J Mol Model (2014) 20:2499
DOI 10.1007/s00894-014-2499-7

ORIGINAL PAPER

Contribution of phenylalanine side chain intercalation to the TATA-box binding protein–DNA interaction: molecular dynamics and dispersion-corrected density functional theory studies

Manas Mondal · Sanchita Mukherjee · Dhananjay Bhattacharyya

Received: 9 July 2014 / Accepted: 13 October 2014 / Published online: 30 October 2014

Abstract Deformation of DNA takes place quite often due to binding of small molecules or proteins with DNA. Such deformation is significant due to minor groove binding and, besides electrostatic interactions, other non-covalent interactions may also play an important role in generating such deformation. TATA-box binding protein (TBP) binds to the minor groove of DNA at the TATA box sequence, producing a large-scale deformation in DNA and initiating transcription. In order to observe the interactions of protein residues with DNA in the minor groove that produce the deformation in the DNA structure, we carried out molecular dynamics simulations of the TBP–DNA system. The results reveal consistent partial intercalation of two Phe residues, distorting stacking interactions at two dinucleotide step sites. We carried out calculations based on dispersion-corrected density functional theory to understand the source of such stabilization. We observed favorable interaction energies between the Phe residues and the base pairs with which they interact. We suggest that salt-bridge interactions between the phosphate groups and Lys or Arg residues, along with the intercalation of Phe residues between two base pair stacks, stabilize the kinked and opened-up DNA conformation.

Keywords Stacking energy · TATA-box TBP · Dispersion-corrected DFT · Base pair orientation · Partial stacking · Molecular dynamics simulation

Electronic supplementary material The online version of this article (doi:10.1007/s00894-014-2499-7) contains supplementary material, which is available to authorized users.

M. Mondal · S. Mukherjee · D. Bhattacharyya (✉)
Computational Science Division, Saha Institute of Nuclear Physics, 1/AF Bidhan Nagar, Kolkata 700064, India
e-mail: dhananjay.bhattacharyya@saha.ac.in

Introduction

Proteins and small molecules bind to DNA by making mostly non-bonded contacts along the major or minor grooves. During this mode of binding, the DNA double helical structure does not alter significantly. The protein–DNA interface is formed primarily by van der Waals contacts, hydrogen bonds, and salt bridges. Some small planer aromatic molecules and aromatic side chains of proteins also interact with DNA by making stacking interactions with the DNA bases via intercalation. With this mode of interaction, the alteration in DNA local structural is quite significant at the binding site. This strength of interaction usually correlates with the biological activity of the molecule, and several energy contributions may be responsible for the binding. Most intercalators bind to DNA by non-covalent stacking with nucleic acid base pairs, often combined with hydrogen bonding and even covalent binding involving the drug side chains [1]. The separation between successive basepairs increases considerably due to the binding of the intercalators, i.e., the rise value of the base pair step is increased and the double helix is partially untwisted. This class of DNA-binding ligands includes dyes, mutagens and carcinogens, and some anti-tumor antibiotics, which bind to DNA with varying degrees of sequence selectivity [2–8]. Such ligand binding can affect many biological properties of DNA, including transcription and replication. Depending on the intercalator, these effects are modulated not only by the chemical characteristics of the chromophore but also by the presence and position of substituent moieties attached to it. The general mode of intercalation was proposed to involve in formation of a kink, i.e., a sharp, localized bend that may provoke the appearance of a wedge-shaped entry notch accessible from the minor or major groove. Initial formation of a kink would be the result of un-stacking of a

single base pair step and rolling open of the base planes. This could be related to the alternation of the normal C2′-endo deoxy-ribose sugar puckering in B-DNA to a mixed puckering pattern. These changes do bring about DNA unwinding and changes in the magnitude of the helical twist [9]. Moreover, the helical twists at dinucleotide steps appear to be inversely related to the roll angles i.e., the amount of wedge formation. Rise value of the dinucleotide step, which is altered due to intercalation, is also strongly correlated to twist. There exists a possibility that intercalation at certain sequences is based, at least in part, on the inherent flexibility of dincleotide steps in terms of base pair parameters like roll, twist, slide, rise, propeller, etc. [10]. Kinking of the DNA by minor groove intercalation has also been observed not only upon binding of small ligands but also in complexes with proteins involved in transcriptional regulation [11].

The three-dimensional (3D) structures of several DNA minor groove binding proteins have been determined, revealing a new mechanism to produce a bend in DNA with a possible role in control and regulation of transcription. The common mechanism involves the insertion of one or several aromatic side chains into the minor groove of DNA, and unstacking of two successive base pairs to produce perceptible kinks. Due to this side chain insertion, the DNA structure is bent strictly toward the major groove, which helps to open the narrow minor groove. The 3D structures of the TATA-box binding protein (TBP) complex with the TATA box [12–15], and of the trimeric complexes between the TBP/TATA box and TFIIA [16, 17] or TFIIB [18, 19] revealed this fact. Besides several hydrogen bonding interactions between charged amino acid residues and the backbone phosphate atoms of DNA as observed in the crystal structures, van der Waals interactions also play an important role [20–22]. These non-polar contacts occur between a series of conserved Val, Leu, and Pro residues and the minor groove edges of the bases, and between the side chains of two pairs of conserved phenylalanine (Phe) residues and the first two (TA) and last two (t/aX) basepairs of the TATA box [23–25]. Insertion of these Phe residues between two base pairs produces two sharp kinks in the helix, unstacking the base-pairs and pulling them partially apart. Any loss of stacking energy is probably compensated by extensive van der Waals interaction or π–π interaction between the bases and the Phe rings. In the case of the human TBPc-DNA complex, Phe-284 and Phe-193 insert into the first and last base pair steps of the TATA box, respectively. In addition, Phe-210 and Phe-301 residues support the penetrating phenylalanines and stabilize the kinks through extensive van der Waals contacts with the ribose groups of A and T, respectively. Depending on the depth of intercalation or insertion of the Phe residues between the base pair steps, the bending kink at the 3′ end of the TATA box is different for plant and human complexes. In case of plants, Phe-193 is inserted 1.0 Å further into the base steps compared with the human complex, leading to a slightly larger kink in the 3′ end of the TATA box (44° vs 39°) [26]. The 3D structure of TBP and its interactions with DNA have remained nearly conserved throughout evolution. The shaded base pair steps in the AdMLP TATA element containing sequence in Fig. 1 are the steps in which the conserved Phe residues are intercalated [15, 27]. The selectivity of TBP for the first and most conserved TA step of the TATA box, compared with the other and less common AT, TT, AA, and CG steps, has been reported [24, 25, 28]. The strongest requirement for an efficient TATA box appears to be an initial T, with a subsequent A mandated so the Phe side-chains of TBP can push them apart to create the upstream kink.

Although structural studies have indicated conserved kinks, stabilization due to these deformations is not well known. We have used molecular dynamics (MD) simulation of the human TBP–DNA complex to study the deformation of DNA due to binding of TBP and interactions of TBP residues with DNA in the minor groove side. Recent advances in ab-initio quantum chemical methods have made them suitable for estimation of such interaction strengths. However, it was pointed out that the most efficient density functional theory (DFT) based methods are incapable of judging stacking interaction by considering the dispersion component. Several dispersion-corrected DFT (DFT-D) methods have been tested by various groups for estimation of stacking interaction [1, 29–34]. By comparing several such DFT-D methods, we recently found that ωB97X-D/6-31G(2d,2p) is most suitable for nucleic acid systems [35]. Here, we also explored the energetic basis of the stability of the intercalated Phe residues to maintain the kinks at the two ends of the TATA box. Both classical and DFT-D calculations were performed to study the energetics of interaction between the base pair steps of DNA and the hydrophobic phenylalanine side chain. It was reported that the DFT-D method yields results with a chemical precision (<1 kcal mol^{-1}) that is comparable to highly accurate CCSD/(T)/complete basis set limit calculations, and is capable of describing the hydrogen bonding and dispersion interactions known to make an important contribution to base pair stacking and ligand–nucleobase interactions [1, 35–45]. We carried out geometry optimizations of two base paired dinucleotide steps, considering the base atoms by replacing the sugar-phosphate backbone atoms with methyl group along with the intercalated Phe residue, replacing the protein main chain with an ethyl group, using the ωB97X-D/6-

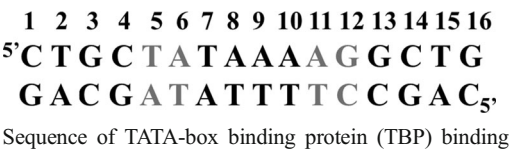

Fig. 1 Sequence of TATA-box binding protein (TBP) binding DNA sequence in h-TBP–DNA complex. *Shaded* base pair steps are phenylalanine (Phe) intercalation sides

31G(2d,2p) method. It was shown recently that results with dinucleotides including the sugar-phosphate backbone are very similar to those obtained by considering bases only [46]. Use of methyl groups mimicking sugars was also seen to give excellent results [35, 40, 47]. Both classical and DFT-D calculations indicate that the interaction between the intercalator and DNA bases is quite strongly attractive and that it has a specific role in stabilizing the kinks formed in the DNA structure.

Methods

Many crystal structures of the TBP–TATA box complex are now available; due to the extreme importance of this binding, most of them were solved at quite high resolution. We have analyzed these PDB structures [48] (PDB ID: 1AIS, 1C9B, 1CDW, 1D3U, 1JFI, 1NGM, 1NH2, 1NVP, 1QN3, 1QN4, 1QN5, 1QN6, 1QNA, 1QNB, 1QNE, 1RM1, 1TGH, 1VOL, 1VTL, 1VTO, 1YTB, 1YTF). Among these, 1CDW is for the human TBP–DNA complex.

MD simulation and structural analysis

The coordinates of the TBP-bound DNA were taken from Protein Data Bank (PDB ID 1CDW [15]). The protein in 1CDW.pdb corresponds to the C-terminal domain but for simplicity we numbered the 155th residue as residue 1. Initial set up and minimizations of the system were performed using the AMBER 8 suite of programs [49] with parm94 [50] parameters, which are well tested for nucleic acids as well as proteins [51, 52]. The system for simulation with periodic boundary condition was solvated in an orthorhombic water box containing TIP3P water molecules in such a way that there were at least 15 Å thick layers of water molecules around the solute in all directions and the required number of sodium ions to maintain electro-neutrality. The system was then energy minimized for 20,000 cycles using a combination of steepest descent and conjugate gradient algorithms and applying periodic boundary condition. Electrostatic energy was calculated using the particle mesh Ewald summation method with 1 Å grid spacing and 10^{-6} convergence criterion. Lennard-Jones and short-range electrostatic interactions were truncated at 10 Å. The MD simulation was carried out using NAMD [53, 54]. Heating to 300 K was carried out slowly during 30 ps with 1 fs time step using NAMD. We continued the simulation for 100 ns at constant temperature (300 K) and pressure (1 bar) using the Langevin-Piston algorithm. Translational and rotational movements of the center of mass were removed at an interval of 5 ps. SHAKE constraints were applied to all bonds involving hydrogen atoms. Conformations after every 1 ps were saved for further trajectory analysis. We also performed the simulation for 30 ns using the same protocol and the AMBER ff14SB [55] force-field, which follows parmbsc0 [56] parameters for nucleic acids to understand the force-field effect.

The interaction energy between the intercalated Phe residue and the corresponding basepair steps was evaluated throughout the simulation time scale using NAMD software [53]. Root mean square deviation (RMSD) and root mean square fluctuation (RMSF) analysis of the DNA and protein structures were performed using CHARMM [57], and all the structural parameters of the DNA were calculated by NUPARM [58]. The hydrogen bonding information between protein and DNA during MD runs of the TBP–DNA complex were obtained by a modified version of pyrHBfind [59] using donor-to-acceptor distance cut-off at 3.0 Å and angle greater than 150°. We used donor-to-acceptor distance cutoff 3.5 Å and hydrogen bond angle cutoff 120° for calculation of the frequency of inter-base-pair hydrogen bonds in the simulation trajectories.

Modeling of the stacked geometry

We took the crystal structure, as well as a few simulation snapshots of the corresponding base pair steps (TA/TA and CT/AG) and intercalating phenyl ring for quantum chemical studies. The details of the system studied are given in Table 1. Instead of the sugar-phosphate backbone we attached a methyl group with the bases as previously done in other studies (Fig. 2) [35, 60]. An ethyl group is added with the phenyl ring instead of s protein chain, which can be considered as ethylbenzene (Fig. 2). An ethyl group was chosen as it is bulkier and may hinder rotation of the phenyl ring to an undesired position and may be able to increase ring charge density as well as mimic the effect of the protein main chain.

DFT calculations

The modeled systems were geometry optimized and the interaction energy was obtained with the DFT-D [39] method, ωB97X-D [42] functional and 6-31G(2d,2p) basis set using Gaussian09 [61]. The interaction energy between the intercalator (L) and DNA base pairs (R), (ΔE_{L-R}) was determined as the difference between the energy of the complex (E_{L-R}) and the sum of the individual energies of the intercalator (E_L) and base pairs (E_R) in the unbound optimized state. If the conformational energy changes undergone by both intercalator and target molecule on binding are neglected (as is often done when series of ligands are compared), the interaction energy can be approximated by

$$\Delta E_{L-R} = E_{L-R} - (E_L + E_R)$$

where E_L and E_R are the intramolecular energies of intercalator and receptor in the bound state. As the ωB97X-

Table 1 Possible hydrogen bonding network between TATA-box binding protein (TBP) and the TATA box

Amino acid residue	Donor atom	DNA residue	Acceptor atom	Percentage of occurrence	Hydrogen bond distance in crystal	Average hydrogen bond distances in simulation (standard deviation)	Secondary structure of the interacting amino acid
Asn (9)	ND2	Thy(24)	O2	37.09	2.72	2.86(0.14)	E
	ND2	Thy(25)	O2	41.34	3.21	2.97(0.20)	
Lys(37)	NZ	Cyt(14)	O3′	5.58	8.23	6.29(2.54)	T
	NZ	Thy(15)	O1P	20.36	7.20	5.58(2.77)	
Arg(38)	NH2	Thy(23)	O1P	5.06	2.62	8.36(2.03)	T
	NH2	Cyt(21)	O3′	3.29	7.66	3.75(0.89)	
	NH2	Thy(22)	O1P	42.82	6.19	3.11(0.42)	
Arg(45)	NH2	Thy(24)	O1P	43.33	3.06	3.16(0.81)	E
	NH2	Thy(23)	O3′	3.12	3.63	3.63(0.83)	
Arg(50)	NH1	Thy(25)	O1P	24.82	2.80	5.34(1.67)	T
	NH1	Ade(26)	O1P	0.10	7.01	8.98(2.08)	
Lys(60)	NZ	Gua(12)	O1P	28.30	4.49	3.86(1.00)	E
	NZ	Ade(11)	O1P	1.63	4.71	5.47(0.98)	
Lys(67)	NZ	Ade(26)	O1P	12.47	4.57	5.23(1.78)	C
	NZ	Thy(25)	O3′	1.45	3.95	5.16(1.56)	
Asn(99)	ND2	Ade(8)	N3	5.48	4.76	3.24(0.28)	E
Arg(136)	NH2	Thy(7)	O1P	18.31	2.81	4.31(0.96)	E
	NH1/NH2	Ade(8)	O2P	92.25	4.58	2.90(0.17)	
Arg(141)	NH1	Ade(9)	O1P	10.53	9.77	7.45(2.79)	T
Lys(151)	NZ	Thy(27)	O1P	9.51	5.80	5.24(1.31)	E
	NZ	Ade(28)	O1P	13.71	4.55	4.47(1.02)	
Lys(158)	NZ	Ade(9)	O1P	1.89	5.69	6.50(1.34)	C
	NZ	Ade(10)	O1P	1.69	5.03	6.16(1.24)	

D functional is recommended to be used without basis set superposition error (BSSE) correction, we have not considered this correction [42].

Results

Deformation of DNA structure

The RMSD of the TBP-bound DNA structure with respect to the initial structure (Fig. S1) indicates that DNA structure does not alter much with respect to the initial structure and the RMSD values become steady within 30 ns. Hence, we analyzed the snapshots of the last 70 ns in our structural studies. It should be noted that deformation of DNA structure is associated with base pair step geometry [62, 63]. We therefore calculated roll, twist, slide, etc., for each base pair step for all snapshots. Comparison of these values within the binding region (6th to 11th base pairs) indicates unusually large roll values and very small twist values for some of those. The roll values for the first and last base pair steps of binding region (i.e. step-5 and step -11) are very high due to partial intercalation of a pair of phenylalanine residues of the TBP protein in these two base pair steps. The variations of these base pair parameters and torsion angles with their standard deviations are given as supplementary materials (Tables S1–S3). Torsion angles are also restricted in the binding region. The δ torsion angles in the binding region are in the 90° range, corresponding to C3′-endo sugar puckers. Thus, the DNA in the complex adopts a regular B-DNA like structure at the both ends but takes up a somewhat unusual deformed conformation in the binding region of the double helix.

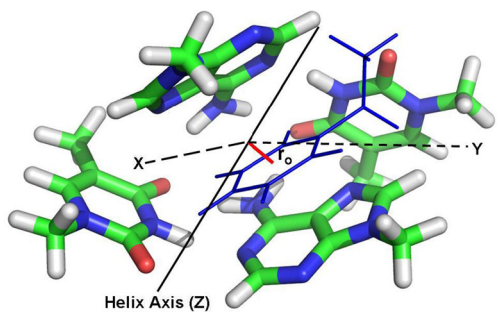

Fig. 2 Model structure of base pairs and the intercalator for quantum chemical study (Phe130 intercalated in TA/TA step)

TBP/DNA interactions

Structural deformation of the duplex DNA can also assist interactions at the binding interface. In many protein/DNA

complexes, the DNA duplex is relatively unperturbed, and sequence recognition occurs via a pattern of hydrogen bonds from amino acid side-chains to amino- and carbonyl- as well as imino-groups of the DNA bases along the surface of the major groove, which is rich in variations of these functional groups. But in the case of the TBP/DNA complex, the interactions occur in the shallow and wide minor groove. It may be noted that variation in disposition of the functional groups is significantly lower in the generally narrower minor groove [64, 65]. In the TBP-DNA complex, most of the H-bonds are formed between phosphate groups of DNA and a few amino acid side chains. A few amino acid residues also make hydrogen bonds with DNA bases as indicated in Table 1. In our simulation of the protein–DNA complex the hydrogen bonding partners are seen to change quite often. Among these bonds, the frequency of two hydrogen bonds involving DNA base atoms are significant: (1) Asn-9 forms a pair of hydrogen bonds with the O2 atoms of Thy-24 and Thy-25, (2) Asn-99 forms a hydrogen bond with the N3 atoms of Ade-8. All other hydrogen bonds are with DNA backbone atoms. We observed hydrogen bonds between Lys-37 and Cyt-14/Thy-15 that were not reported in the crystal structure [15]. The H-bonds between Lys-37, Lys-60, Arg-136, Arg-141 with one strand of the DNA and those involving Arg-38, Arg-45, Arg-50, Lys-67, Lys-151 with the second strand of DNA are quite stable and similar to electrovalent bonds. As shown earlier from quantum chemical studies in gas-phase, these give rise to about -100.0 kcal mol^{-1} energy [59, 62] as compared to an interaction energy of about -10.0 kcal mol^{-1} for a regular hydrogen bond [60]. These hydrogen bonds therefore can give huge stabilization energy to the deformed DNA as well as to the complex even in solution. On the other hand, a few other hydrogen bonds, tabulated in Table 1, were found to be unstable in the simulation. These atoms possibly form stronger hydrogen bonds with water in the simulation setup. Two pairs of Phe residues (Phe-39 and Phe-56, Phe-130 and Phe-147) are observed partially intercalated between the DNA bases at the two ends of the TATA box, and making large van der Waals contacts. Most of the charged Arg and Lys residues, which formed strong hydrogen bonds with the backbone phosphates atoms, are also near these intercalated Phe residues. RMSF fluctuations of all these charged residues are also comparatively lower (Fig. S2). It is possible that, along with hydrogen bonding interactions, these Phe residues also play some role in stabilizing the kinks formed at the two ends of the TATA box.

Extent of intercalation of Phe side chain

To study the stability of the inserted Phe residues, we calculated the variation in distance between the center of mass of partially intercalated Phe residues and the corresponding base pair step centers (r_{inter}) from the simulation trajectory of the TBP-DNA complex. The average value of r_{inter} for Phe-39 and the CT/AG step is 4.24(0.7) Å, the Phe-56 and CT/AG step is 7.11(0.4) Å, the Phe-130 and TA/TA step is 4.67(0.4) Å, Phe-147 and that of the TA/TA step is 7.43(0.3) Å (Fig. 2). We also evaluated the values of r_{inter} for the corresponding Phe residues and intercalated steps from the 22 available crystal structures of TBP-DNA complexes. The values of r_{inter} for Phe-39, Phe-56, Phe-130 and Phe-147 are 5.07(0.7) Å, 7.36(0.3) Å, 4.74(0.2) Å and 7.50(0.2) Å, respectively, in the crystal structure. This indicates that, of the four Phe residues, Phe-130 and Phe-39 are intercalated deeply (Fig. 3). The values of r_{inter} for Phe-39 and Phe-130 are 6.1 Å and 4.8 Å, respectively, in the initial crystal structure of our system (1CDW). However, during simulation Phe-39 intercalated more deeply although it is slightly more dynamic than Phe-130 (Fig. 4). The dependence of intercalation on the parallel positioning of the ring structure was also analyzed by measuring the angles between each base pair axis and phenyl ring axis. For this, we defined the phenyl ring axis system with respect to the center of mass of the phenyl ring, where the *z*-axis is defined as normal to the plane of the phenyl ring, and the *x*-axis is along the line connecting the ring carbon atoms C^{γ} and C^{ζ} (Fig. 5). Therefore, the angle between the base pair normal, i.e., base pair *z*-axis and *z*-axis of the phenyl ring (θ), can be an estimation of their parallel orientation. The base pair and phenyl ring will be parallel if this angle (θ) becomes zero; this stacked geometry plays an important role in stacking interactions. The variations in these angles (θ) during simulation for Phe-39 (CT/AG) and Phe-130 (TA/TA) are shown in Fig. 4. The mean values of θ for Phe-39 (CT/AG) and Phe-

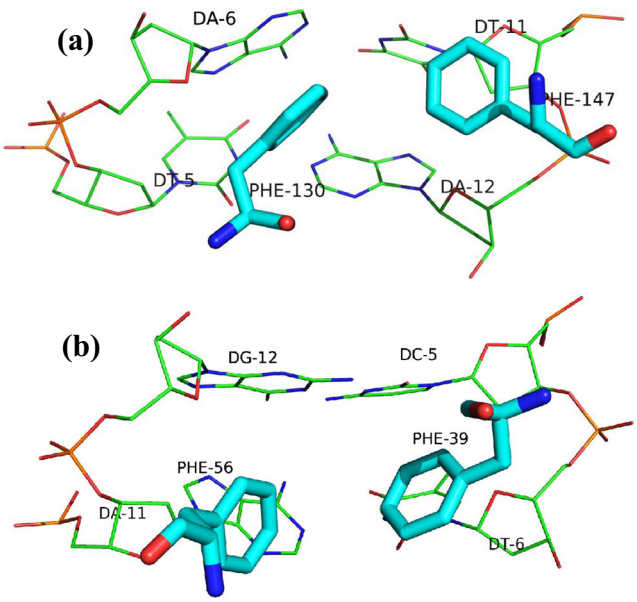

Fig. 3 Intercalation of Phe in the crystal structure. **a** Phe130 and Phe147 intercalated in TA/TA step, **b** Phe39 and Phe56 intercalated in CT/AG step

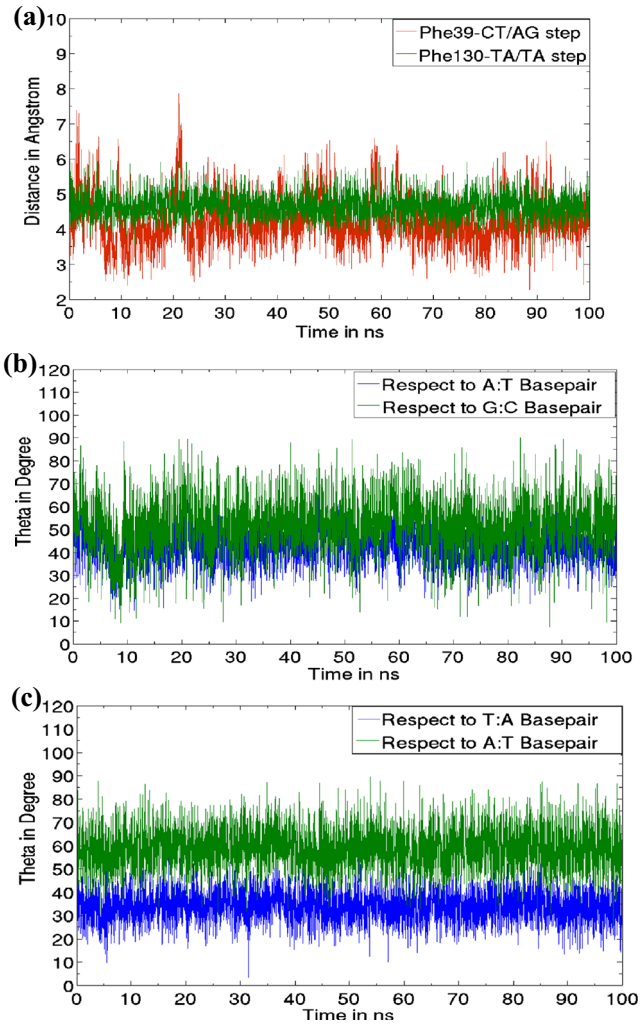

Fig. 4 Variation in **a** the distance between the base pair center and Phe residues (r_{inter}) with time, **b** the angle between the *z*-axis of the Phe-39 ring and base pairs *z*-axes of the AG/CT step (θ), **c** the angle between the *z*-axis of the Phe-130 ring and base pairs *z*-axes of the TA/TA step (θ)

130 (TA/TA) were 42.08° (7.98), 51.17° (11.9) and 58.95° (9.14), 34.33° (6.55), respectively, indicating that, for both steps, the intercalation is not parallel. In the case of the CT/AG step, Phe-39 intercalated nearly parallel to the second A:T basepair, and for the TA/TA step, Phe-130 intercalated parallely to first T:A basepair. A similar feature of intercalation was also observed in the simulation using the recent AMBER ff14SB force field (Fig. S3). In the crystal structures, the values of θ for Phe-39 and corresponding intercalated step were 44.76° (7.7) and 63.22° (10.4). Similarly, for Phe-130 and corresponding intercalated step, the values were 57.94° (5.2) and 37.40° (5.8). For the other Phe residues, the values of θ are comparatively higher. These observations as a whole indicate that the Phe side chains (Phe-130, Phe-39) are intercalated more deeply as well as stacked nearly parallel with the base pairs during the simulation time scale as compared to the initial crystal observation.

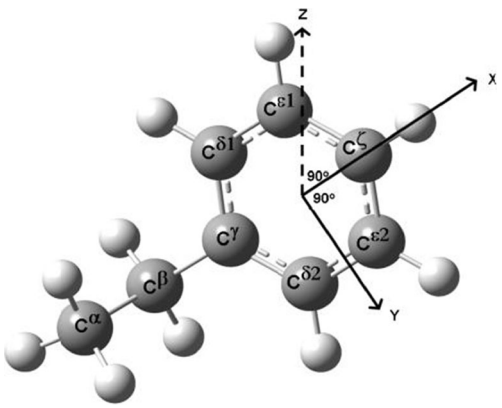

Fig. 5 Axis system of the Phe ring. Atom names are given according to IUPAC notation

Interaction of intercalated phenylalanine with bases

Classical calculation

The distorted conformation of Phe intercalated with the TA/TA and CT/AG base pair dinucleotide steps may not be able to accommodate itself within a double helix structure. These distorted conformations may gain stability due to interaction of the Phe ring within basepair steps. Thus, the interaction energies between the Phe ring and the corresponding basepair steps including the sugar-phosphate backbone were first evaluated by classical forcefields where we considered electrostatic, van der Waals and total interaction energies seperately as shown in Fig. 6. For both TA/TA and CT/AG steps, van der Waals interaction energies were found to contribute more towards the total interaction energies independently. This is due to the hydrophobic character of the Phe residue without any dipole moment. Additionally, more fluctuations in van der Waals and total interaction energies were observed in the case of the CT/AG step (Fig. 6b) as compared to total interaction energy in the TA/TA step (Fig. 6a). From the simulation using ff14SB force field also we observed a similar kind of interactions (Fig. S4). Calculations of the extent of intercalation also indiacte that Phe-39 is more dynamic within the CT/AG step. Correlation between the extent of intercalation (r_{inter}) of the Phe-39 side chain and the van der Waals component of the interaction energy was also reflected in this step (Figs. 4, 6b).

In order to estimate normal stacking energy in TA/TA and CT/AG steps, we generated B-DNA models of these sequences from X-ray fiber diffraction data [66]. Stacking energy between these base pairs were calculated using the ωB97X-D/6-31G(2d,2p) method and found to be −18.46 and −19.71 kcal mol^{-1} for the TA/TA and CT/AG step, respectively. This indicates that the base pair stacking energies are of a magnitude similar to those between the bases of a base pair and Phe residues. However, classical calculations are probably not sufficiently accurate to estimate the intercalation or stacking interactions involving C–H···π and π···π types [20, 34, 67].

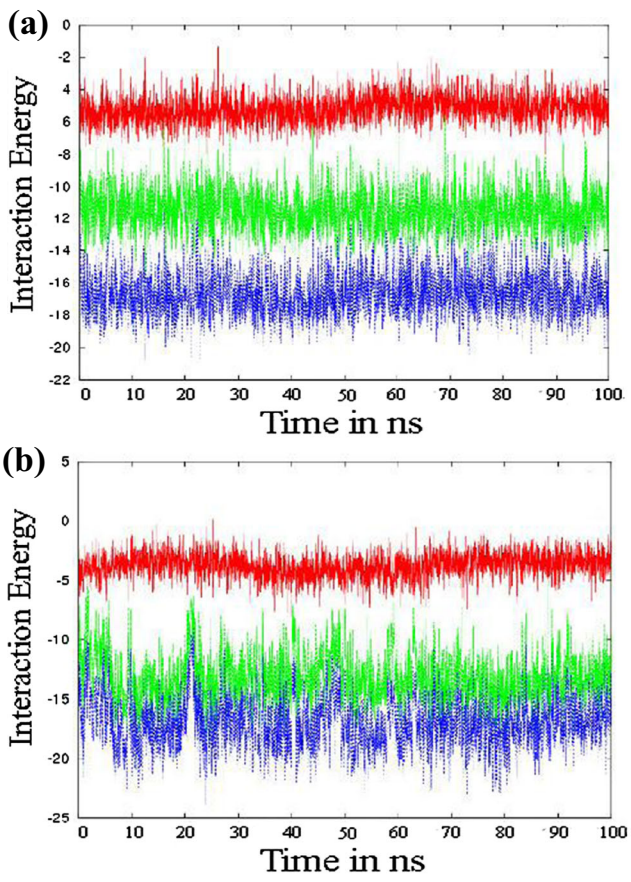

Fig. 6 Interaction energy (in kcal mol^{-1}) between **a** Phe-130 and TA/TA step, and **b** Phe-39 and CT/AG step. Components: *Red* Electrostatic, *green* van der Waals, *blue* total interaction energy

Table 2 Systems used for quantum chemical study. *Phe:* Phenylalanine

Base pair step and intercalated Phe residue	Structure	Description
TA/TA- Phe 130	Crystal structure	Shallow intercalation in TA/TA step
	System 1	Deep intercalation in TA/TA step as observed during simulation
	System 2	Moderate intercalation in TA/TA step as observed during simulation
CT/AG- Phe 39	Crustal structure	Shallow intercalation in CT/AG step
	System 1	Deep intercalation in CT/AG step as observed during simulation
	System 2	Moderate intercalation in CT/AG step as observed during simulation

Table 3 Extent of intercalation for the systems before and after optimization with dispersion-corrected density functional theory (DFT-D)

Base pair step and intercalated Phe residue	Structure	Distance between base pair step center and center of mass of phenyl ring (r_{inter}) in Å		Angle between 1st base pair normal and phenyl ring normal (θ)		Angle between 2nd base pair normal and phenyl ring normal (θ)		Inter base pair hydrogen bond parameter in the optimized geometry		Interaction energy (kcal mol^{-1})
		Before optimization	After optimization	Before optimization	After optimization	Before optimization	After optimization	(O4 – N6) distance	Angle (O4-H…N6)	
TA/TA- Phe 130	Crystal structure	4.76	3.30	59.51°	50.60 °	38.93°	34.62°	3.0 Å, 3.02 Å	166.9°, 160.9°	−26.900
	System 1	3.13	2.76	43.10°	44.63°	24.86°	45.84°	3.0 Å, 3.0 Å	163.1°, 162.6°	−25.454
	System 2	3.68	3.29	57.60°	34.81°	25.52°	50.91°	3.0 Å, 3.0 Å	161.0°, 167.0°	−31.581
CT/AG- Phe 39	Crystal structure	6.07	5.77	62.0°	40.91°	48.0°	39.91°	–		−12.756
	System 1	2.27	2.60	38.37°	32.31 °	39.0°	25.97°			−24.518
	System 2	2.48	2.32	35.65°	22.30°	26.0°	26.84°			−32.393

Optimized geometry and DFT calculation

As exact estimation of stacking interaction between DNA heterocyclic base pairs and the phenyl rings is difficult, they were characterized by ab initio quantum chemical methods. We calculated the interaction energy between the corresponding base pair steps and intercalated Phe ring of optimized model structures as mentioned in Table 2 using DFT-D. The geometric parameters of the systems before and after optimization are given in Table 3. The results indicated that, in all systems, the distance between the center of the base pair step and the center of mass of the phenyl ring (r_{inter}) decreases in the DFT-D optimized geometries. The angles between the Phe ring and the base pairs (θ) also reduced in most cases for the optimized geometries. However, the variations in these angles differed for the two steps. In the case of CT/AG, both angles were reduced significantly (Table 3). Such a reduction was not observed in case of TA/TA, where the angles change after optimization but neither reduce (Table 3). The optimized geometries for all these structures are shown in Fig. 7. The general trend in interaction energy (ΔE) shows that higher penetration, i.e., smaller r_{inter}, favors the interaction, which directly indicates favorable intercalation of the phenyl ring (Table 3). As indicated earlier, the Phe residues did not penetrate deep into the stacked base pair of the crystal structure for the CT/AG and TA/TA step, hence the order of interaction energies is: -12.76 kcal mol^{-1} and -26.9 kcal mol^{-1} respectively. In the case of system 1 and system 2 for the CT/AG step, the interaction energy increased significantly (Table 3) and this may be favored by the higher intercalation by the reduced distances of r_{inter}. The interaction energy of system 2 for the CT/AG step (-32.39 kcal mol^{-1}) is higher than that of system 1 (-24.52 kcal mol^{-1}). We found that, in system 2, the phenyl ring stacked parallely and the angles (θ) were less than those in system 1 (Table 3). Thus, interaction energy also correlates with the geometry of intercalation. In the case of TA/TA, the interaction energy in the crystal structure is almost comparable to the ΔE values of system 1. This can be attributed to the fact that, in crystal as well as in simulation, r_{inter} attains an almost constant value at 4.5 Å for the TA/TA step. However, for system 2 the interaction energy is higher. This indicates that, after optimization, interaction of Phe with bases become more favorable in this case.

Optimized geometry for the TA/TA step, in the presence of Phe, shows that deformed base pair steps having high positive roll values can gain stability by forming inter-base pair hydrogen bonds between thymine and adenine of both strands (Fig. 7), as predicted earlier [24]. The O4 atom of thymine formed hydrogen bonds with the N6 donor atom of adenine along the strand. We also analyzed the hydrogen bond distances (O4–N6 distance) and angles (O4···H–N6 angle) corresponding to optimized geometries as well as simulation trajectories of the TBP-DNA complex. In the optimized geometries, the hydrogen bond distances and angles indicate their extremely high strength (Table 3). In the simulation, these inter-base pair hydrogen bonds along the strand persist very frequently (89%). Thus, the efficiency of the inter-base hydrogen bond network of the TA/TA basepair step helps stabilize the deformed geometry. Furthermore, we found that co-planarity plays a crucial role alongside distance.

Electronic charge distribution of the phenyl ring due to intercalation

There is no possibility of hydrogen bond formation or polar interaction between Phe and the heterocyclic bases. However, due to the adjacent interaction, the Phe residue may acquire some charge redistribution to allow it to interact favorably with the base pairs. Hence, we analyzed these effects through NBO charge analysis. In the case of the Phe residue, we have compared the total natural bond orbital (NBO) charges of the phenyl ring atoms (excluding the ethyl group) in the intercalated-state and free-state for all systems (Table S4). This indicated that the total NBO charge of the Phe ring became more negative. For the TA/TA step, this effect is more pronounced compared to the CT/AG step. However, in the

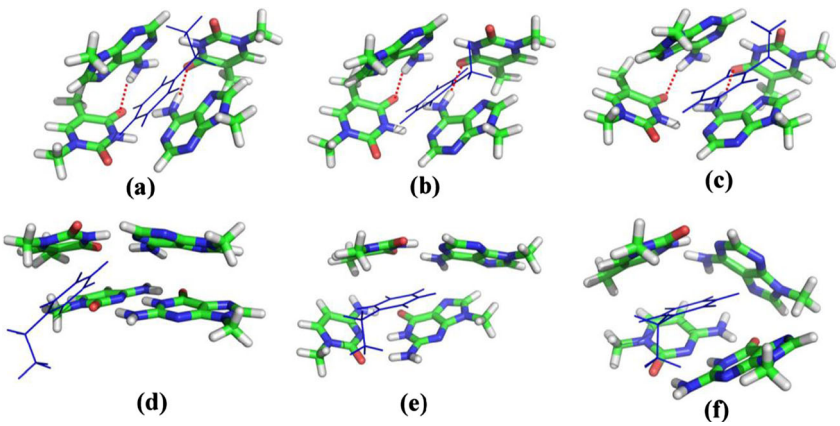

Fig. 7 Optimized geometrics of TA/TA-Phe 130. **a** Crystal structure [3.0 Å, 166.9°; 3.02 Å, 160.9°]. **b** System 1 [3.0 Å, 163.1°; 3.0 Å, 162.6°]. **c** System 2 [3.0 Å, 161.0°; 3.0 Å, 167.0°] and CT/AG-Phe 39. **d** Crystal structure. **e** System 1. **f** System 2. *Dotted lines* Intermolecular hydrogen bonds

case of the crystal geometry of the CT/AG step, the total NBO charge of the phenyl ring is almost equal to the free-state phenyl ring charge. This could be related to the fact that, in the CT/AG crystal structure, the Phe residue did not intercalate deeply (larger r_{inter}). The change in total charge of the phenyl ring may indicate that some redistribution of charge is taking place in the intercalated Phe ring. This could facilitate intercalation, as evident from the interaction parameters shown in Table 3.

Conclusions

This study indicates that deformation of DNA due to binding of TBP is associated with the interactions of charged side chains of Lys and Arg amino acids of TBP with the backbone phosphates atoms of DNA. In addition to formation of these strong hydrogen bonds, side chains of two Phe residues become involved in stacking interactions with the DNA bases. Two such intercalations occur at a TA/TA step and a CT/AG step. The stacking geometry of these bases appears to depend on the sequence of base pairs at the intercalation site—better stacking is observed at the CT/AG step with larger interaction energy. On the other hand, the TA/TA step becomes stabilized by additional cross-strand hydrogen bonds between bases of the two pairs. The geometry optimized structures obtained by DFT-D methods also indicate that stacking of the ethylbenzene between two nearly planar aromatic base pairs does not favor high π–π overlap between the phenyl ring and any of the bases or base pairs, while a kink appears mandatory. Thus, it is expected that the Phe residue of the TBP protein also would give a similar effect. The energetic study signifies that the insertion or intercalation of Phe rings between these basepair steps, at the 5′ and 3′ ends of the TATA box may play an important role in producing and stabilizing the kinks at the two ends of the TATA box, which is very important for the entire biological kingdom. Mutational studies, by perturbing the Phe residues of the TBP, might confirm the importance of such induced kinks during TATA box–TBP interaction and hence their importance to transcriptional activity.

Acknowledgments The authors are thankful to Department of Atomic Energy, Government of India for support through the BARD project.

References

1. Reha D, Kabelac M, Ryjacek F, Sponer J, Sponer JE, Elstner M, Suhai S, Hobza P (2002) Intercalators. 1. Nature of stacking interactions between intercalators (ethidium, daunomycin, ellipticine, and 4′,6-diaminide-2-phenylindole) and DNA base pairs. Ab initio quantum chemical, density functional theory, and empirical potential study. J Am Chem Soc 124(13):3366–3376. doi:10.1021/Ja011490d
2. Geierstanger BH, Wemmer DE (1995) Complexes of the minor groove of DNA. Annu Rev Biophys Biomol Struct 24:463–493. doi:10.1146/annurev.bb.24.060195.002335
3. Rentzeperis D, Marky LA, Dwyer TJ, Geierstanger BH, Pelton JG, Wemmer DE (1995) Interaction of minor groove ligands to an AAAT T/AATTT site: correlation of thermodynamic characterization and solution structure. Biochemistry 34(9):2937–2945. doi:10.1021/bi00009a025
4. Wemmer DE (2000) Designed sequence-specific minor groove ligands. Annu Rev Biophys Biomol Struct 29:439–461. doi:10.1146/annurev.biophys.29.1.439
5. Peek ME, Lipscomb LA, Haseltine J, Gao Q, Roques BP, Garbay-Jaureguiberry C, Williams LD (1995) Asymmetry and dynamics in bis-intercalated DNA. Bioorg Med Chem 3(6):693–699. doi:10.1016/0968-0896(95)00064N
6. Zimmer C, Wahnert U (1986) Nonintercalating DNA-binding ligands: specificity of the interaction and their use as tools in biophysical, biochemical and biological investigations of the genetic material. Prog Biophys Mol Biol 47(1):31–112. doi:10.1016/0079-6107(86)90005-2
7. de Pascual-Teresa B, Gallego J, Ortiz AR, Gago F (1996) Molecular dynamics simulations of the bis-intercalated complexes of ditercalinium and Flexi-Di with the hexanucleotide d(GCGCGC)2: theoretical analysis of the interaction and rationale for the sequence binding specificity. J Med Chem 39(24):4810–4824. doi:10.1021/jm9604179
8. Gallego J, Ortiz AR, de Pascual-Teresa B, Gago F (1997) Structure-affinity relationships for the binding of actinomycin D to DNA. J Comput Aided Mol Des 11(2):114–128. doi:10.1023/A:1008018106064
9. Gago F (1998) Stacking interactions and intercalative DNA binding. Methods 14(3):277–292. doi:10.1006/meth.1998.0584
10. Kolar M, Kubar T, Hobza P (2010) Sequence-dependent configurational entropy change of DNA upon intercalation. J Phys Chem B 114(42):13446–13454. doi:10.1021/jp1019153
11. Baker CM, Grant GH (2007) Role of aromatic amino acids in protein-nucleic acid recognition. Biopolymers 85(5–6):456–470. doi:10.1002/bip.20682
12. Juo ZS, Chiu TK, Leiberman PM, Baikalov I, Berk AJ, Dickerson RE (1996) How proteins recognize the TATA box. J Mol Biol 261(2):239–254. doi:10.1006/jmbi.1996.0456
13. Kim JL, Nikolov DB, Burley SK (1993) Co-crystal structure of TBP recognizing the minor groove of a TATA element. Nature 365(6446):520–527. doi:10.1038/365520a0
14. Kim Y, Geiger JH, Hahn S, Sigler PB (1993) Crystal structure of a yeast TBP/TATA-box complex. Nature 365(6446):512–520. doi:10.1038/365512a0
15. Nikolov DB, Chen H, Halay ED, Hoffman A, Roeder RG, Burley SK (1996) Crystal structure of a human TATA box-binding protein/TATA element complex. Proc Natl Acad Sci USA 93(10):4862–4867. doi:10.1073/pnas.93.10.4862
16. Tan S, Hunziker Y, Sargent DF, Richmond TJ (1996) Crystal structure of a yeast TFIIA/TBP/DNA complex. Nature 381(6578):127–134. doi:10.1038/381127a0
17. Geiger JH, Hahn S, Lee S, Sigler PB (1996) Crystal structure of the yeast TFIIA/TBP/DNA complex. Science 272(5263):830–836. doi:10.1126/science.272.5263.830
18. Kosa PF, Ghosh G, DeDecker BS, Sigler PB (1997) The 2.1-angstrom crystal structure of an archaeal preinitiation complex: TATA-box-binding protein/transcription factor (II)B core/TATA-box. Proc Natl Acad Sci USA 94(12):6042–6047. doi:10.1073/pnas.94.12.6042
19. Nikolov DB, Chen H, Halay ED, Usheva AA, Hisatake K, Lee DK, Roeder RG, Burley SK (1995) Crystal-structure of a Tfiib-Tbp-TATA-element ternary complex. Nature 377(6545):119–128. doi:10.1038/377119a0

20. Umezawa Y, Nishio M (2000) CH/pi interactions in the crystal structure of TATA-box binding protein/DNA complexes. Bioorg Med Chem 8(11):2643–2650. doi:10.1016/S0968-0896(00)00197-8
21. Nishinaka T, Ito Y, Yokoyama S, Shibata T (1997) An extended DNA structure through deoxyribose-base stacking induced by RecA protein. Proc Natl Acad Sci USA 94(13):6623–6628. doi:10.1073/pnas.94.13.6623
22. Nishinaka T, Shinohara A, Ito Y, Yokoyama S, Shibata T (1998) Base pair switching by interconversion of sugar puckers in DNA extended by proteins of RecA-family: a model for homology search in homologous genetic recombination. Proc Natl Acad Sci USA 95(19):11071–11076. doi:10.1073/pnas.95.19.11071
23. Strahs D, Barash D, Qian XL, Schlick T (2003) Sequence-dependent solution structure and motions of 13 TATA/TBP (TATA-box binding protein) complexes. Biopolymers 69(2):216–243. doi:10.1002/bip.10409
24. Pardo L, Pastor N, Weinstein H (1998) Selective binding of the TATA box-binding protein to the TATA box-containing promoter: analysis of structural and energetic factors. Biophys J 75(5):2411–2421. doi:10.1016/S0006-3495(98)77685-4
25. Pastor N, Pardo L, Weinstein H (1997) Does TATA matter? a structural exploration of the selectivity determinants in its complexes with TATA box-binding protein. Biophys J 73(2):640–652. doi:10.1016/S0006-3495(97)78099-8
26. Starr DB, Hoopes BC, Hawley DK (1995) DNA bending is an important component of site-specific recognition by the TATA binding protein. J Mol Biol 250(4):434–446. doi:10.1006/jmbi.1995.0388
27. Patikoglou GA, Kim JL, Sun LP, Yang SH, Kodadek T, Burley SK (1999) TATA element recognition by the TATA box-binding protein has been conserved throughout evolution. Genes Dev 13(24):3217–3230. doi:10.1101/gad.13.24.3217
28. Wong JM, Bateman E (1994) TBP-DNA interactions in the minor groove discriminate between A:T and T:A base pairs. Nucleic Acids Res 22(10):1890–1896. doi:10.1093/nar/22.10.1890
29. Zhao Y, Truhlar DG (2008) The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. Theor Chem Acc 120(1–3):215–241. doi:10.1007/s00214-007-0310-x
30. Zhao Y, Truhlar DG (2008) Density functionals with broad applicability in chemistry. Acc Chem Res 41(2):157–167. doi:10.1021/Ar700111a
31. Hohenstein EG, Chill ST, Sherrill CD (2008) Assessment of the performance of the M05-2X and M06-2X exchange-correlation functionals for noncovalent interactions in biomolecules. J Chem Theory Comput 4(12):1996–2000. doi:10.1021/Ct800308k
32. Hargis JC, Schaefer HF, Houk KN, Wheeler SE (2010) Noncovalent Interactions of a Benzo[a]pyrene diol epoxide with DNA base pairs: insight into the formation of adducts of (+)-BaP DE-2 with DNA. J Phys Chem A 114(4):2038–2044. doi:10.1021/Jp911376p
33. Goerigk L, Grimme S (2011) A thorough benchmark of density functional methods for general main group thermochemistry, kinetics, and noncovalent interactions. Phys Chem Chem Phys 13(14):6670–6688. doi:10.1039/C0cp02984j
34. Sedlak R, Janowski T, Pitonak M, Rezac J, Pulay P, Hobza P (2013) Accuracy of quantum chemical methods for large noncovalent complexes. J Chem Theory Comput 9(8):3364–3374. doi:10.1021/Ct400036b
35. Mukherjee S, Kailasam S, Bansal M, Bhattacharyya D (2014) Energy hyperspace for stacking interaction in AU/AU dinucleotide step: dispersion-corrected density functional theory study. Biopolymers 101(1):107–120. doi:10.1002/Bip.22289
36. Svozil D, Hobza P, Sponer J (2010) Comparison of intrinsic stacking energies of ten unique dinucleotide steps in A-RNA and B-DNA duplexes. can we determine correct order of stability by quantum-chemical calculations? J Phys Chem B 114(2):1191–1203. doi:10.1021/jp910788e
37. Morgado CA, Jurecka P, Svozil D, Hobza P, Sponer J (2010) Reference MP2/CBS and CCSD(T) quantum-chemical calculations on stacked adenine dimers. comparison with DFT-D, MP2.5, SCS(MI)-MP2, M06-2X, CBS(SCS-D) and force field descriptions. Phys Chem Chem Phys 12(14):3522–3534. doi:10.1039/b924461a
38. Ford AR, Janowski T, Pulay P (2007) Array files for computational chemistry: MP2 energies. J Cmput Chem 28(7):1215–1220. doi:10.1002/jcc.20630
39. Morgado C, Vincent MA, Hillier IH, Shan X (2007) Can the DFT-D method describe the full range of noncovalent interactions found in large biomolecules? Phys Chem Chem Phys 9(4):448–451. doi:10.1039/b615263e
40. Cooper VR, Thonhauser T, Puzder A, Schroder E, Lundqvist BI, Langreth DC (2008) Stacking interactions and the twist of DNA. J Am Chem Soc 130(4):1304–1308. doi:10.1021/ja0761941
41. Grimme S (2004) Accurate description of van der Waals complexes by density functional theory including empirical corrections. J Comput Chem 25(12):1463–1473. doi:10.1002/Jcc.20078
42. Chai JD, Head-Gordon M (2008) Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections. Phys Chem Chem Phys 10(44):6615–6620. doi:10.1039/B810189b
43. Grimme S, Antony J, Ehrlich S, Krieg H (2010) A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. J Chem Phys 132:15. doi:10.1063/1.3382344
44. Hobza P (2012) Calculations on noncovalent interactions and databases of benchmark interaction energies. Acc Chem Res 45(4):663–672. doi:10.1021/ar200255p
45. Morgado CA, Svozil D, Turner DH, Sponer J (2012) Understanding the role of base stacking in nucleic acids. MD and QM analysis of tandem GA base pairs in RNA duplexes. Phys Chem Chem Phys 14(36):12580–12591. doi:10.1039/C2cp40556c
46. Barone G, Guerra CF, Bickelhaupt FM (2013) B-DNA structure and stability as function of nucleic acid composition: dispersion-corrected DFT Study of dinucleoside monophosphate single and double strands. Chemistryopen 2(5–6):186–193. doi:10.1002/open.201300019
47. Parker TM, Hohenstein EG, Parrish RM, Hud NV, Sherrill CD (2013) Quantum-mechanical analysis of the energetic contributions to pi stacking in nucleic acids versus rise, twist, and slide. J Am Chem Soc 135(4):1306–1316. doi:10.1021/Ja30633091
48. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The protein data bank. Nucleic Acids Res 28(1):235–242. doi:10.1093/nar/28.1.235
49. Pearlman DA, Case DA, Caldwell JW, Ross WS, Cheatham TE, Debolt S, Ferguson D, Seibel G, Kollman P (1995) AMBER, a package of computer-programs for applying molecular mechanics, normal-mode analysis. molecular-dynamics and free-energy calculations to simulate the structural and energetic properties of molecules. Comput Phys Commun 91(1–3):1–41. doi:10.1016/0010-4655(95)00041-d
50. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA (1996) A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. J Am Chem Soc 118(9):5179–5197. doi:10.1021/ja00124a002
51. Dong C, Yong-Zhi L, Zhi-Chao W, Bo L (2014) Performance of four different force fields for simulations of dipeptide conformations: GlyGly, GlyGly-, GlyGly. Cl-, GlyGly. Na+and GlyGly. (H2O)2. J Mol Model 20(6):2279. doi:10.1007/s00894-014-2279-4
52. An Y, Raju RK, Lu TX, Wheeler SE (2014) Aromatic interactions modulate the 5 ′-base selectivity of the DNA-binding autoantibody ED-10. J Phys Chem B 118(21):5653–5659. doi:10.1021/Jp502069a
53. Kale L, Skeel R, Bhandarkar M, Brunner R, Gursoy A, Krawetz N, Phillips J, Shinozaki A, Varadarajan K, Schulten K (1999) NAMD2: greater scalability for parallel molecular dynamics. J Comput Phys 151(1):283–312. doi:10.1006/jcph.1999.6201

54. Nelson M, Humphrey W, Kufrin R, Gursoy A, Dalke A, Kale L, Skeel R, Schulten K (1995) MDSCOPE - a visual computing environment for structural biology. Comput Phys Commun 91(1–3):111–133. doi:10.1016/0010-4655(95)00045-h
55. D.A. Case VB, J.T. Berryman, R.M. Betz, Q. Cai, D.S. Cerutti, T.E. Cheatham III, T.A. Darden, R.E. Duke, H. Gohlke, A.W. Goetz, S. Gusarov, N. Homeyer, P. Janowski, J. Kaus, I. Kolossvary, A. Kovalenko, T.S. Lee, S. LeGrand, T. Luchko, R. Luo, B. Madej, K.M. Merz, Jr., F. Paesani, D.R. Roe, A. Roitberg, C. Sagui, R. Salomon-Ferrer, G. Seabra, C.L. Simmerling, W.L. Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu, P.A. Kollman (2014) The FF14SB force field. AMBER 14 Reference Manual:29–31
56. Perez A, Marchan I, Svozil D, Sponer J, Cheatham TE, Laughton CA, Orozco M (2007) Refinenement of the AMBER force field for nucleic acids: improving the description of alpha/gamma conformers. Biophys J 92(11):3817–3829. doi:10.1529/biophysj.106.097782
57. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M (1983) Charmm—a program for macromolecular energy, minimization, and dynamics calculations. J Comput Chem 4(2):187–217. doi:10.1002/jcc.540040211
58. Bansal M, Bhattacharyya D, Ravi B (1995) NUPARM and NUCGEN—software for analysis and generation of sequence-dependent nucleic-acid structures. Comput Appl Biosci 11(3):281–287. doi:10.1093/bioinformatics/11.3.281
59. Mukherjee S, Majumdar S, Bhattacharyya D (2005) Role of hydrogen bonds in protein–DNA recognition: effect of nonplanar amino groups. J Phys Chem B 109(20):10484–10492. doi:10.1021/Jp0446231
60. Roy A, Panigrahi S, Bhattacharyya M, Bhattacharyya D (2008) Structure, stability, and dynamics of canonical and noncanonical base pairs: quantum chemical studies. J Phys Chem B 112(12):3786–3796. doi:10.1021/Jp076921e
61. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Mennucci B, Petersson GA, Nakatsuji H, Caricato M, Li X, Hratchian HP, Izmaylov AF, Bloino J, Zheng G, Sonnenberg JL, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Montgomery JA Jr, Peralta JE, Ogliaro F, Bearpark M, Heyd JJ, Brothers E, Kudin KN, Staroverov VN, Kobayashi R, Normand J, Raghavachari K, Rendell A, Burant JC, Iyengar SS, Tomasi J, Cossi M, Rega N, Millam JM, Klene M, Knox JE, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Martin RL, Morokuma K, Zakrzewski VG, Voth GA, Salvador P, Dannenberg JJ, Dapprich S, Daniels AD, Farkas O, Foresman JB, Ortiz JV, Cioslowski J, Fox DJ (2009) Gaussian, Inc, Wallingford CT
62. Samanta S, Chakrabarti J, Bhattacharyya D (2010) Changes in thermodynamic properties of DNA base pairs in protein-DNA recognition. J Biomol Struct Dyn 27(4):429–442. doi:10.1080/07391102.2010.10507328
63. Dickerson RE (1998) DNA bending: the prevalence of kinkiness and the virtues of normality. Nucleic Acids Res 26(8):1906–1926. doi:10.1093/nar/26.8.1906
64. Calladine CR, Drew HR, Luisi BF, Travers AA (2004) Understanding DNA, the molecule and how it works, 3rd edn. Elsevier, London
65. Halder S, Bhattacharyya D (2013) RNA structure and dynamics: a base pairing perspective. Prog Biophys Mol Biol 113(2):264–283. doi:10.1016/j.pbiomolbio2013.07.003
66. Chandrasekaran R, Arnott S (1996) The structure of B-DNA in oriented fibers. J Biomol Struct Dyn 13(6):1015–1027. doi:10.1080/07391102.1996.10508916
67. Umezawa Y, Nishio M (2002) Thymine-methyl/pi interaction implicated in the sequence-dependent deformability of DNA. Nucleic Acids Res 30(10):2183–2192. doi:10.1093/nar/30.10.2183