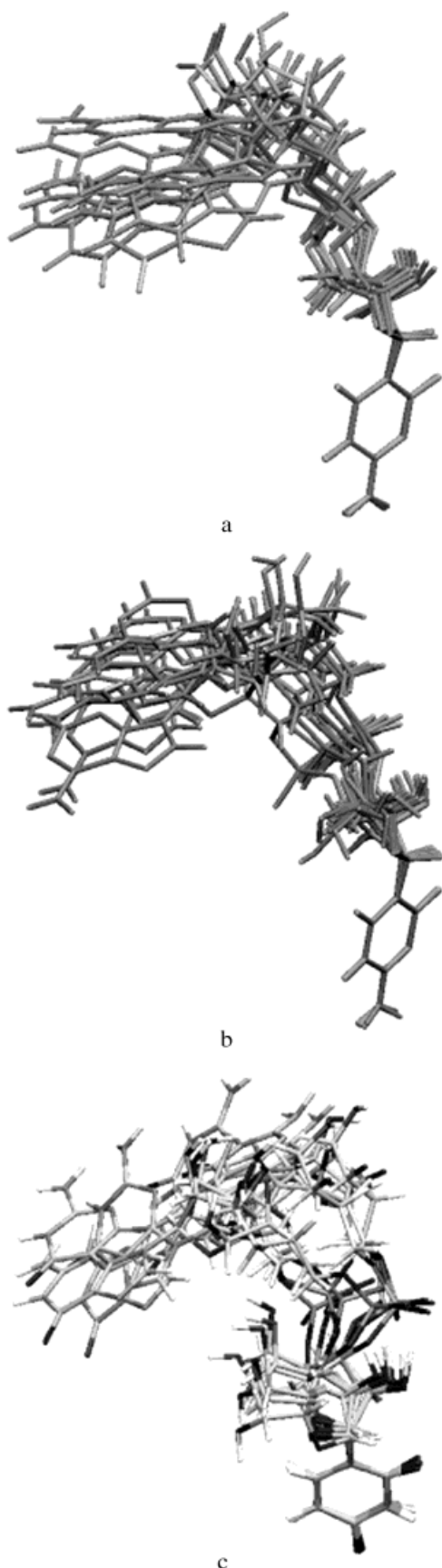
### Comparing With Human Angiogenin-Dinucleotide Complexes

The conservation of residues in the B1 and P1 sites between the human and bovine systems warrants conservation in interactions between the base and the phosphate of the ligand and the protein. In the CpA complex of bovine angiogenin, the N4 atom of the base additionally interacts with the O $\epsilon$  atoms of E118. This has been discussed in an earlier section. No such interaction between Q117, the equivalent of E118 in the human, and the cytosine base was obtained in the modeled complexes with human angiogenin. In bovine angiogenin, the sequential neighbor of the catalytic H115 is F116. Nevertheless, in both systems the amide nitrogen of this residue forms a hydrogen bond with the phosphate oxygen of the substrate. In

the bovine case, the side chain of this F115 stacks with the first base, a feature that is different from the human angiogenin in which the corresponding interaction is with a leucine residue. In this regard, the bovine angiogenin-substrate complex is like the RNase A-ligand complexes in which such stacking interactions are formed between the first base and F120.<sup>4</sup>

Major differences between the complexes of the two systems come up in the interactions of the second nucleotide base with protein, which is not surprising, given that the sequence of residues in the regions that interact with the second base is not conserved. The B2 base (both adenine and guanine) in our models interact with the residues R67 and D69. In CpG, the N1 and O6 atoms of the guanine base interact with the O $\delta$  atom of D69 and the NH2 atom of R67. Sometimes the main chain O of D69 interacts with the N2 of guanine. This is the only complex in which both residues R67 and D69 interact with the ligand. In the CpA and UpG complexes, it is only the O and O $\delta$  atoms of D69 that interact with the N1, N2, O6, N6 (in the case of adenine) atoms of the second base. In the human angiogenin complexes, only N68 interacts directly with the second base of the ligand.<sup>7</sup> In the CpA complex, there is a water-mediated hydrogen bond between the second base and E67. R66 is never within direct hydrogen-bonding distance in the human complexes. In both the human and bovine systems, the number of interactions with the most favored substrate, CpA in human and CpG in bovine, are consistently higher than of that with other substrates.

In human angiogenin, it has been experimentally shown that T80 plays an important role in binding ligands. This residue, though not close enough to directly bind the ligand, assists in orienting T44 so that it binds to the first base. The residue equivalent to T80 in the bovine system is I81. The absence of the T80-equivalent residue to direct the T45 to bind to the first base could be one reason why



bovine angiogenin is less active than its human counterpart. The present investigation points to the fact that the differences in substrate specificity is a cumulative effect of the various factors, such as the lack of a specific interaction, a shift in the substrate position that can affect the interaction even if relevant groups are present.

### Comparing Substrate-Bound Complexes With Product-Bound Complexes

In an earlier study,<sup>6</sup> we had looked at bovine angiogenin-monomucleotide complexes. The mononucleotides docked onto angiogenin were CMP and UMP. These monophosphates are not only the starting templates for docking the dinucleotide substrate but also products of the enzymatic cleavage. In the product-bound cases, the phosphate group of the nucleotide brings up one end of the docked ligand and is hence freer to move than it is in the dinucleotide complexes. This free movement of the phosphate group results in protein-ligand interactions that are different when the ligand is a dinucleotide. The catalytic residue H14 interacts with O2' of the ribose R1 in the protein-substrate complex. This is a crucial interaction for catalysis of RNA cleavage.<sup>8</sup> In the catalytic mechanism, a proton is abstracted from O2' of the ribose by H14. This interaction is absent in the product-protein complexes. H14 instead interacts with the phosphate oxygens.

Another difference between the product-bound and substrate-bound complexes is the interaction of the ligand with K41, another catalytic residue. In the substrate-bound complexes, the N $\zeta$  atom of the K41 side chain interacts with the O2' of the first ribose sugar and O2 atom of the pyrimidine base. In the CMP bound complex, the interaction with O2' is not always present in the simulation of the docked structure. In the UMP-bound complex, however, this interaction is a stable one during the simulation. The instability of this interaction in the CMP complex could indicate that binding modes of the products and substrates could be different in subtle ways. We stress here that a substrate-docked model is a more appropriate model than substrate analogs or product complexes to deduce enzyme mechanistic interactions of the protein. Substrate-docked models could also offer a more comprehensive picture to deduce substrate specificity.

### Consensus Substrate Conformation and a Model of EDN-Substrate Complex

We have refined the 10 models of each of the angiogenin-substrate complexes by simulated annealing procedure. From these structures ( $3 \times 10$ ) of bovine angiogenin and structures from the previous human angiogenin complexes ( $2 \times 10$ ), we can derive a consensus conformation for the dinucleotide that can bind to angiogenin. A comprehensive list of the average values (along with standard deviation) of all the substrate torsion angles [indicated in the sche-

Fig. 6. A superimposition of the substrates (a) CpG, (b) CpA, and (c) UpG in the 10 annealed models. Only the atoms of the first base of the dinucleotides are used for superimposition.

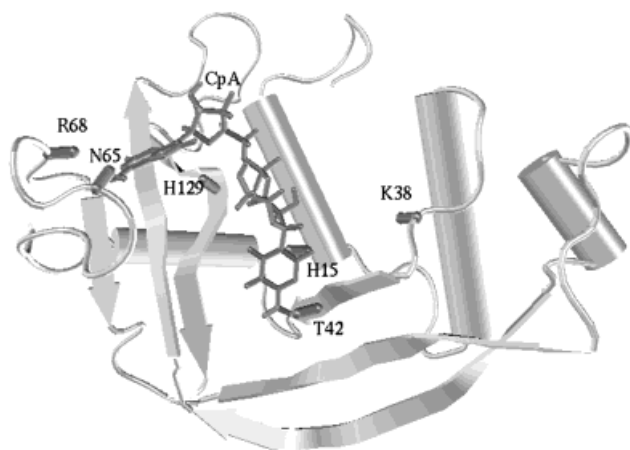


Fig. 7. Modeled structure of EDN-CpA complex obtained by superposing an annealed structure of human angiogenin-CpA on EDN. The protein in cartoon model and CpA in stick model were generated by VMD.<sup>17</sup> The positions of the active-site residues interacting with CpA are marked.

matic diagram Fig. 1(a)] in the 10 models of each complex is given in Table III. The corresponding values obtained in human angiogenin complexes are also reported in Table III. The conformational details of the substrates are discussed in "Simulated Annealing," and it is clear that only limited ranges of conformational parameters are accessed by the annealed structures. These consensus values of the torsion angles describing the dinucleotide conformation results in a defined shape of the substrate required to bind to the enzyme. This desired shape is brought out by the superimposed annealed structures in Figures 6(a–c) and also from the human angiogenin complex structures [Figs. 10(a, b) of Ref. 7]. The difference in interaction energies in different models is not apparently correlated to the difference of any one substrate torsion angle. Subtle variations in the structure seem to account for the range in interaction energies.

Having obtained a consensus dinucleotide conformation that fits into the active site of angiogenins, we have tested the fitting of such a conformation into an RNase A family protein—eosinophil-derived neurotoxin (EDN). This protein acts as a neurotoxin apart from possessing the ribonuclease activity. The crystal structure of human EDN is reported<sup>16</sup> and, as is in the case of angiogenins, no ligand-bound structures are available. We have considered one of the CpA conformations from the human angiogenin CpA complex obtained by simulated annealing procedure. The backbone plus C $\beta$  coordinates of the chosen human angiogenin-CpA complex and EDN are superimposed. The CpA position thus superimposed onto EDN is shown in Figure 7. The fit is remarkably good, and the structure of the complex has many of the desired features. The features, such as the first base interaction with N and O $\gamma$  of T42 (equivalent of T45 in bovine angiogenin), O2' of ribose R1 with N $\epsilon$  of K38 (equivalent of K41), proximity of H15 (equivalent of H14) to O2' of ribose R1, H129 (equivalent of H114) to the second base and phosphate oxygen lend credential to the modeled complex. Furthermore, it is known that the RNase A family proteins differ in their

interaction with the second base. We have earlier seen that the residues interacting with the second base are Q69 and N71 in RNase A, R66 and N68 in human angiogenin, and R67 and D69 in bovine angiogenin, respectively. In our present model of EDN-CpA, we see that the residues N65 and R68 are proximal to the second base, which have the potential to interact with the second base. Thus, the modeled complex is an excellent starting model for future refinement.

## CONCLUSIONS

This is the first report of substrate docking onto bovine angiogenin. In the previous study, we looked at substrate docking onto human angiogenin. The modeled complexes consisted of most of the desired protein-substrate interactions important for the binding of base B1, ribose R1 and phosphate P1 as deduced from experiments. This model also consists of important interactions of angiogenin with the base B2 for which no experimental data are available. The modeled complexes were subjected to 1 ns of MD simulations, and snapshots from these simulations were refined by using simulated annealing. Our models clearly bring out the substrate specificity of the protein in both first and second base positions. This is seen not only from the number of interactions but also from the energy of interaction with the different substrates. A comparison of the substrate-bound complexes with that of the product-bound ones indicate that binding modes in the two could be subtly altered. The substrate-bound complex is a better gauge of enzymatically crucial interaction than the product or substrate analog-bound complexes. The substrate-bound complexes are also better indicators of ligand specificity of the protein. Our models of bovine angiogenin-substrate complexes differentiate effectively between bases at both the first and second base positions. The specificity observed in the models is in accordance with experimental data. Finally, a consensus conformation of the dinucleotide that fits into both the bovine and human angiogenin active sites has emerged. The conformation is described in various torsion angles and possible allowed fluctuations in these values. The conformation of CpA with the well-defined topology obtained from these studies has been shown to fit into the active site of EDN, a protein belonging to RNase A family. The modeled complex not only exhibits the desired protein-substrate interactions but also predicts that the residues N65 and R68 will interact with the second base.

## ACKNOWLEDGMENTS

We thank Dr. K. Seshadri for useful discussions. We also thank Dr. James MNG for providing us with coordinates of eosinophil-derived neurotoxin. We thank the Super Computer Education and Research Centre (SERC) of the Indian Institute of Science for computational facilities and the Department of Science and Technology for supporting the project. B.S.S. thanks CSIR for fellowship.

## REFERENCES

1. Beintema JJ, Breukelman HJ, Carsana A, Furia A. Evolution of vertebrate ribonucleases: ribonuclease A superfamily. In: D'Alessio

- G, Riordan JF, editors. Ribonucleases: structures and functions. New York: Academic Press; 1997. p 245–269.
2. Seshadri K, Rao VSR, Vishveshwara S. Interaction of substrate Uridyl 3',5'-adenosine with ribonuclease A: a molecular dynamics study *Biophys J* 1995;69:2185–94.
  3. Seshadri K, Rao VSR, Vishveshwara S. Characterization of substrate UpA binding to RNase A—computer modelling and energetics approach. *J Biomol Struct Dyn* 1994;12:581–603.
  4. Nadig G, Vishveshwara S. Effects of substrate binding on the dynamics of RNase A: molecular dynamics simulations of UpA bound and native RNase A. *Biopolymers* 1997;42:505–520.
  5. Madhusudhan MS, Vishveshwara S. Modeling of angiogenin—3-NMP complex. *J Biomol Struct Dyn* 1998;16:715–722.
  6. Madhusudhan MS, Vishveshwara S. Molecular dynamics simulations of modeled angiogenin-monomonucleotide complexes. *Curr Sci* 2000;78:852–857.
  7. Madhusudhan MS, Vishveshwara S. Computer modeling of human angiogenin dinucleotide substrate interaction. *Proteins* 2001;42:125–135.
  8. Riordan JF. Structure and function of angiogenin. In: D'Alessio G, Riordan JF, editors. Ribonucleases: structures and functions. New York: Academic Press; 1997. p 445–489.
  9. Acharya KR, Shapiro R, Riordan JF, Vallee BL. Crystal structure of bovine angiogenin at 1.5 Å resolution. *Proc Acad Natl Sci USA* 1995;92:2949–2953.
  10. Acharya KR, Shapiro R, Allen SC, Riordan JF, Vallee BL. Crystal structure of human angiogenin reveals the structural basis for its functional divergence from ribonuclease. *Proc Acad Natl Sci USA* 1994;91:2915–2919.
  11. 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.
  12. Bond MD, Strydom DJ. Amino acid sequence of bovine angiogenin. *Biochemistry* 1989;28:6110–6113.
  13. Pearlman DA, et al. AMBER4.1 San Francisco: University of California, 1995.
  14. Russo N, Acharya KR, Vallee BL, Shapiro R. A combined kinetic and modeling study of the catalytic center subsites of human angiogenin. *Proc Natl Acad Sci USA* 1996;93:804–808.
  15. Gorenstein DG, Findlay JB, Luxon BA, Kar DJ. Stereoelectronic control in C-O and P-O bond breaking processes: ab-initio calculations and speculations on the mechanism of action of ribonuclease A, staphylococcal nuclease and lysozyme. *J Am Chem Soc* 1977;99:3473.
  16. Mosimann SC, Newton DL, Youle RJ, James MNG. X-ray crystallographic structure of recombinant eosinophil-derived neurotoxin at 1.83 Å resolution. *J Mol Biol* 1996;260:540–552.
  17. Humphery W, Dalke A, Schulten K. VMD—visual molecular dynamics. *J Mol Graph* 1996;14.1:33–38.
  18. Madhusudhan MS, Vishveshwara S. Comparison of the dynamics of bovine and human angiogenin: a molecular dynamics study. *Biopolymers* 1999;49:131–144.