J Mol Model (2011) 17:2805–2816
DOI 10.1007/s00894-011-0954-2

ORIGINAL PAPER

Molecular dynamics simulations and binding free energy analysis of DNA minor groove complexes of curcumin

Mathew Varghese Koonammackal · Unnikrishnan Viswambharan Nair Nellipparambil · Chellappanpillai Sudarsanakumar

Received: 17 August 2010 / Accepted: 3 January 2011 / Published online: 2 February 2011

Abstract Curcumin is a natural phytochemical that exhibits a wide range of pharmacological properties, including antitumor and anticancer activities. The similarity in the shape of curcumin to DNA minor groove binding drugs is the motivation for exploring its binding affinity in the minor grooves of DNA sequences. Interactions of curcumin with DNA have not been extensively examined, while its pharmacological activities have been studied and documented in depth. Curcumin was docked with two DNA duplexes, d(GTATATAC)$_2$ and d(CGCGATATCGCG)$_2$, and molecular dynamics simulations of the complexes were performed in explicit solvent to determine the stability of the binding. In all systems, the curcumin is positioned in the minor groove in the A·T region, and was stably bound throughout the simulation, causing only minor modifications to the structural parameters of DNA. Water molecules were found to contribute to the stability of the binding of the ligand. Free energy analyses of the complexes were performed with MM-PBSA, and the binding affinities that were calculated are comparable to the values reported for other similar nucleic acid–ligand systems, indicating that curcumin is a suitable natural molecule for the development of minor groove binding drugs.

Keywords DNA · Curcumin · Molecular dynamics · MM-PBSA · Binding free energy

Electronic supplementary material The online version of this article (doi:10.1007/s00894-011-0954-2) contains supplementary material, which is available to authorized users.

M. V. Koonammackal · U. V. Nellipparambil · C. Sudarsanakumar (✉)
School of Pure and Applied Physics, Mahatma Gandhi University, Kottayam, Kerala 686560, India
e-mail: csudarsan1@sify.com

M. V. Koonammackal
e-mail: mathewkv@gmail.com

Introduction

DNA is known to act as a target for different types of drug molecules [1, 2]. Drug–DNA interactions involve covalent or noncovalent binding in the minor or major grooves, or intercalation [3]. DNA minor groove binding molecules usually have a concave shape which complements the convex shape of the floor of the minor groove, formed by the edge of the bases [4], and these molecules have attracted considerable attention because of their sequence specificity [5]. Moreover, they cause minimal perturbation to the DNA structure [6]. Two well-studied examples of minor groove binders are netropsin and distamycin [7].

Curcumin, 1,7-bis(4-hydroxy-3-methoxyphenyl)-1,6-heptadine-3,5 dione, is a natural coloring pigment that is the major component in the spice *Curcuma longa*. It shows a wide range of pharmacological activities, such as anticancer, anti-inflammatory, antioxidant and pro-oxidant properties [8–10]. The anticancer and antitumor properties of curcumin have been studied in depth, and it has been shown to exhibit therapeutic value for the treatment of various types of cancers, including leukemia and lymphoma [11]. Curcumin is a good inhibitor of angiogenesis [12], which is the fundamental step in the transition of tumors from the dormant state to the malignant state. Aflatoxin is a carcinogen that forms an adduct with DNA, catalyzed by the cytochrome P450 enzymes (CYP450 superfamily of hemoproteins); this adduct formation is inhibited by curcumin [13]. Chemoprevention is the use of natural or synthetic chemicals to suppress, retard or reverse carcinogenesis, with reduced side effects in comparison with classical chemotherapy [14]. There are two types of chemopreventive agents: blocking agents and suppressing agents. Curcumin can act as both, since it has multiple mechanisms of action. Various transcription factors, cytokines, protein kinases, and

enzymes that play some role in inflammation have been shown to be regulated by curcumin [15]. The usual molecular targets of curcumin include various transcription factors, protein kinases, enzymes, growth factors, anti-apoptotic proteins, cell cycle regulatory proteins, etc [16].

The curcumin molecule is similar in shape to the DNA minor groove binding molecules netropsin [17] and distamycin [18] (Fig. 1), and is proposed to bind selectively to AT-rich sequences [19–21]. Netropsin and distamycin are nitrogen rich, which is helpful for forming hydrogen bonds with the receptor DNA, and is advantageous to binding.

It was recently proposed that one of the decisive factors in the drug-binding process is the shape complementarity of the drug with the minor groove of DNA, and the associated favorable van der Waals contribution [22, 23]. While we were performing our calculations, spectroscopic studies on the binding of curcumin with calf thymus DNA and yeast RNA were reported [24]. Also, circular dichroism spectroscopic studies revealed pH-dependent binding of curcumin in the minor grooves of nucleic acids [21]. Evaluating the binding free energy in receptor–ligand complexes is one of the most important steps in theoretical drug design. There are molecular dynamics (MD) simulation-based methods like free energy perturbation (FEP) [25] and thermodynamic integration (TI) that can be used to do this [26], but these are computationally very expensive. Docking methods are less reliable because of the fact that they normally do not take into account the flexibility of the receptors. The MM-PBSA method [27, 28] of evaluating binding affinity is computationally more efficient than FEP and TI [29], and is more reliable than simple empirical scoring methods like XSCORE [30]. It uses the conformations of the receptor, ligand and the complex extracted from the trajectory of a MD simulation, typically in explicit solvent, and the energy calculations are done with the continuum solvent method. The solvation free energy is calculated by solving the Poisson–Boltzmann equation, and entropy contributions are determined with normal mode analysis. However, it was recently reported that the MM-PBSA method has some shortcomings regarding the calculation of entropy and the evaluation of the electrostatic component of the binding free energy [31].

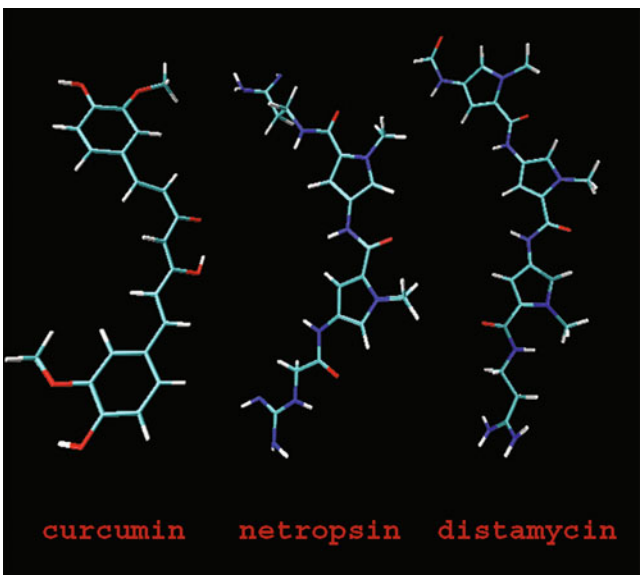

Fig. 1 Shape similarity among curcumin, netropsin, and distamycin

The binding of the ligand with its receptor is a dynamic process, and hence the stability of the complex should be ascertained. We have used docking and MD simulation methods to evaluate the binding of curcumin to AT-rich DNA oligomers, and to calculate the binding energy of the complexes. The MM-PBSA method was successfully applied to calculate the binding free energies of protein–ligand systems [32–34], as well as nucleic acid–ligand systems [35]. NMR and X-ray studies indicate that at least 4 AT base pairs are required to bind a netropsin or a distamycin molecule [36, 37]. Since curcumin has a similar shape to these molecules, we have used two B-form DNA structures with four and six AT base pairs as targets for the curcumin molecule.

Curcumin can exist in diketo and keto-enol forms in solution, and at equilibrium the amount of the keto-enol form is more than 95% [38]. In the solid state the keto-enol form is again more stable, and is responsible for the pharmacological qualities of curcumin [39]. Here we present the results from three unrestrained MD simulations of DNA–curcumin complexes; curcumin in the keto-enol and the diketo forms complexed with $d(GTATATAC)_2$ (at6a and at6b respectively), and the keto-enol form complexed with $d(CGCGATATCGCG)_2$ (at4).

Computational methods

The structure of curcumin was extracted from the Cambridge Crystallographic Data Center (CCDC ID BINMEQ01) [40], and the starting receptor DNAs were taken from the protein data bank (PDB). In the first case, the DNA $d(GTATATAC)_2$ was taken from a drug–DNA complex structure (PDB ID 378d) that contains two distamycin molecules bound side-by-side in the minor groove [41]. The distamycin molecules were removed and curcumin was docked into this duplex in its keto-enol and diketo forms, and the systems were denoted at6a and at6b, respectively. In the second case, the DNA $d(CGCGATATCGCG)_2$ was taken from a complex with netropsin bound to the DNA in the minor groove (PDB ID 1dne) [42]. The netropsin was removed and curcumin in the keto-enol form was docked to the duplex; this system was

denoted at4. In all three cases, docking was done with DOCK6 [43]. This is a previously reported [44] methodology for producing the receptor–ligand complexes for MD simulations and subsequent free-energy analysis. The rigid docking protocol was used and the best-scoring binding mode of the curcumin was used for the simulations.

The MD simulations were done with the *sander* module of the *AMBER 8* suite of programs [45]. The parameters for the curcumin were generated with *ANTECHAMBER1.26* [46] with the general AMBER force-field (*gaff*) [47], and the ff99 force-field [48] was used for the nucleic acids. Three additional simulations (~7ns each) of the complexes (at6a, at6b, and at4) were performed with the parmbsc0 [49] parameters for DNA, and are designated at6a-bsc, at6b-bsc, and at4-bsc, respectively. The parm99bsc is a refinement of the parm99 parameters for the correct representation of α/γ concerted rotation in nucleic acids. A 6 ns simulation of the at6a–distamycin complex was also performed to evaluate the binding free energy. AM1-BCC partial atomic charges were used for curcumin atoms [50]. Even though the AM1-BCC charge method is based on calculations performed at a lower level of theory (the AM1 wavefunction) than HF (6-31G*), the charges are of comparable quality to RESP charges [50], and have been used for many similar studies. They work well with AMBER force-field parameters [51–54]. Figure 2 shows the curcumin molecule in its keto-enol and diketo forms along with atom numbers and the assigned atom types. The atom types and the partial atomic charges assigned to the curcumin atoms are given in Table 1 of the "Electronic supplementary information." The Lennard–Jones parameters from the general AMBER force-field were used for the curcumin atoms. The same simulation protocol was used for all three systems. The systems were built with xLEaP, Na^+ counterions were used to neutralize, and they were solvated in a truncated octahedron box of TIP3P water [55]. The waterbox extends to at least 11 Å from the solute in all directions. The simulations were performed under periodic boundary conditions, and the electrostatic interactions were evaluated with the particle mesh Ewald method [56, 57] using a grid spacing of approximately 1 Å and a direct space sum cutoff distance of 10 Å. Lennard–Jones interactions were also cut off at 10 Å. The built systems were energy minimized, heated to 300 K at constant volume, and then simulated at constant pressure to equilibrate density, with restraints on the DNA and curcumin. The restraints were then gradually decreased, as reported earlier [58], which provided a total equilibration of 500 ps. After that, the production simulations were run for 10 ns, and the atomic coordinates were saved at an interval of 1 ps for further analysis.

The analysis of the trajectories was done with the *ptraj* module, which was used to calculate the rms deviations, hydrogen bonding analysis, hydration analysis, average structures, and to track the backbone torsions and distances as a function of time. Using the grid option in *ptraj*, the hydration and ion association sites were determined. *VMD* [59] was used to visualize the trajectories generated. Principal component analysis (PCA), a statistical technique that is employed to reduce the dimensionality of a problem which involves a significantly large number of degrees of freedom, can be used for the conformational analysis of MD trajectories. PCA of the trajectories were performed to

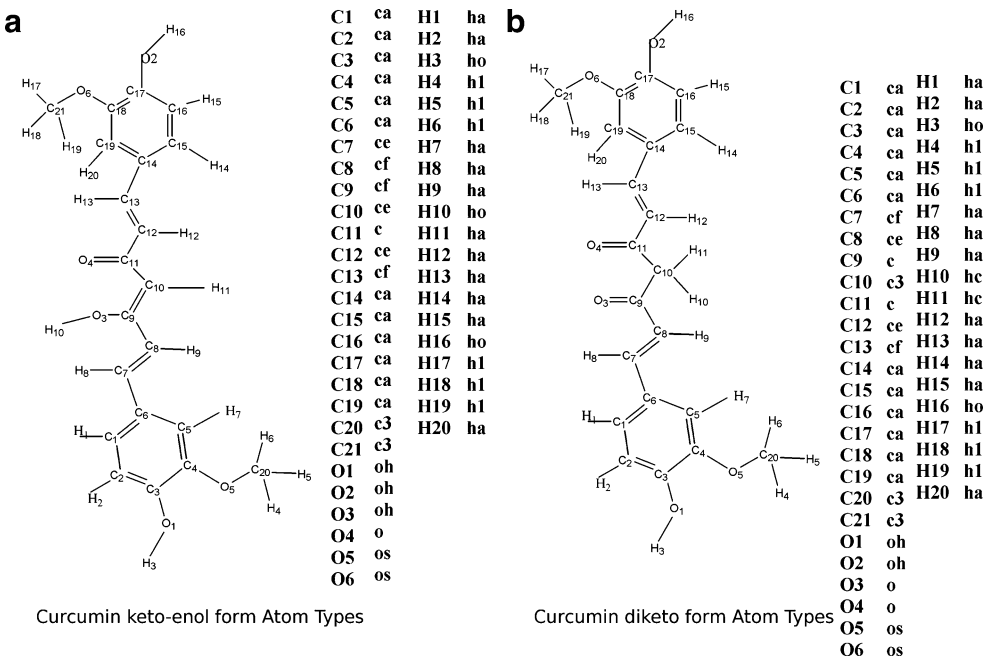

Fig. 2 Atom names and atom types. **a** Curcumin in keto-enol form, **b** curcumin in diketo form

visualize the essential motions and distinguish them from random thermal motions. The mass covariance matrix of the coordinates was built, and their eigenvectors were calculated using *ptraj*. The results were visualized using the Interactive Essential Dynamics (IED) [60] script extension in *VMD*.

The conformations generated were clustered into different groups using *MMTSB* [61], with the rms deviation from the average structure used as the clustering criterion. The conformation with the lowest rms deviation from the centroid of each cluster was used as the representative structure to calculate various structural parameters. The helical parameters were calculated with *3DNA* [62].

Free energy analysis

The binding free energy of a receptor–ligand system can be calculated using the equation [32]

$$\Delta G_{\text{bind}} = G_{\text{complex}} - \left[G_{\text{receptor}} + G_{\text{ligand}}\right], \quad (1)$$

where

$$\Delta G_{\text{molecule}} = \langle E_{\text{MM}} \rangle + \left\langle G_{\text{polar-solvation}} \right\rangle + \left\langle G_{\text{nonpolar-solvation}} \right\rangle - TS. \quad (2)$$

Here, the molecular mechanical energy

$$\langle E_{\text{MM}} \rangle = \langle E_{\text{internal}} \rangle + \langle E_{\text{electrostatic}} \rangle + \langle E_{vdW} \rangle, \quad (3)$$

while

$\left\langle G_{\text{polar-solvation}} \right\rangle$, and $\left\langle G_{\text{nonpolar-solvation}} \right\rangle$

are the polar solvation and nonpolar solvation energies, respectively, and S is the solute entropy. The binding energy can then be represented as

$$\Delta G_{\text{bind}} = \Delta E_{\text{MM}} + \Delta G_{\text{solv}} - T\,\Delta S. \quad (4)$$

Free energy calculations were performed with the MM-PBSA [27, 63] method as implemented in AMBER 8 with the single trajectory approach: only the simulation of the complex was performed, and the snapshots of the complex, receptor and ligand were extracted from this trajectory. The snapshots for the free energy analysis were taken from the last 4 ns of the production run with an interval of 100 ps, which provided a total of 40 frames. The snapshots of curcumin from the trajectory of curcumin alone are also extracted and are used to calculate the binding energy. The molecular mechanical energy was calculated with *sander* without any cutoff, using the continuum solvent method. The PB (Poisson–Boltzmann) calculation was done with *pbsa* [64] with an internal dielectric constant of 1 and an external dielectric constant of 80, along with a grid spacing of 0.33 Å. Nonpolar solvation energies were calculated with the equation $\Delta G_{\text{nonpolar}} = \gamma A + \beta$. Here, γ and β are constants and were derived experimentally [65] as $\gamma = 0.00542$ kcal mol^{-1} Å^{-2} and $\beta = 0.92$ kcal mol^{-1}, and the solvent-accessible surface area A was estimated using *molsurf* with a probe radius of 1.4 Å. The solute entropic contribution $T\Delta S$ was estimated using nmode, although the entropy is reported to be overestimated in MM-PBSA [31, 35].

Results and discussion

The 10 ns trajectories of all three complexes (at6a, at6b, and at4) were stable, as indicated by the stabilization of the rms deviations (Fig. 3), and the curcumin was bound to the duplex until the end of the simulations. The rms deviation of the at6a structure stabilized at an average value of ~2.4 Å, and that of the curcumin (~1.2 Å) was low, as expected. The rms deviation of at6b increased to ~3 Å after equilibration but decreased to ~2.4 Å after 5.6 ns, while the rms deviation of the curcumin molecule was ~2 Å. The high rms deviation of the curcumin is due to the change in its conformation caused by the variations in the dihedral angle about the C9–C10 bond in diketo form. In the at4 system, the rms deviation stabilized at an average of 3.3 Å, and that of curcumin was ~0.9 Å, which is the lowest in all three simulations. The rms deviation of the central AT region alone stabilized at ~1.4 Å. The additional simulations using parmbsc0 force-field parameters are also stable, and the rms deviation stabilized at ~3 Å (see Fig. 1 of the "Electronic supplementary information"). The potential and kinetic energies of the trajectories also stabilized after the equilibration run. In the at6a system, all of the base-pair hydrogen bonds in the duplex were retained throughout the dynamics run with high occupancies.

Principal component analyses of all of the systems were performed to identify the predominant motions. We calculated the first three eigenvectors of the mass-weighted covariant matrix with *ptraj*, and the visualization of the results indicated that the most important motion, represented by the first principal component, was that of the curcumin in all three systems. The movement of the curcumin molecule in the minor groove of the duplex is also visible in the trajectories. In at6a, curcumin moved along the minor groove of the duplex and shifted its position by about two base-pair steps. The ligand molecule has approximately the length of four base-pair steps, and it was initially resting between the base pairs $A_3{\cdot}T_{14}$ and $T_6{\cdot}A_{11}$. At 4.2 ns, it approached the terminal base pair $G_1{\cdot}C_{16}$ and again moved back to the initial position at ~7.8 ns; after that, it resided in that position. In the at6a system, the ligand was oriented such that one of the methyl groups (C20) was directed into the groove, and the other (C21) at the opposite end of the ligand was directed away from the groove. The second and third principal components account

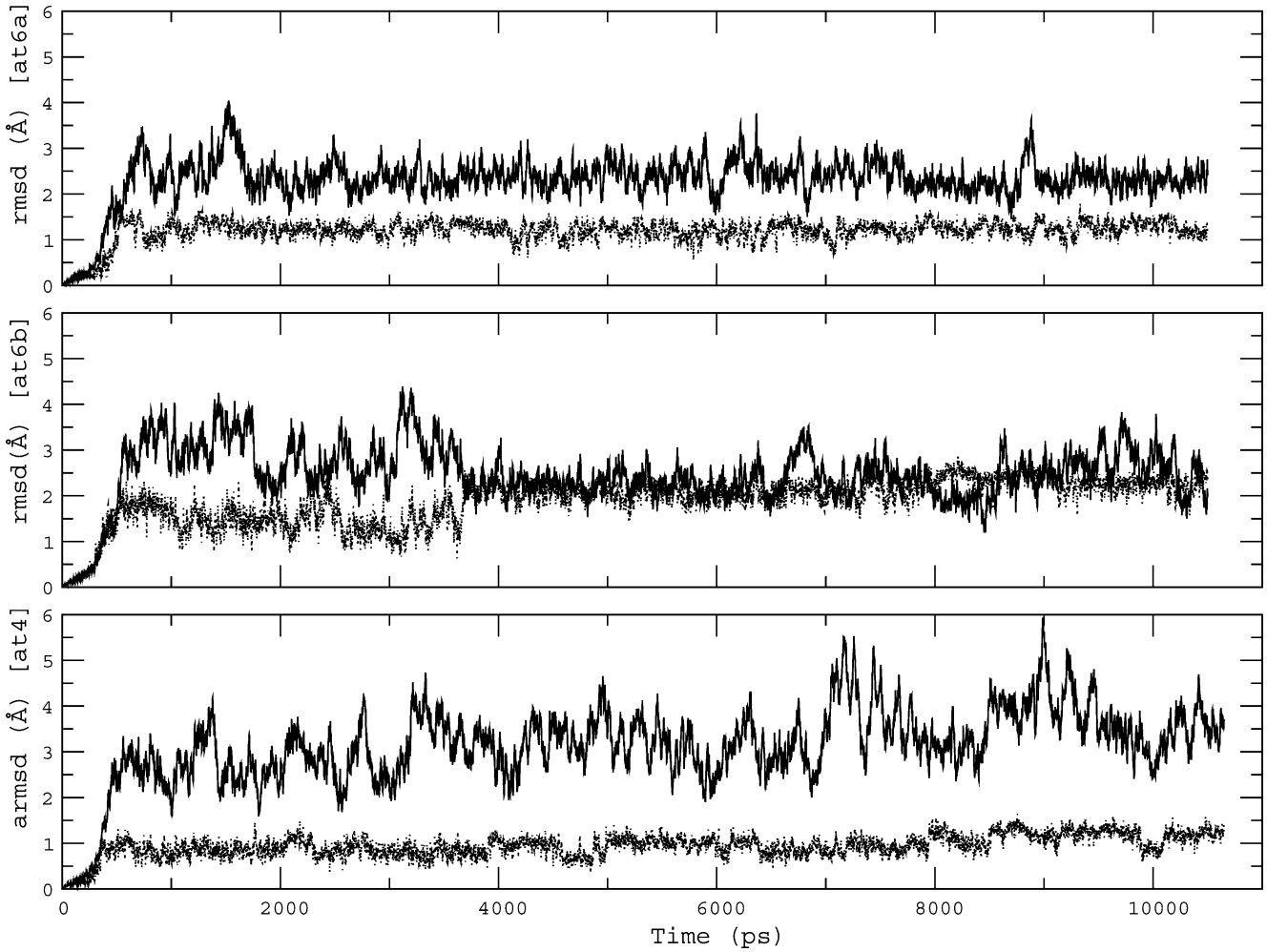

Fig. 3 Rms deviations from the starting structures of the DNA (*black*) and curcumin (*gray*) in different trajectories

for the variations in the groove widths and the fluctuations in both terminal base pairs. The minor groove width of the DNA duplex in all three systems was changed during the simulation to keep the curcumin bound in the groove. This change in groove width could be important to the binding of the ligand in the groove, as reported earlier [66, 67]. The variation of the groove width independent of the base motions comprises the second principal component. The deformability of the backbone, which is a soft degree of freedom, is important for the variations in the groove width. The starting duplex taken from the PDB (ID:1dne) has a wider minor groove (~15.4 Å) due to the side-by-side binding of two distamycin molecules, and this groove width is reduced to ~11.3 Å during the simulation at ~1.7 ns. Variations of minor groove width were found in other MD simulations of DNA minor groove complexes with an analog of H43254 [68], and with bis-benzimidazoles [69].

In the at6b system, the ligand was initially between $A_3{\cdot}T_{14}$ and $T_6{\cdot}A_{11}$, and then it moved towards the 3′ end of the first strand at 4 ns and lodged itself between the base pairs $T_4{\cdot}A_{13}$ and $A_7{\cdot}T_{10}$. In this system, the curcumin is in diketo form with single bond character for C9–C10, and hence its dihedral angles cur7 (C8-C9-C10-C11) and cur8 (C9-C10-C11-C12) are subjected to variations in comparison with the keto-enol form, where C9–C10 is a double bond. The dihedral angle cur7 was initially at 150° and changed to 86° after the equilibration. It shifted to 200° at ~3.6 ns and again to 180° at ~8 ns. Then at ~9.1 ns, cur7 changed to 220°. Similarly, cur8 was initially at 110°, showed high fluctuations between 100° and 150°, and at ~3.6 ns changed to 200°. Again, at ~8 ns it shifted to 250°, and at ~9.1 ns it changed to 210°. These variations in the dihedral angles resulted in variations in the alignment of the curcumin in the groove and also in the orientations of the phenyl groups at its ends. Both the CH_3 groups attached to the phenyl rings (C20 and C21) were directed outward from the groove. In at6b, the first principal component represents the movement of the curcumin, and the second one the

backbone movements in the ligand-bound region. The third component accounts for the movement of the terminal regions. In at6b, the minor groove width showed similar variations to those for at6a; it was reduced from ~14 Å to ~11.3 Å, but again increased to ~13.7 Å towards the end of the simulation due to the change in conformation and alignment of the ligand, especially the orientation of the phenyl groups at the ends. In this system, 12 base-pair hydrogen bonds were observed with occupancies of >95%, and two were found with 41% occupancy. The remaining four, in the ligand-bound region, showed occupancies of less than 10%.

In at4, the curcumin molecule was bound strongly to the duplex throughout the trajectory. It was closer to the backbone of the first strand but showed movement along the groove, as in the other two systems. The curcumin was attached to the duplex with one CH_3 (C20) in the groove, and the other (C21) directed outward from the groove, as in at6a. In PCA, the first component represents the motion of the curcumin along with the movements of both terminal base pairs. The three base pairs at the 5′ end of the first strand of the duplex showed much movement, which constitutes the second and third components. Initially the ligand was bound between the base pairs $A_7{\cdot}T_{18}$ and $G_{10}{\cdot}C_{15}$, aligned nearer to the first strand. At 6.4 ns it moved to the $T_6{\cdot}A_{19}$–$C_9{\cdot}G_{16}$ region, and the alignment of the curcumin molecule changed; it moved to the middle of the minor groove. The rms deviation between these two ligand conformations is 1.4 Å, and this deviation is mainly due to the movements of the two CH_3 and OH groups at both ends. In the at4 system, 32 base-pair hydrogen bonds were expected and 15 were observed with occupancies of greater than 85%, three with occupancies greater than 36%, and the remaining base-pair hydrogen bonds showed occupancies of less than 10%. As in the other two systems, low occupancies for base-pair hydrogen bonds were observed for the base pairs at the binding site and for the terminal base pairs.

Free energy calculations

The structural stability of the complexes confirms that the binding of curcumin in the minor groove is firm, and this view is supported by the MM-PBSA free energy analysis. The values of the calculated binding affinities are tabulated in Table 1. The van der Waals component is almost the same in every system, and both the van der Waals and the electrostatic contributions from the molecular mechanical component (E_{MM}) are favorable for binding. The nonpolar component of the solvation free energy is favorable, but the polar component is highly unfavorable, making the total solvation free energy unfavorable. The total binding energy calculated for at6a is −13.87 kcal mol^{-1}, that for at6b is −19.17 kcal mol^{-1}, and that for at4 is −14.28 kcal mol^{-1}, without considering the entropic contributions. In at6b, both of the terminal CH_3 groups are pointing out of the DNA minor groove, which allows for better van der Waals contacts with the duplex. This is reflected in the binding energy of at6b, which is favorable by 5.30 kcal mol^{-1} and 4.89 kcal mol^{-1} when compared to the at6a and at4 systems, respectively.

The entropic contributions of the solute for at6a and at6b are 19.0 and 22.72 kcal mol^{-1}, respectively, favoring at6a by 3.72 kcal mol^{-1}. In at4, the entropic contribution is 19.63 kcal mol^{-1}. The translational and rotational components are unfavorable in all three cases, and the vibrational component is slightly favorable. The loss of translational and rotational degrees of freedom associated with the binding is responsible for the unfavorable entropic contribution. The estimated losses of the translational and rotational components are 13.01 and 11.11 kcal mol^{-1}, respectively, for at6a. The corresponding values for at6b are 13.03 and 11.15 kcal mol^{-1}. These values were calculated based on ideal gas values with the rigid rotator harmonic approximation, and the values obtained from these assumptions are reported to be high in the free energy analysis of the complex of 4′,6-diamidino-2-phenylindole (DAPI) with

Table 1 Binding energies (kcal mol^{-1}) obtained from the MM-PBSA calculations of at6a, at6b, and at4. **a** *ELE* electrostatic, **b** *vdW* van der Waals, **c** *MM* molecular mechanical, **d** *POL* polar solvation, **e** *NON-POL* nonpolar solvation, **f** total solvation, **g** ΔG Poisson–Boltzmann free-energy, and **h** $-T\Delta S$ entropy term. *at6a-bsc*, *at6b-bsc* and *at4-bsc* are the values obtained for at6a. at6b, and at4 with the parm99bsc force-field parameters for DNA. *at6-dista* is at6 with distamycin in the minor groove

Complex	ΔE_{ELE} a	ΔE_{vdW} b	ΔE_{MM} c=a+b	ΔG_{POL} d	$\Delta G_{NON\text{-}POL}$ e	ΔG_{SOL} f=d+e	ΔG g=c+f	$-T\Delta S$ h
at6a	−19.11	−38.11	−57.22	48.15	−4.80	43.35	−13.87	19.00
at6b	−14.73	−37.74	−52.47	37.79	−4.49	33.30	−19.17	22.72
at4	−19.02	−36.18	−55.20	45.58	−4.66	40.92	−14.28	19.63
at6a-bsc	−20.79	−31.82	−52.61	43.39	−4.27	39.12	−13.49	17.54
at6b-bsc	−20.81	−41.11	−61.92	46.35	−4.72	41.63	−20.29	22.19
at4-bsc	−22.70	−32.08	−54.78	41.94	−4.31	37.63	−17.15	24.54
at6-dista	−456.81	−62.96	−519.77	481.06	−6.15	474.91	−44.86	25.06

DNA duplexes by Spackova et al. [35] using the MMPBSA method. The values reported are comparable to those obtained in the present work. In this case (DAPI complex), the binding energies reported are in the range of −18 to −22 kcal mol^{-1}, and the entropic contributions are in the range of 20–30 kcal mol^{-1}. The calculated relative binding energies, neglecting the entropic components, are comparable with experimental values.

The free energy analysis was also done with the coordinates extracted from the simulations with parm99bsc parameters for the DNA. The values obtained for at6a-bsc, at6b-bsc, and at4-bsc are −13.49, −20.29, and −17.15 kcal mol^{-1}, respectively. Different components of the binding energy are shown in Table 1.

In the at6a–distamycin complex simulation, the binding energy obtained is −44.86 kcal mol^{-1}, in comparison with approximately −13 to −20 kcal mol^{-1} for the curcumin systems (Table 1). The highly favorable electrostatic component is responsible for this high negative value, and this is expected, as the distamycin molecule is positively charged. The enhanced electrostatic term is expected on account of the hydrogen bonding potential of the distamycin molecule. The van der Waals contribution is also high due to the increased length of distamycin in comparison with curcumin: both of these contribute to the binding energy in distamycin. However, curcumin also showed comparable values for the van der Waals term, and this contributes to the binding.

In all three systems in the present work, binding was found to be driven by van der Waals interactions, and this is expected because the molecule exhibits shape complementarity with the minor groove of the duplex. The van der Waals interactions were found to be the predominant factor in protein–ligand interactions [70, 71]. Shape complementarity is also an important factor in DNA–drug systems also, and shape-based screening and docking is a successful strategy followed in rational drug design [22, 23].

DNA conformation and drug-binding interactions in the minor groove

The conformations generated in the simulations were clustered using *MMTSB*. The structure with the lowest rms deviations from the centroid of each cluster was used as the representative structure when calculating the helical parameters. Table 2 shows the rms deviations (from their corresponding starting structures) of the duplexes and curcumin of the representative structures from different clusters of the three systems, and Fig. 4 shows the starting structures and the superpositions of representative structures from each cluster of all three systems. Table 3 shows the structural parameters, and Table 4 shows groove widths. Hydrogen-bonded contacts between the curcumin and the duplex of the representative structures from different clusters of all three systems are shown in Table 2 of the "Electronic supplementary material."

In at6a, the conformations were grouped into three clusters. In the most crowded cluster of at6a, cluster 1, the ligand molecule is arranged parallel to the groove with the concave surface facing the groove, and is almost equidistant from both sugar phosphate backbone chains. Table 4 shows a comparison of the groove widths of different clusters from the three systems with those reported for some crystal structures of DNA minor groove complexes with related sequences. The ligand has three hydrogen bond contacts with the duplex. The carbonyl and hydroxyl groups of the curcumin protrude out of the groove without involving themselves in any hydrogen bonding with the DNA atoms. The hydroxy substitutions in the phenyl rings are also parallel to the groove. Figure 5a shows the contacts of the ligand in the groove of the duplex in cluster 1. Cluster 2 shows similar structural parameters, with the ligand bound in the middle of the minor groove, which is narrow at the binding region. The mode of binding for cluster 3 is different from those for clusters 1 and 2; the ligand is

Table 2 Number of conformations in each cluster and the rms deviations of the representative structures from their corresponding starting structures

System	Cluster 1	Cluster 2	Cluster 3	Cluster 4
at6a				
Number	7004 (70%)	2568 (25.7%)	428 (4.3%)	
Rmsd (Å)	2.21	2.06	2.58	
Rmsd-curcumin (Å)	0.62	0.43	0.51	
at6b				
Number	195 (2.0%)	5057 (50.6%)	4725 (47.2%)	23 (0.2%)
Rmsd (Å)	2.85	2.64	2.84	3.34
Rmsd-curcumin (Å)	2.21	2.13	2.21	2.78
at4				
Number	3770 (37.7%)	4315 (43.2%)	1202 (12.0%)	713 (7.1%)
Rmsd (Å)	3.80	3.04	3.33	4.02
Rmsd-curcumin (Å)	0.42	0.41	0.34	0.33

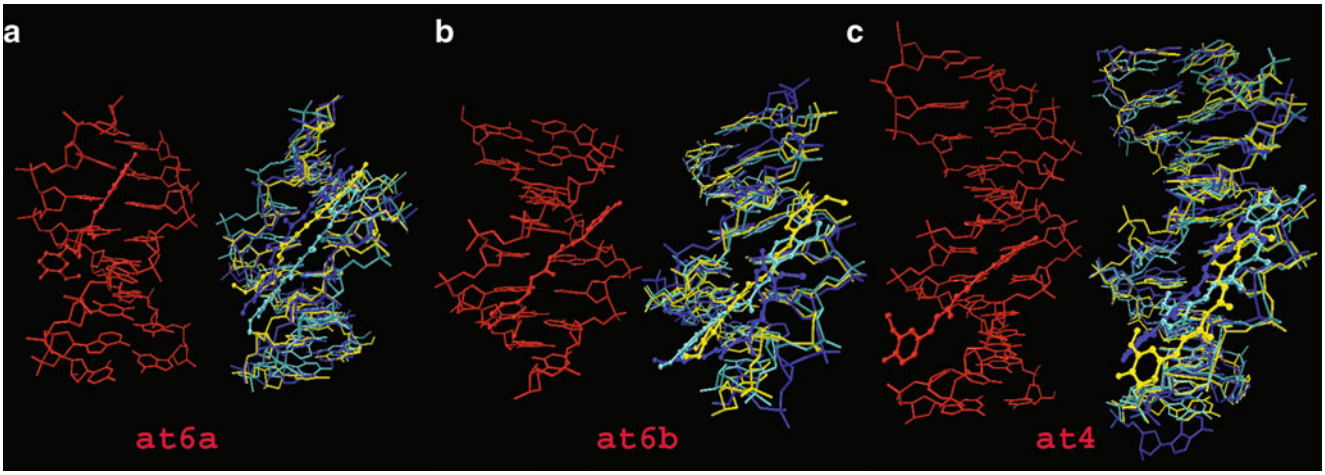

Fig. 4 Starting structures (*violet*) and superpositions of representative structures from the clusters of different systems (*blue* cluster 1, *yellow* cluster 2, and *cyan* cluster 3). Hydrogen atoms are not shown. **a** at6a: in clusters 1 and 2, the curcumin is bound at the center of the groove, while it is aligned closer to the first strand in cluster 3. **b** at6b: the curcumin in clusters 2 and 3 have similar conformations, while curcumin shows a different conformation in cluster 1, which has a population of only 2.0%. **c** at4: in clusters 1 and 3, the curcumin conformations are similar and bound at the middle of the groove, but in cluster 2 the ligand is aligned near to the first strand

aligned near the first strand, and the minor groove width is higher in this cluster (13.7 Å) at the binding site. The main common feature of clusters 1 and 2 is the reduced minor groove width at the binding site. Similar groove width patterns were observed in netropsin complexes with d $(GGCCAATTGG)_2$ [72] and d$(CGCAATTGCG)_2$ [73], and in the distamycin complex with d$(GGCCAATTGG)_2$ [74].

In the second system, at6b, the conformations were clustered into four (Table 2). Starting structures and the superpositions of representative structures from the first three clusters are shown in Fig. 4b. The rms deviations of the curcumin in various clusters are higher than that found for at6a due to the higher flexibility of the C9–C10 bond in the diketo form. In the most populated cluster, cluster 2, the curcumin was bound between the base pairs $T_4{\cdot}A_{13}$ and $A_7{\cdot}T_{10}$ in the center of the minor groove, as in at6a, and there were two hydrogen bond contacts with the duplex (Fig. 5b). The conformation of the curcumin molecule is rather similar to that in at6a. In cluster 3, the second largest cluster, the ligand occupied the same site as in cluster 2. Here again the curcumin is at the center of the groove with a similar curcumin conformation, and the groove widths showed similar behavior to those in cluster 2. In cluster 1, which contains 195 conformations, the ligand was bound between the base pairs $A_5{\cdot}T_{12}$ and $C_8{\cdot}G_9$, and its conformation was different from the other two clusters. The molecule is twisted about the central region (about the bonds C9–C10 and C10–C11), and the two phenyl groups are in different orientations. One end of the ligand is near the first strand, while the other end is near the second. The effect of this is an increased groove width in comparison with the other clusters: 13.8 Å at the binding region. From the groove widths of the different clusters in at6b, it is evident that the conformations of the ligand have an effect on the groove width. In at6b, the flexibility of the torsion angle about C9–C10 of the curcumin results in different orientations of the phenyl groups at both ends of the molecule and the attached OH and CH_3 groups. This in turn results in the widening of the minor groove at the binding site, similar to the starting crystal structure (14.7 Å), where

Table 3 Structural parameters of representative structures of different systems

System	Rise (Å)	Twist (°)	Roll (°)	Propeller (°)	Opening (°)	Number of bp/turn
At6a-cluster1	3.1	35.2	3.7	−11.0	1.7	10.4
At6a-cluster2	3.4	33.4	4.7	−9.7	1.1	10.8
At6a-cluster3	3.3	31.6	3.7	−13.4	4.2	11.4
At6b cluster1	3.3	30.8	6.3	−8.8	3.2	11.7
At6b-cluster2	3.3	33.8	3.1	−8.8	2.7	10.6
At6b-cluster3	3.4	34.1	2.2	−11.7	2.7	10.6
At4-cluster1	3.3	32.3	4.7	−15.4	2.8	11.2
At4-cluster2	3.4	32.0	5.2	−14.6	−1.9	11.2
At4-cluster3	3.3	31.6	1.6	−11.2	−1.4	11.4

Table 4 Groove widths of representative structures from different clusters in the ligand-bound region and the region away from the ligand. Groove widths of some crystal structures of minor groove complexes of DNA duplexes are also shown. The width is reduced by ~2 Å in the binding region

	Minor groove width (Å)	
System	Near the ligand	Away from the ligand
At6a		
Cluster 1	11.3	13.2
Cluster 2	11.2	12.8
Cluster 3	13.7	
At6b		
Cluster 1	14.8	12.9
Cluster 2	11.6	
Cluster 3	11.8	11.0
Cluster 4	13.8	13.4
At4		
Cluster 1	11.7	13.3
Cluster 2	11.2	14.1
Cluster 3	11.8	14.2
Cluster 4	11.3	13.8
Crystal structures(PDB ID and ligand)		
1dne (netropsin)	11.6	13.9
1z8v (netropsin)	11.2	13.4
261d (netropsin)	11.4	13.8
1jtl (distamycin)	11.3	13.9
1k2z (distamycin)	11.2	12.7

two distamycin molecules are bound side-by-side in the minor groove [41].

In the at4 system, four clusters were obtained (Table 2). Figure 4c shows the superposition of representative structures of different clusters and the starting structure. In cluster 2, the most crowded one, the curcumin is between $A_7{\cdot}T_{18}$ and $G_{10}{\cdot}C_{15}$, and is aligned closer to the backbone of the first strand, with one end of the ligand protruding out of the groove. There are two hydrogen-bond contacts between the ligand and the DNA in this cluster (Fig. 5c). In the second largest cluster, cluster 1, the ligand is bound between the base pairs $A_5{\cdot}T_{20}$ and $T_8{\cdot}A_{17}$, and there are three hydrogen bond contacts with the duplex. In cluster 3, the ligand is between the base pairs $T_6{\cdot}A_{19}$ and $C_9{\cdot}G_{16}$. In all of the systems, hydrogen bonds were observed between the duplex and the ligand, and the minor groove is reduced at the binding region.

Hydration and water-mediated binding interactions

It is interesting to note that water molecules mediate the binding of the curcumin in all three systems. The snapshot from each trajectory corresponding to the representative structure of the largest cluster in each system was extracted to analyze the water-mediated contacts between the ligand and the duplex. In Fig. 6, the water molecules interacting with the ligand and the duplex are shown in the representative snapshots of at6a, at6b, and at4. The water-mediated interactions between the curcumin and the duplex in different systems are shown in Table 3 of the "Electronic supplementary material." Water molecules are reported to play a crucial role in DNA minor groove recognition by small molecules [75], and similar water-mediated contacts

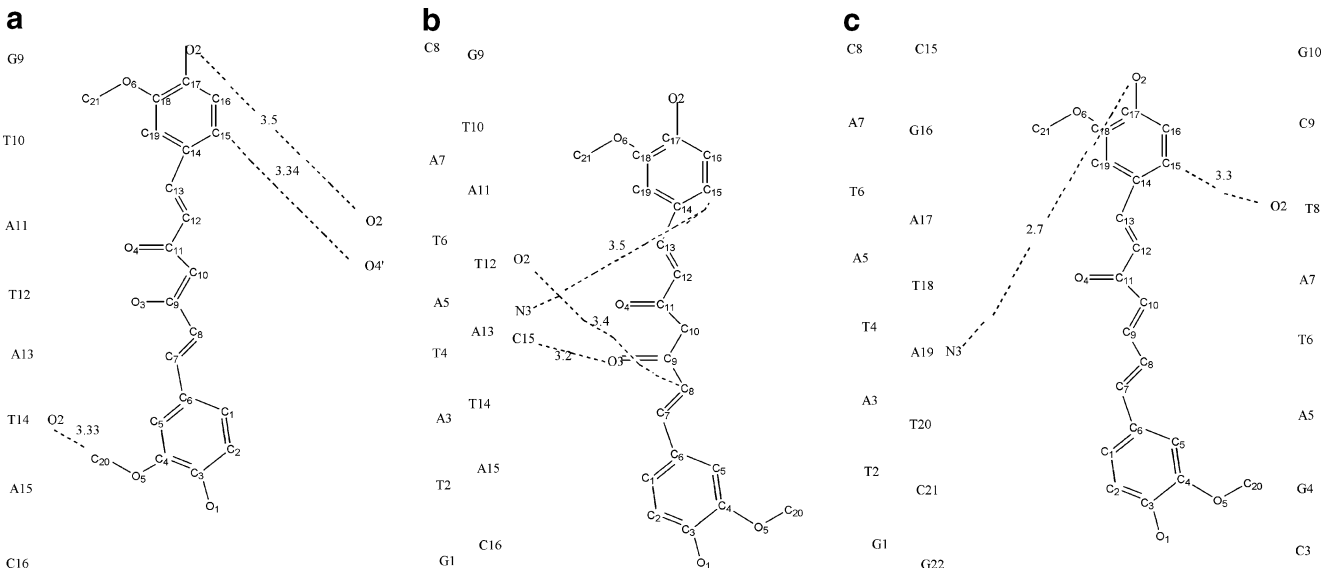

Fig. 5 Hydrogen bonds between the curcumin and DNA duplex in the most populated cluster in each system: **a** at6a, **b** at6b, and **c** at4

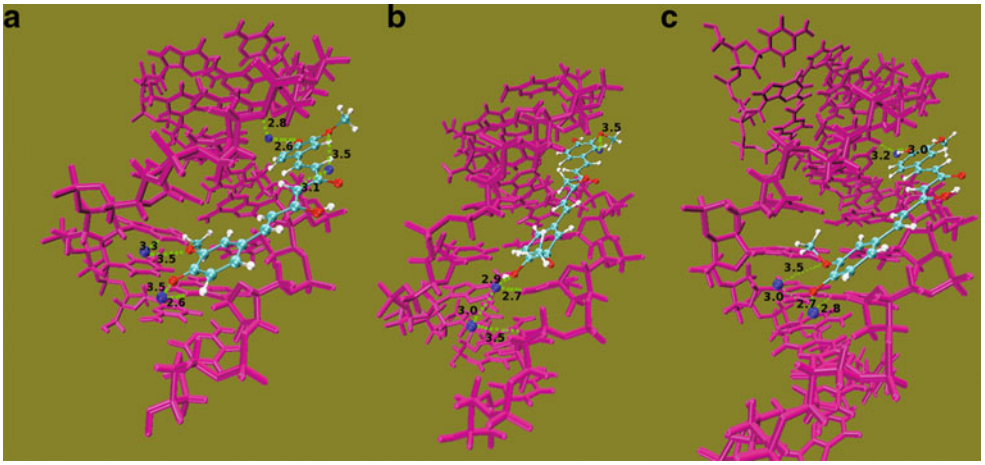

Fig. 6 Water-mediated hydrogen bonds between the curcumin molecule and the duplexes in **a** at6a, **b** at6b, and **c** at4

were seen in MD simulations of DNA minor groove complexes with DAPI [35].

Duplex hydration was examined by performing a hydrogen-bonding analysis with water, as well as with the grid option in *ptraj*. All three systems showed similar hydration patterns. As expected, the phosphates and the minor grooves of the duplexes were hydrated except for the ligand-bound regions. In at6a, on average, ~2.5 water molecules were attached to the phosphates throughout the simulation. Two water molecules each were found near the curcumin molecule at both ends in the minor groove of the duplex, and they showed long residence times. Similarly, one water molecule was present with a long residence time in the minor groove, in-between the duplex and the ligand, and was hydrogen bonded to the T_{14}–O2 with 91.78% occupancy. This water molecule was near ligand atom C5. Another water molecule was attached to T_{12}–O2 with an occupancy of ~100%, which is near the ligand atom O1. The major groove side shows a similar hydration pattern to normal B-DNA, with long residence times. In at6b, approximately two water molecules were attached to all of the phosphates with high residence times. In at4 too, all of the phosphates were hydrated with ~2.6 water molecules throughout the simulation, and there was a spine of hydration in the minor groove, with variations in the region of the ligand. The major groove side also had some hydration sites. Figure 7 shows the hydration patterns of at6a, at6b, and at4, obtained from *ptraj* grid analysis (contouring level is three times the bulk water density). The duplex is hydrated in a similar fashion to a normal DNA duplex, with deviations at the binding region in all three systems. Changes in the hydration pattern of the DNA duplex are seen in other minor groove complexes as well [76].

Modification of helical parameters on account of curcumin binding

It is known that in the minor groove complexes of DNA, the helical parameters usually show deviations from their canonical DNA forms [3]. In order to accommodate the ligand in the groove, bases as well as the backbone deviate from their usual conformations, resulting in varied structural parameters, especially for the bases in the binding region and the flanking bases. In our systems, the helical parameters of the structures also showed minor deviations from the standard values, and the results for the representative structures are shown in Fig. 2 of the "Electronic

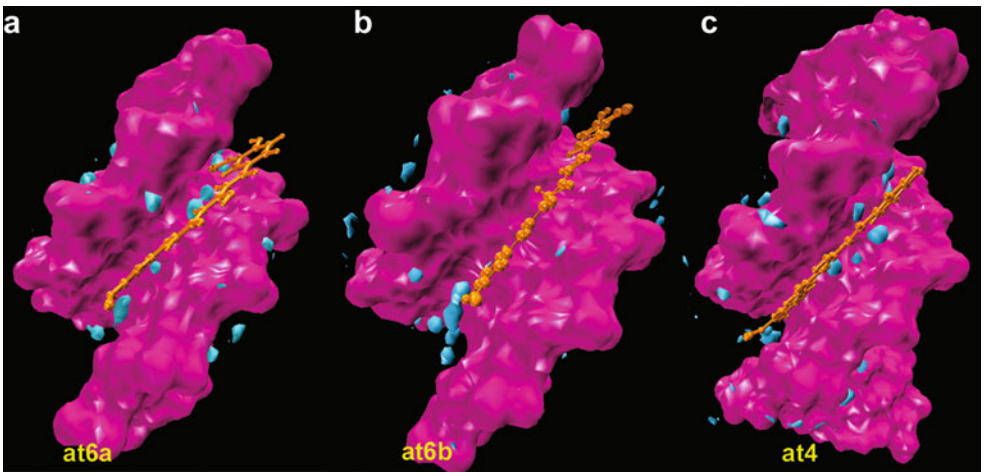

Fig. 7 Hydration patterns of the duplexes. Water is contoured at ~3 times the bulk water density

supplementary information." Even though the average twist is close to the standard B-form geometry in at6a, a large twist is observed in the base-pair step next to the ligand site, the $T_6{\cdot}A_{11}$-$A_7{\cdot}T_{10}$ step, in all three clusters (49.7°, 40.5°, and 36.1°, respectively). Similarly large positive roll values were seen in the $T2{\cdot}A_{15}$- $A_3{\cdot}T_{14}$ step in all three clusters near the other end of the ligand, and overall the roll angles showed variations. A large negative propeller twist is seen in the central region of the duplex where the ligand is bound. The at6b duplex also showed similar trends for all helical parameters. A large twist at the base-pair step next to the binding region and a negative propeller twist at the binding site are seen in all of the clusters. Roll values showed variations, with the maximum positive value seen for the central base pair. This is in contrast to the negative roll in the central region seen in the starting crystal structure [41]. In at4, the variation in twist is higher than in other systems. High twist is observed for the base-pair steps flanking the ligand-bound region, as in the starting crystal structure. Roll values also showed variations and a large negative propeller twist is seen in the central region in all clusters. A large twist in the region next to the bound ligand has been reported for the starting structures as well as other minor groove complexes [77].

Generally the backbone torsion angles were in standard conformations. The α-γ torsions were in the preferred −sc/+sc conformations in all of the clusters for all of the systems except T2 of at6a, cluster1, for which it was +sc/+ap. The glycosidic torsion χ was in the −ac region for all of the residues in all of the systems. In all three systems at6a, at6b, and at4, the sugar puckers were mainly distributed among C2′ endo and C1′ exo, both of which belong to the B-form family. O4′ endo sugar puckers, which are the lowest-energy intermediates between the A and B forms, were seen for some residues. This was also reported for the starting crystal structures. Overall, the binding of curcumin causes variations in groove width, high twist to the base pairs adjacent to the region of binding, and high negative propeller twist in the binding region.

Conclusions

The shape complementarity between the B-form DNA minor groove and the curcumin molecule was the reason for performing this work, in which we docked the curcumin molecule to AT-rich DNA duplexes. MD simulations followed by free energy analysis of the complexes were performed to assess the binding affinity, and our results show that curcumin binds in the minor grooves of AT-rich sequences of DNA, like the minor groove binding drugs netropsin and distamycin. The binding interactions of curcumin are mainly favored by van der Waals forces, as expected on account of its electrical neutrality, unlike the electrostatic interactions that favor the binding of netropsin and distamycin (both of which are charged molecules).

The starting B-form geometry of the DNA duplex is retained in all three systems investigated throughout the simulations, as indicated by the average helical parameters and the sugar pucker. However, the helical parameters show variations in the ligand-bound region. High twist, observed at the base-pair steps next to the binding site, could be due to the reduced groove width at the binding region. The groove widths are reduced by about 2 Å in the binding region in all of the systems and show higher values in other regions (Table 4); this variation in groove width is one of the most important factors in the stabilization of the complex. Reduced groove widths are seen in the simulations, and are found in crystal structures of minor groove complexes and also in other MD simulations of similar systems. The binding of the curcumin was stable in all three systems, and the interactions are mainly van der Waals and weak hydrogen bonding, such as C-H...N and C-H...O. Water-mediated hydrogen bonds between the curcumin molecule and the DNA atoms, seen in all three complex systems, are also found to contribute to the stable binding. The phosphates, as well as the minor groove, are hydrated in all three systems. The spine of hydration in the minor groove is slightly modified due to the ligand binding. The non-covalent binding of the curcumin in the minor grooves of the duplexes is stable, and the binding free energies obtained are favorable for binding. In support of our findings, recent spectroscopic investigations on curcumin complexes with calf thymus DNA and yeast RNA systems by Nafisi et al. [24] demonstrate the binding of curcumin, even though the sequences we studied are different. The results indicate that curcumin could be an appropriate choice to develop minor groove binding natural drugs.

Acknowledgments The simulations were carried out at the Supercomputing Facility for Bioinformatics and Structural Biology at the Indian Institute of Technology Delhi, and at the University Computer Center, Mahatma Gandhi University, Kottayam. The authors are thankful to the University Grants Commission, Government of India for financial assistance through the Special Assistance Programme—Departmental Research Support Programme.

References

1. Neidle S, Thurston DE (2005) Nat Rev Cancer 5:285–296
2. Nelson SM, Ferguson LR, Denny WA (2007) Mutat Res 623:24–40
3. Geierstanger BH, Wemmer DE (1995) Annu Rev Biophys Biomol Struct 24:463–493
4. Neidle S (2001) Nat Prod Rep 18:291–309
5. Ren J, Chaires JB (1999) Biochemistry 38:16067–16075

6. Dolenc JZ, Oostenbrink C, Koller JZ, van Gunsteren WF (2005) Nucleic Acids Res 33:725–733
7. Zimmer C, Wahnert U (1986) Prog Biophys Mol Biol 47:31–112
8. Aggarwal BB, Kumar A, Bharti AC (2003) Anticancer Res 23:363–398
9. Brouet I, Ohshima H (1995) Biochem Biophys Res Commun 206:533–540
10. Rajakumar DV, Rao MN (1994) Mol Cell Biochem 140:73–79
11. Anand P, Sundaram C, Jhurani S, Kunnumakkara AB, Aggarwal BB (2008) Cancer Lett 267:133–164
12. Arbiser JL, Klauber N, Rohan R, Van Leeuwen R, Huang MT, Fisher C, Flynn E, Byers HR (1998) Mol Med 4:376–383
13. Firozi PF, Aboobaker VS, Bhattacharya RK (1996) Chem Biol Interact 100:41–51
14. Duvoix A, Blasius R, Delhalle S, Schnekenburger M, Morceau F, Henry E, Dicato M, Diederich M (2005) Cancer Lett 223:181–190
15. Aggarwal BB, Harikumar KB (2009) Int J Biochem Cell Biol 41:40–59
16. Aggarwal BB, Sung B (2009) Trends Pharmacol Sci 30:85–94
17. Kopka ML, Yoon C, Goodsell D, Pjura P, Dickerson RE (1985) J Mol Biol 183:553–563
18. Coll M, Frederick CA, Wang AHJ, Rich A (1987) Proc Nat Acad Sci USA 84:8385–8389
19. Ahsan H, Parveen N, Khan NU, Hadi SM (1999) Chem Biol Interact 121:161–175
20. Ahsan H, Hadi SM (1998) Cancer Lett 124:23–30
21. Zsila F, Bikadi Z, Simonyi M (2004) Org Biomol Chem 2:2902–2910
22. Spitzer GM, Wellenzohn B, Markt P, Kirchmair J, Langer T, Liedl KR (2009) J Chem Inf Model 49:1063–1069
23. Kirchmair J, Distinto S, Markt P, Schuster D, Spitzer GM, Liedl KR, Wolber G (2009) J Chem Inf Model 49:678–692
24. Nafisi S, Adelzadeh M, Norouzi Z, Sarbolouki MN (2009) DNA Cell Biol 28:201–208
25. Zwanzig RW (1954) J Chem Phys 22:1420–1426
26. Kirkwood JG (1935) J Chem Phys 3:300–313
27. Srinivasan J, Cheatham TE, Cieplak P, Kollman PA, Case DA (1998) J Am Chem Soc 120:9401–9409
28. Massova I, Kollman PA (1999) J Am Chem Soc 121:8133–8143
29. Wang W, Kollman PA (2000) J Mol Biol 303:567–582
30. Obiol-Pardo C, Rubio-Martinez J (2007) J Chem Inf Model 47:134–142
31. Singh N, Warshel A (2010) Proteins 78:1705–1723
32. Huo S, Wang J, Cieplak P, Kollman PA, Kuntz ID (2002) J Med Chem 45:1412–1419
33. Stoica I, Sadiq SK, Coveney PV (2008) J Am Chem Soc 130:2639–2648
34. Combelles C, Gracy J, Heitz A, Craik DJ, Chiche L (2008) Proteins 73:87–103
35. Spackova N, Cheatham TE III, Ryjacek F, Lankas F, van Meervelt L, Hobza P, Sponer J (2003) J Am Chem Soc 125:1759–1769
36. Kopka ML, Yoon C, Goodsell D, Pjura P, Dickerson RE (1985) Proc Nat Acad Sci USA 82:1376–1380
37. Patel DJ (1982) Proc Nat Acad Sci USA 79:6424–6428
38. Khopde S, Priyadarsini KI, Palit DK, Mukherjee T (2000) Photochem Photobiol 72:625–631
39. Mague JT, Alworth WL, Payton FL (2004) Acta Crystallogr C 60: o608–o610
40. Ishigami Y, Goto M, Masuda T, Takizawa Y, Suzuki S (1999) J Jpn Soc Colour Mater 72:71–77
41. Mitra SN, Wahl MC, Sundaralingam M (1999) Acta Crystallogr D 55:602–609
42. Coll M, Aymami J, Van der Marel GA, Van Boom JH, Rich A, Wang AH (1989) Biochemistry 28:310–320
43. Lang PT, Brozell SR, Mukherjee S, Pettersen EF, Meng EC, Thomas V, Rizzo RC, Case DA, James TL, Kuntz ID (2009) RNA 15:1219–1230
44. Wang J, Morin P, Wang W, Kollman PA (2001) J Am Chem Soc 123:5221–5230
45. Case DA, Darden TA, Cheatham TE III, Simmerling CL, Wang J, Duke RE, Luo R, Merz KM, Wang B, Pearlman DA, Crowley M, Brozel S, Tsui V, Gohlke J, Mongan J, Hornack V, Cui G, Beroza P, Schafmeister C, Caldwell JW, Ross WA, Kollman PA (2004) AMBER8. University of California, San Francisco
46. Wang J, Wang W, Kollman PA, Case DA (2006) J Mol Graph Model 25:247–260
47. Wang J, Wolf RM, Caldwell JW, Kollamn PA, Case DA (2004) J Comput Chem 25:1157–1174
48. Wang J, Cieplak P, Kollman PA (2000) J Comput Chem 21:1049–1074
49. Perez A, Marchan I, Svozil D, Sponer J, Cheatham TE, Laughton CA, Orozco M (2007) Biophys J 92:3817–3829
50. Jakalian A, Jack DB, Bayly CI (2002) J Comput Chem 23:1623–1641
51. Kuhn B, Gerber P, Schulz-Gasch T, Stahl M (2005) J Med Chem 48:4040–4048
52. Anconi CPA, Nascimento CS, De Almeida WB, Dos Santos HF (2009) J Phys Chem B 113:9762–9769
53. Mobley DL, Bayly CI, Cooper MD, Dill KA (2009) J Phys Chem B 113:4533–4537
54. Taliani S, Cosimelli B, Da Settimo F, Marini AM, Motta CL, Simorini F, Salerno S, Novellino E, Greco G, Cosconati S, Marinelli L, Salvetti F, Abbate GL, Trasciatti S, Montali M, Costa B, Martini C (2009) J Med Chem 52:3723–3734
55. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) J Chem Phys 79:926–935
56. Darden TA, Perera L, Li L, Penderson L (1999) Structure 7:R55–R60
57. Darden TA, York D, Penderson L (1993) J Chem Phys 98:10089–10092
58. Varghese MK, Thomas R, Unnikrishnan NV, Sudarsanakumar C (2009) Biopolymers 91:351–360
59. Humphrey W, Dalke A, Schulten K (1996) J Mol Graph 14:33–38
60. Mongan J (2004) J Comput Aided Mol Des 18:433–436
61. Feig M, Karanicolas J, Brooks CL III (2004) J Mol Graph Model 22:377–395
62. Lu X-J, Olsen WK (2003) Nucleic Acids Res 31:5108–5121
63. Kollman PA, Massova I, Reyes C, Kuhn B (2000) S Huo, Chong L, Lee M, Lee T-S, Duan Y, Wang W, Donini O, Cieplak P, Srinivasan J, Case D, Cheatham TE, III. Acc Chem Res 33:889–897
64. Luo R, David L, Gilson MK (2002) J Comput Chem 23:1244–1253
65. Sitkoff D, Sharp KA, Honig B (1994) J Phys Chem 98:1978–1988
66. Laughton C, Luisi B (1998) J Mol Biol 288:953–963
67. Wood AA, Nunn CM, Czarny A, Boykin DW, Neidle S (1995) Nucleic Acids Res 23:3678–3684
68. Bostock-Smith CE, Harris SA, Laughton CA, Searle MS (2001) Nucleic Acids Res 29:693–702
69. Bailly C, Chessari G, Carrasco C, Joubert A, Mann J, Wilson WD, Neidle S (2003) Nucleic Acids Res 31:1514–1524
70. Barratt E, Bingham RJ, Warner DJ, Laughton CA, Phillips SEV, Homans SW (2005) J Am Chem Soc 127:11827–11834
71. Kasper C, Lunn M-L, Liljefors T, Gouaux E, Egebjerg J, Kastrup JS (2002) FEBS Lett 531:173–178
72. Van Hecke K, Nam PC, Nguyen MT, van Meervelt L (2005) FEBS J 272:3531–3541
73. Nunn CM, Garman E, Neidle S (1997) Biochemistry 36:4792–4799
74. Uytterhoeven K, Sponer J, van Meervelt L (2002) Eur J Biochem 269:2868–2877
75. Nguyen B, Neidle S, Wilson WD (2009) Acc Chem Res 42:11–21
76. Han F, Taulier N, Chalikian TV (2005) Biochemistry 44:9785–9794
77. Abrescia NGA, Malinina L, Subirana JA (1999) J Mol Biol 294:657–666