J Mol Model (2014) 20:2500
DOI 10.1007/s00894-014-2500-5

ORIGINAL PAPER

Understanding the recognition mechanisms of Zα domain of human editing enzyme ADAR1 ($hZ\alpha_{ADAR1}$) and various Z-DNAs from molecular dynamics simulation

Qianqian Wang · Lanlan Li · Xiaoting Wang · Huanxiang Liu · Xiaojun Yao

Received: 10 April 2014 / Accepted: 13 October 2014 / Published online: 26 October 2014

Abstract The Z-DNA-binding domain of human double-stranded RNA adenosine deaminase I ($hZ\alpha_{ADAR1}$) can specifically recognize the left-handed Z-DNA which preferentially occurs at alternating purine-pyrimidine repeats, especially the CG-repeats. The interactions of $hZ\alpha_{ADAR1}$ and Z-DNAs in different sequence contexts can affect many important biological functions including gene regulation and chromatin remodeling. Therefore it is of great necessity to fully understand their recognition mechanisms. However, most existing studies are aimed at the standard CG-repeat Z-DNA rather than the non-CG-repeats, and whether the molecular basis of $hZ\alpha_{ADAR1}$ binding to various Z-DNAs are identical or not is still unclear on the atomic level. Here, based on the recently determined crystal structures of three representative non-CG-repeat Z-DNAs ($d(CACGTG)_2$, $d(CGTACG)_2$ and $d(CGGCCG)_2$) in complex with $hZ\alpha_{ADAR1}$, 40 ns molecular dynamics simulation together with binding free energy calculation were performed for each system. For comparison, the standard CG-repeat Z-DNA ($d(CGCGCG)_2$) complexed with $hZ\alpha_{ADAR1}$ was also simulated. The consistent results demonstrate that nonpolar interaction is the driving force during the protein-DNA binding process, and that polar interaction mainly from helix α3 also provides important contributions. Five common hot-spot residues were identified, namely Lys169, Lys170, Asn173, Arg174 and Tyr177. Hydrogen bond analysis coupled with surface charge distribution further reveal the interfacial information between $hZ\alpha_{ADAR1}$ and Z-DNA in detail. All of the analysis illustrate that four complexes share the common key features and the similar binding modes irrespective of Z-DNA sequences, suggesting that Z-DNA recognition by $hZ\alpha_{ADAR1}$ is conformation-specific rather than sequence-specific. Additionally, by analyzing the conformational changes of $hZ\alpha_{ADAR1}$, we found that the binding of Z-DNA could effectively stabilize $hZ\alpha_{ADAR1}$ protein. Our study can provide some valuable information for better understanding the binding mechanism between $hZ\alpha_{ADAR1}$ or even other Z-DNA-binding protein and Z-DNA.

Keywords Binding free energy · Molecular dynamics simulation · Protein-DNA interaction · Recognition mechanism

Introduction

Left-handed Z-DNA preferentially consists of alternating pyrimidine-purine nucleotides especially the CG-repeat sequences with the alternating *anti* and *syn* conformations, which binds to the Z-DNA-binding protein to affect various important biological functions such as gene regulation and chromatin remodeling [1–5]. Z-DNA has a high-energy conformation relative to the classic right-handed B-DNA [6]. With the unstable thermodynamic structure, Z-DNA can be stabilized by many factors including Z-DNA-binding proteins, negative supercoiling in vivo and high salt condition in vitro [6, 7]. Although a series of crystal structures of Z-DNA complexed with different Z-DNA-binding proteins (e.g., ZBP [8], E3L [9], DLM-1 [10]) have been discovered, human double-stranded RNA adenosine deaminase I ($hZ\alpha_{ADAR1}$) is regarded as the best-characterized one. As the human RNA editing enzyme, ADAR1 has the capability of A to I RNA editing [11], and can alter the important

Electronic supplementary material The online version of this article (doi:10.1007/s00894-014-2500-5) contains supplementary material, which is available to authorized users.

Q. Wang · L. Li · H. Liu (✉)
School of Pharmacy, Lanzhou University, Lanzhou 730000, China
e-mail: hxliu@lzu.edu.cn

X. Wang · X. Yao
State Key Laboratory of Applied Organic Chemistry and Department of Chemistry, Lanzhou University, Lanzhou 730000, China

sequential information of RNA which further leads to the diversity of the corresponding protein. Structural study showed that ADAR1 had two binding motifs, Zα and Zβ, but only the former could specifically bind to Z-DNA and/or Z-RNA with high affinity [12–15]. Zα domain of ADAR1 containing a winged helix-turn-helix (HTH) structure can not only recognize the Z-DNA among numerous B-DNAs, but also induce B-Z junction formation of DNA duplexes [16–18].

In early research, Rich et al. [13] observed that B-DNAs could be converted to the Z-DNAs by binding to $hZ\alpha_{ADAR1}$. It was shown that $hZ\alpha_{ADAR1}$ preferred to bind to Z-DNA rather than B-DNA due to the unfavorable steric hindrance of the latter. Moreover, Z-DNA recognition by $hZ\alpha_{ADAR1}$ mainly depended on the level of negative superhelicity or context rather than base composition. After the solution structure of free $hZ\alpha_{ADAR1}$ was determined [11], they found that the interaction surface between $hZ\alpha_{ADAR1}$ and Z-DNA in the solution agreed well with that in crystal structure of their complex with interaction mapping strategy, indicating that minor conformational variations occurred upon binding to Z-DNA. In their work, they also exhibited the important roles of hot-spot residues determined by previous mutagenesis, which were confirmed again by subsequent studies of this team such as E3L [9], DLM-1 [10]. In reality, the identification of these key residues lays the foundation for understanding the binding mode of Z-DNA-binding proteins and Z-DNA. Apart from these, Lee et al. [17, 19] took full advantage of NMR spectroscopy to explore the B-Z transition mechanism of a DNA duplex induced by $hZ\alpha_{ADAR1}$ and even the conformational and dynamic features of $hZ\alpha_{ADAR1}$ during the B-Z transition. Recently, a X-ray study [20] reported three representative co-crystal structures of non-CG-repeat Z-DNAs, namely $d(CACGTG)_2$, $d(CGTACG)_2$ and d $(CGGCCG)_2$, complexed with $hZ\alpha_{ADAR1}$. On this basis, Lee's group further investigated the sequence discrimination of $hZ\alpha_{ADAR1}$ during B-Z transition of DNA duplexes [21]. The results showed that $hZ\alpha_{ADAR1}$ had much more sequence preference for the CG-repeat Z-DNA than for the non-CG-repeats through multiple sequence discrimination steps, which explained how $hZ\alpha_{ADAR1}$ selectively recognized the alternating CG-repeat in a long genomic DNA.

In recent years, molecular dynamics simulation has been widely used for the systems of different DNA-binding proteins and DNA to explain the protein-DNA recognition mechanism, such as p53-DNA [22], TALE-DNA [23] and CBFα-DNA [24]. However, few computational studies are aimed at the interaction of $hZ\alpha_{ADAR1}$ and Z-DNA although experiments have revealed that $hZ\alpha_{ADAR1}$ recognized Z-DNAs in a conformation-specific manner regardless of DNA sequence [12, 13, 20]. In addition, most existing experimental works focus on the standard CG-repeat Z-DNA rather than the non-CG-repeats. The existence of these questions motivates our present study. In this paper, starting from crystal structures of the CG-repeat and three representative non-CG-repeat Z-DNAs complexed with $hZ\alpha_{ADAR1}$, 40 ns molecular dynamics simulations were performed for each system. Our aims are to reveal the recognition mechanism of $hZ\alpha_{ADAR1}$ and Z-DNA from the perspective of structure and energy, and investigate whether the molecular basis of different Z-DNAs binding to $hZ\alpha_{ADAR1}$ are identical on the dynamic level. In order to explore the conformational changes of $hZ\alpha_{ADAR1}$ induced by Z-DNA, the apo-protein and apo-DNA were also simulated. Our work can provide a complement for the existing experimental results in theory, and is very useful for better understanding the molecular recognition mechanism of Zα domain of Z-DNA-binding protein and Z-DNA.

Materials and methods

Molecular systems

Initial structures were based on the co-crystal structures of CG-repeat DNA and three representative non-CG-repeat DNAs with $hZ\alpha_{ADAR1}$, which were simplified as Ori (PDBID: 1QBJ), F21 (PDBID: 3F21), F22 (PDBID: 3F22) and F23 (PDBID: 3F23), respectively [20]. In these four systems (Ori, F21, F22 and F23), the corresponding Z-DNA sequences were $d(CGCGCG)_2$, d $(CACGTG)_2$, $d(CGTACG)_2$ and $d(CGGCCG)_2$, respectively. The components of each system and the molecular model of $hZ\alpha_{ADAR1}$/Z-DNA complex were shown in Fig. 1. From the inset of Fig. 1, each system is an asymmetric unit composed of three $hZ\alpha_{ADAR1}$ chains (assigned as **a**, **b**, **c**) and three Z-DNA chains (assigned as **d**, **e**, **f**). Each complex consists of one $hZ\alpha_{ADAR1}$ chain and one strand of a DNA duplex, therefore **ad**, **be** and **cf** make up the corresponding three $hZ\alpha_{ADAR1}$/Z-DNA complexes in each system. Chain **d** forms a DNA duplex with chain **e**, but chain **f** loses the matched half. In each $hZ\alpha_{ADAR1}$/Z-DNA

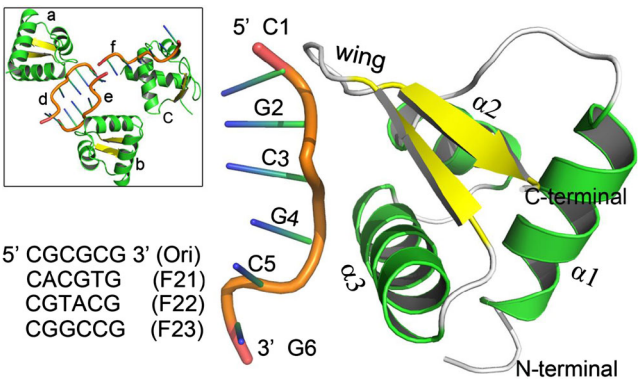

Fig. 1 The overall structure of each system and the molecular model of $hZ\alpha_{ADAR1}$/Z-DNA complex. Z-DNA with "zig-zag" conformation represents four different kinds of DNA sequences, namely $d(CGCGCG)_2$, $d(CACGTG)_2$, $d(CGTACG)_2$ and $d(CGGCCG)_2$

complex, the backbone of Z-DNA exhibits the "zig-zag" conformation. All the helices of hZα_{ADAR1} are almost perpendicular to each other and the helix-turn-helix domain (α2-turn-α3) locates besides the antiparallel β-sheet. In our work, the apo-DNA and apo-hZα_{ADAR1} were extracted from the complex.

Prior to molecular dynamics (MD) simulations, hydrogen atoms were added using the tleap module of Amber 10.0 package [25], and all of the crystal waters were kept. To maintain the system as neutral, a proper number of Na+ions were added. Then, the corresponding system was solvated using TIP3P [26] water in a cubic box by setting the minimum distance from the solute to the edge of box to 10.0 Å.

Molecular dynamics simulation

All the MD simulations were carried out using Amber 10.0 package [25] with ff99SB force field [27] under period boundary conditions. To relieve possible steric clashes and unreasonable overlaps of side chains, we first conducted three rounds of energy minimization using the steepest descent algorithm followed by the conjugate gradient method. Positional restraints were applied to all the non-hydrogen atoms of hZα_{ADAR1}/Z-DNA complex in the first and second rounds with harmonic restraint weights of 2.0 kcal mol^{-1} Å^{-2} and 1.0 kcal mol^{-1} Å^{-2}, respectively. In the third round, the whole system was minimized without any restraint. The system was then gradually heated from 0 to 310.0 K over a period of 50 ps under the NVT ensemble. Subsequently, five separate 100 ps equilibrations were performed to adjust the solvent density with the decreased restraint forces from 2.0, 1.5, 1.0, 0.5 and 0.1 kcal mol^{-1} Å^{-2}, respectively. These were followed by an additional 500 ps equilibration by releasing all the restraints. Last, the production phase was performed for a total of 40 ns without any restraint. Both the equilibration and production phases were carried out in the NPT ensemble at a temperature of 310.0 K and a pressure of 1 atm. The initial velocity was assigned from a Maxwellian distribution at the initial temperature. During the simulation, bond lengths involving hydrogen atom were constrained using SHAKE algorithm [28] and the equations of motion were integrated with 2 fs time step. The particle mesh Ewald (PME) method [29] was used to calculate long-range electrostatic interactions, and the non-bonded cutoff value was set to 10.0 Å. Coordinate trajectory was recorded every 1 ps for later analysis. For apo-DNA and apo-hZα_{ADAR1} systems, similar input parameters were used.

Binding free energy calculation

The molecular mechanics/generalized Born surface area (MM/GBSA) and molecular mechanics/Poisson Boltzmann surface area (MM/PBSA) methods, which have been successfully applied for the binding free energy calculation in a broad range of systems [30–37], were used here to compute the binding free energies of Z-DNAs and hZα_{ADAR1}. Generally, the binding free energy is obtained from the difference:

$$\Delta G_{bind} = G_{complex} - G_{protein} - G_{DNA} \quad (1)$$

The free energy (G) is calculated based on an average over the extracted snapshots from a single-trajectory MD simulation. Each state is estimated from the molecular mechanics energy (E_{gas}), the solvation free energy (G_{sol}) and the solute entropy (−TS).

$$G = H - TS = E_{gas} + G_{sol} - TS \quad (2)$$

$$E_{gas} = E_{ele} + E_{vdW} + E_{int} \quad (3)$$

$$G_{sol} = G_{sol-polar} + G_{sol-np} \quad (4)$$

$$G_{sol-np} = \gamma * SAS \quad (5)$$

$$G_{polar} = E_{ele} + G_{sol-polar} \quad (6)$$

where E_{gas} is the gas-phase energy; E_{int} is the internal energy; E_{ele} and E_{vdW} are the Coulomb and van der Waals energies, respectively. G_{sol} is the solvation free energy and can be decomposed into the polar and nonpolar terms. G_{sol_polar} is the polar solvation contribution obtained by solving the GB and PB equations [35, 38]. In the MM/GBSA and MM/PBSA calculation for polar solvation energy, the solute dielectric constant (ε_{in}) was a key parameter and was testified to increase with the rising charged binding interface [39]. Because the interfaces of hZα_{ADAR1} and Z-DNA in our systems were highly polar, ε_{in} value was then set to 4. G_{sol_np} is the nonpolar solvation contribution and can be estimated by the solvent accessible surface with Amber molsurf module [40] using a 1.4 Å radius of water probe. The surface tension constant γ was set to 0.0072 kcal mol^{-1} Å^{-2} [41, 42]. G_{polar} is the total polar contribution, composed of Coulomb term (E_{ele}) and polar solvation term (G_{sol_polar}). The conformational entropy (S) was estimated by normal-mode analysis of Amber 10.0 nmode program [25, 43].

In addition to the binding free energy calculations, another important and attractive feature of MM/GBSA method is that it can decompose the total binding free energy into group contributions in a structurally non-perturbing formalism by considering molecular mechanics energy and solvation energy without considering the contribution of entropy [30, 34]. We therefore can use this method to obtain the detailed contribution of individual residue of hZα_{ADAR1} and further identify the key residues during the hZα_{ADAR1}-DNA binding process.

Clustering analysis and principal component analysis

The clustering analysis [44] is usually used to generate the representative structure from the MD trajectory. Here, the conformational clustering was carried out for backbone atoms of $hZ\alpha_{ADAR1}$ and/or Z-DNA using the ptraj module of Amber 10.0 [25]. By using the root-mean-square deviation (RMSD) as a measure of the distance between any two given conformations within the trajectory, SOM algorithm was applied to produce clusters. The cutoff value of RMSD was set to 1.4 Å.

Principal component analysis (PCA) is an effective tool to investigate the motion tendency of a protein during the MD simulation [45, 46]. As a linear dimension reduction method, it can identify the most significant fluctuation modes of the protein using a few dimensions [47]. The concerted motions of this protein then become much easier to be monitored and visualized. In our work, PCA was performed for α3 and wing domains of $hZ\alpha_{ADAR1}$ with the ptraj module of Amber 10.0 package [25]. We first calculated the covariance matrix C using the coordinates of Cα atoms of α3 and wing domains, and then diagonalized them to obtain the principal component eigenvectors. Finally, each structure extracted from the trajectory was projected in the collective coordinate space which was defined by the first two largest principal component eigenvectors (PC1 and PC2). In theory, principal components of the protein motion were estimated using the following formula:

$$C_{ij} = \langle (x_i - \langle x_i \rangle) \rangle \langle (x_j - \langle x_j \rangle) \rangle \, pt(i, j = 1, 2, 3, \ldots, 3N) \quad (7)$$

where x_i and x_j are the Cartesian coordinates of the *i*th and *j*th of Cα atom, respectively. N represents the number of Cα atoms, and x_i and x_j are the time average over all the configurations from the simulation [47]. Prior to clustering and principal component analysis, each snapshot was aligned with the first structure of the selected trajectory to remove rigid-body motions.

Results and discussion

The global structural properties

The equilibration of molecular dynamics simulations was monitored from the convergence of the root-mean-square deviation (RMSD) of $hZ\alpha_{ADAR1}$ and Z-DNA backbone atoms relative to the initial structure (Fig. 2). As shown in Fig. 2, RMSD values of **cf** complex in four systems (Ori, F21, F22 and F23) fluctuate in a wider range than that of the corresponding **ad** and **be** complexes, indicating that **cf** complex has larger conformational changes than **ad** and **be** in each system. The similar phenomenon happens to **c** and **f** chains as well. For **ad** complex (Fig. 2a), Ori and F21 remain stable around 1.7 Å during the whole simulation and F22 becomes stable from 25 ns at 2.3 Å. RMSD values of F23 are very stable from 5 ns, but have a sudden fluctuation at 25 ns. By extracting the structures of 25 ns and 35 ns, we found that this obvious bounce of RMSDs was mainly ascribed to the large conformational changes of helix α2 of $hZ\alpha_{ADAR1}$, which also could gain information from chain **a** of F23 system (Fig. 2d). Compared with **ad** and **cf**, RMSDs of **be** complex fluctuate in a relatively small range. As shown in Fig. 2b, RMSD values of **be** complex in all four systems converge from 20 ns. Specifically, Ori/F21 and F22/F23 stay around 1.4 Å and 2.2 Å, respectively. Furthermore, conformational changes of the $hZ\alpha_{ADAR1}$ (chain **b**) and the Z-DNA (chain **e**) in **be** complex are also kept very small, as shown in Fig. 2e and Fig. 2h. Therefore, **be** complex is considered as the most stable one in each system.

To investigate the effect of Z-DNA's binding on the internal compactness of $hZ\alpha_{ADAR1}$ in different complexes, we calculated the windowed average values of the radius of gyration (R_g) of $hZ\alpha_{ADAR1}$ of every 200 ps during the whole simulation and the average values of R_g of the last 10 ns trajectory (Fig. 3). In each system, we observed that R_g values of chain **c** fluctuated more notably than that of chain **a** and **b**, which was consistent with the changes of RMSD values above. Chain **a** of the F21 system also has an obvious bounce at 12 ns, but with a smaller range than chain **c**. Relatively, chain **b** of each system remains the most stable with the minimum R_g value during the whole trajectory. For instance, the average value of R_g of chain **b** in Ori system is 10.9 Å, notably lower than those (11.1 Å and 11.2 Å) of chain **a** and **c**. These suggest that the internal structure of $hZ\alpha_{ADAR1}$ corresponding to chain **b** is more compact than other chains in the same system. Additionally, the root-mean-square fluctuations (RMSFs) of backbone atoms for each residue/nucleotide in different $hZ\alpha_{ADAR1}$/Z-DNA complexes were calculated by averaging the 10,000 frames of the last 10 ns using the Amber 10.0 software. As shown in Fig. 4, the overall tendencies of $hZ\alpha_{ADAR1}$/Z-DNA complex are similar in four systems, and in each system **be** complex has lower RMSF values than **ad** and **cf** complexes.

Taken together, the $hZ\alpha_{ADAR1}$/Z-DNA complex corresponding to **be** in each system is the most stable, and therefore our analysis and discussions below will focus on the **be** complex.

Binding free energy calculation

To explore the interaction features of $hZ\alpha_{ADAR1}$ and Z-DNA from the perspective of energy, the binding free energy calculations were performed. The snapshots used to calculate the binding free energy were obtained from the last equilibrated

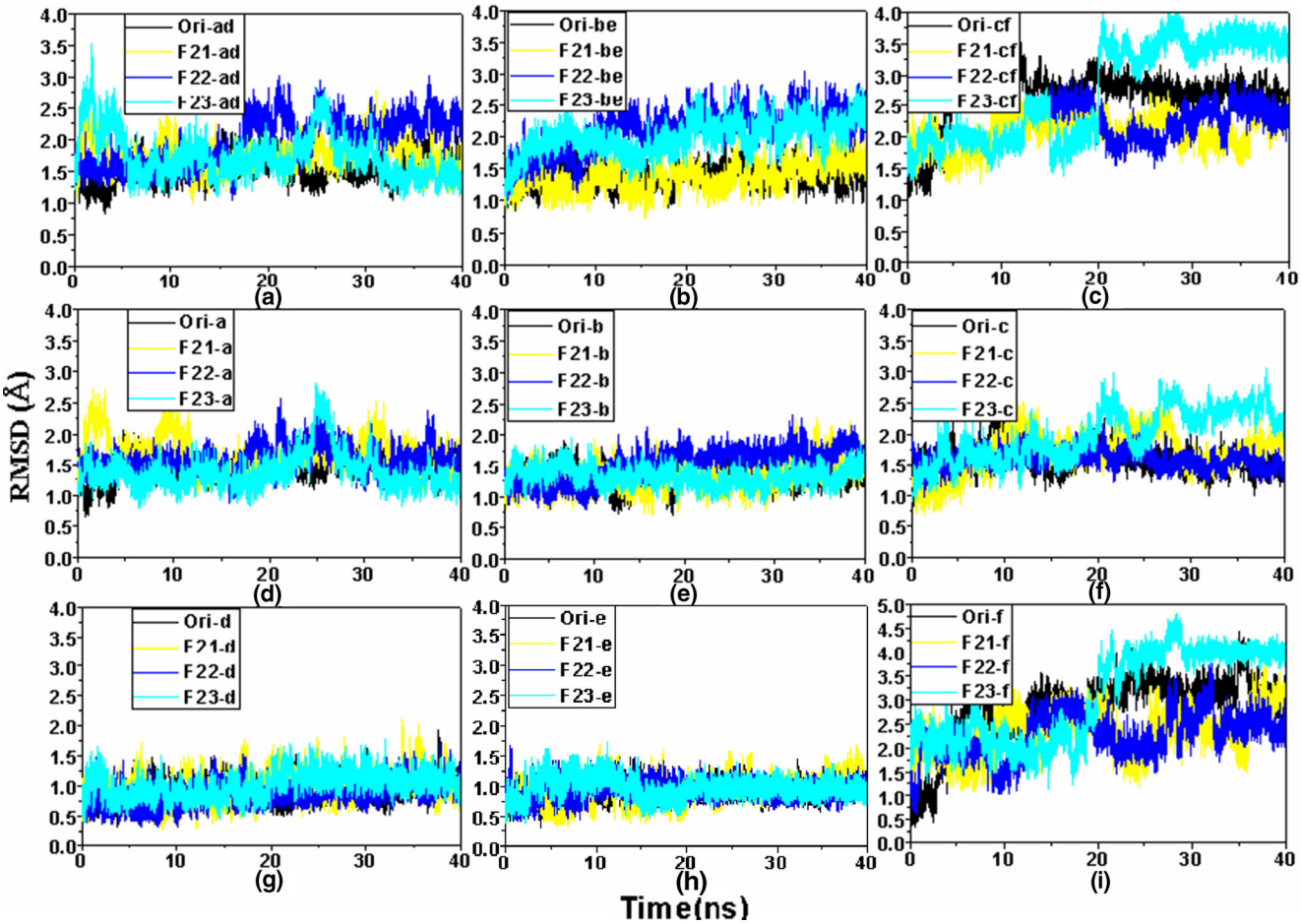

Fig. 2 **A-I** Time dependences of RMSDs of backbone atoms relative to the initial structure for (**a**) **ad** complex (hZα_{ADAR1}/Z-DNA), (**b**) **be** complex (hZα_{ADAR1}/Z-DNA) and (**c**) **cf** complex (hZα_{ADAR1}/Z-DNA); (**d**) chain **a** (hZα_{ADAR1}), (**e**) chain **b** (hZα_{ADAR1}) and (**f**) chain **c** (hZα_{ADAR1}); (**g**) chain **d** (Z-DNA), (**h**) chain **e** (Z-DNA) and (**i**) chain **f** (Z-DNA)

10 ns trajectory. Here, five hundred snapshots were extracted at 20 ps intervals for the enthalpy calculation. Given the high computational demand, two hundred snapshots were used to calculate the entropy. The average binding free energies and the detailed contributions of various energy components were shown in Table 1. As can be seen, the van der Waals contributions (ΔE_{vdw}) commonly dominate during the binding of hZα_{ADAR1} to Z-DNAs in all four systems. The resulting balance (ΔG_{polar}) of electrostatic interactions in vacuum and solvent, namely $\Delta E_{ele}+\Delta G_{sol\_polar}$, provides favorable contributions for their binding as well. Totally, the enthalpic contributions of hZα_{ADAR1}/Z-DNA complex in Ori (−55.34 kcal mol^{-1}), F22 (−55.64 kcal mol^{-1}) and F23 (−55.76 kcal mol^{-1}) systems were similar, and lower than that (−54.32 kcal mol^{-1}) in F21 system. Additionally, as we know, the changes of solute entropy derive from the burial of hydrophobic groups on binding and the loss of solute conformational degrees of freedom. In this work, for F21 and F22 systems, their losses of configurational entropy upon binding to hZα_{ADAR1} were larger than that for Ori and F23 systems by ~1−2 kcal mol^{-1}, resulting in that their final binding free energies were slightly higher than that for Ori and F23 systems.

The identification of key residues and binding mode analysis of hZα_{ADAR1} and Z-DNA

By decomposing the binding free energy into each residue, we can obtain the energy contribution of each residue of hZα_{ADAR1} protein and further identify the hot-spot residues for the hZα_{ADAR1}/Z-DNA binding. In this paper, we considered the residue with energy contribution larger than 2 kcal mol^{-1} as a hot-spot residue. In order to directly find out the hot-spot residues, the DNA-residue spectra was then depicted, shown in Fig. 5. From Fig. 5 (a, c, e and g), it is easy for us to observe that almost all the hot-spots locate at the helix α3 and wing domains, confirming the importance of these two domains. Previous studies [11, 20] found that hZα_{ADAR1} interacted with Z-DNA mainly through helix α3 and wing domains, which was consistent with our result. To quantify the contributions of helix α3 and wing domains, we further summed the contributions of residue 169–181 (helix α3)

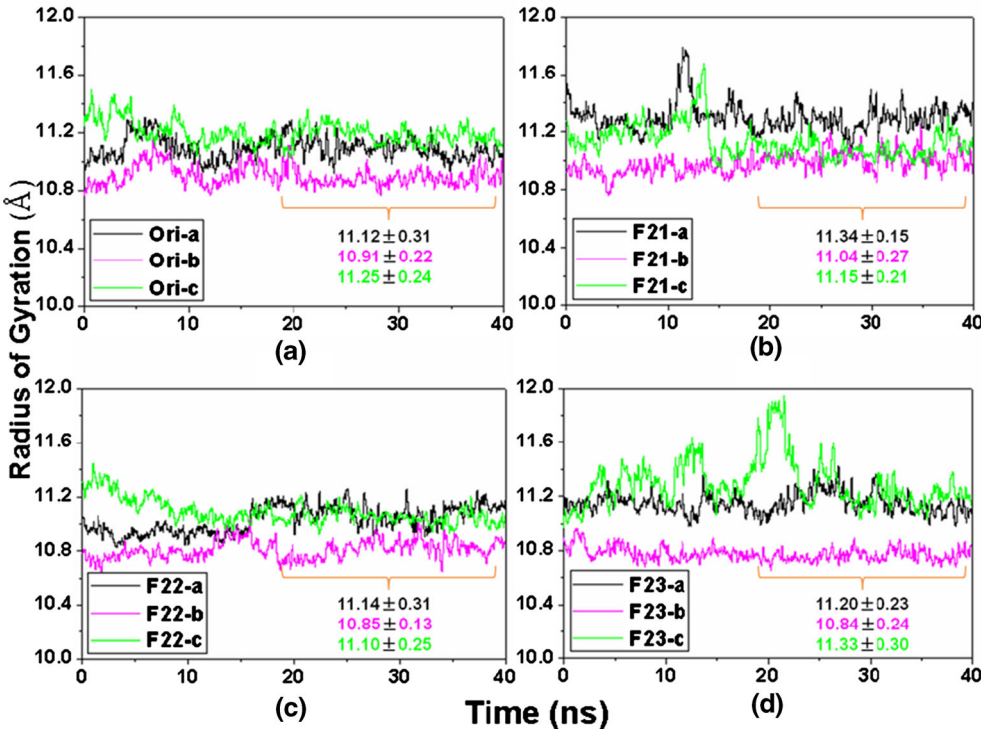

Fig. 3 **A-D** The windowed average values of the radius of gyration (R_g) of every 200 ps of $hZ\alpha_{ADAR1}$ for (**a**) Ori, (**b**) F21, (**c**) F22 and (**d**) F23 systems. For clarity, the average values of R_g of the last 10 ns trajectory were also shown

and residue 190–193 (wing domain), respectively. As can be seen from Table S1, helix α3 is the dominant binding domain for $hZ\alpha_{ADAR1}$/Z-DNA complex and provides 40.48 %, 43.12 %, 43.35 % and 41.12 % contributions for Ori, F21, F22 and F23 systems, respectively. Although the wing domain only consists of four residues, this domain contributes 10.84 %, 7.24 %, 8.99 % and 10.92 % to the binding process in four systems, respectively. Moreover, the energy contribution sum of helix α3 and wing domains is larger than 50 % of the total enthalpy, indicating that these two domains indeed are the key binding domains of DNA during the $hZ\alpha_{ADAR1}$/Z-DNA interaction.

Five common hot-spot residues were identified in four systems, namely Lys169, Lys170, Asn173, Arg174 and Tyr177, in accordance with the previous experimental result [11, 12, 20]. For Z-DNAs, most of the energy contributions came from nucleotides 2–5. On the basis of the identified hot-spot residues, we further decomposed their contributions into polar and nonpolar terms to reveal the major source of interaction. From Fig. 5 (b, d, f and h), most of the residues of helix α3 provide both polar and nonpolar contributions with the polar contribution more notable. Different from helix α3, the hot-spot residues in wing domain interact with Z-DNA mainly by nonpolar interaction. Residue Pro192, as the most important one in this domain, plays a key role in each system, especially in Ori, F22 and F23 systems (nonpolar interactions under −2 kcal mol^{-1}). Its dominant nonpolar interaction mainly comes from the strong hydrophobic contacts between the pyrrole ring of Proline and Z-DNA, which can be reflected from the early experiments [12, 20].

To identify the source of the large polar contribution of helix α3, we further analyzed the hydrogen bond formations between $hZ\alpha_{ADAR1}$ and Z-DNA. From the 40 ns trajectory 40,000 snapshots were monitored to identify all direct hydrogen bond (H-bond) interactions (Table 2 and Table S2-S4). A hydrogen bond is considered to be formed if the donor-acceptor distance is less than 3.5 Å and the donor-hydrogen-acceptor angle is larger than 120°. Here, we only gave out hydrogen bonds whose percentage of occupation exceeded 20 %. With regard to the Ori system (shown in Table 2), it is notable that all of the H-bond acceptor atoms locate on the sugar-phosphate backbone of Z-DNA. Most of the protein residues involving hydrogen bond interaction belong to helix α3, which is in good agreement with the binding free energy calculation result that polar interaction dominates in α3 domain. Four H-bonds are of great importance in Ori system, and they exist between the side-chain N−H of Asn173 and O1P of G4 nucleotide of Z-DNA, the side-chain N−H of Arg174 and O1P of C5 nucleotide, the O−H of Tyr177 and O2P of C3 nucleotide, and the side-chain N−H of His159 and O2P of G2 nucleotide, respectively, with the occupation over 50 %. The hydrogen bonds between the side-chain N−H of Asn173 and O5' of G4, the O−H of Tyr177 and O1P/P of C3 also possess large occupied percentages, namely 38.62 %, 46.15 % and 40.07 %, respectively. In addition, Lys169 with G4 nucleotide and Lys170 with G4/C5 nucleotide also have obvious hydrogen bond formations. Combined with the

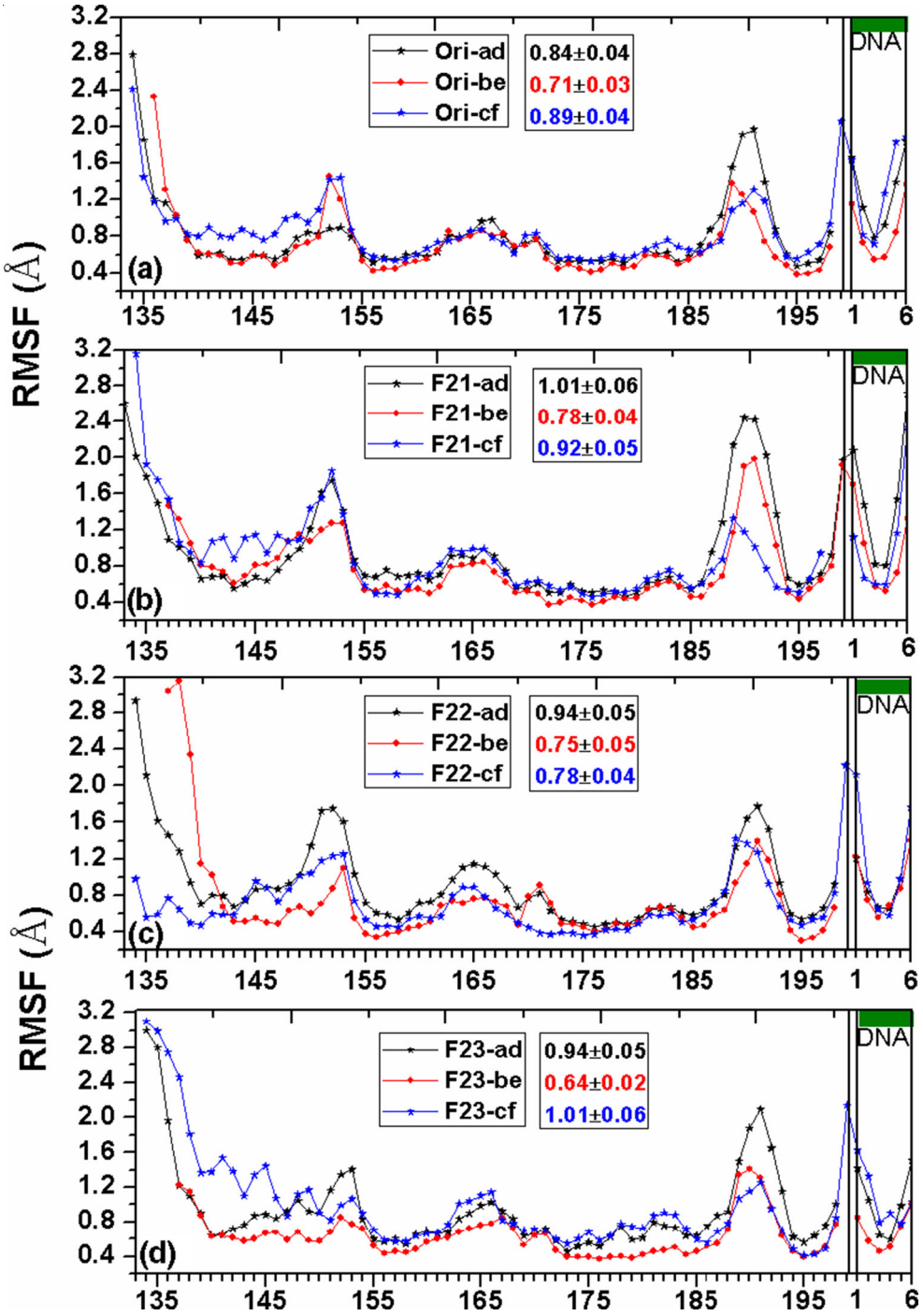

Fig. 4 **A-D** The calculated RMSF values of backbone atoms as a function of residue/nucleotide of hZα_{ADAR1}/Z-DNA complex for (**a**) Ori, (**b**) F21, (**c**) F22 and (**d**) F23 system. Only the last equilibrated 10 ns trajectory was considered. For comparison, the average RMSFs of all involved residues for each system were also given

spectra of residue's contribution, the hydrogen bond formations from residue His159, Lys169, Lys170, Asn173 and Arg174 make these residues have large polar contributions. However, why does Tyr177 participate in many H-bond formations and the polar contribution is almost zero? According to the calculated result of energy decomposition, the direct Coulombic interaction (−2.85 kcal mol^{-1}) from residue Tyr177 indeed is obviously favorable for the binding, but most of which is neutralized by its unfavorably polar solvation contribution (2.77 kcal mol^{-1}).

Hydrogen bond formations of hZα_{ADAR1} and Z-DNA in F21, F22 and F23 systems are similar to that in Ori system. As shown in Table S2-S4, the H-bond occupied percentages between Asn173 and O1P of the fourth nucleotide in three systems are up to 91.22 %, 97.21 % and 97.02 %, respectively. Tyr177 also forms obvious hydrogen bonds with O1P of C3/T3/G3 (occupation 97.17 %, 96.96 % and 99.75 %, respectively). However, contrary to simultaneous H-bond formations of Tyr177 with P, O1P and O2P of C3 nucleotide in Ori system, Tyr177 forms the unique hydrogen bond in F21,

Table 1 The binding free energy and its components obtained from the MM/GB (PB) SA calculations for hZα_{ADAR1}/Z-DNA complex (kcal mol^{-1})

Contributions	Ori-be	F21-be	F22-be	F23-be
ΔE_{ele}	−258.62	−256.81	−289.57	−290.04
ΔE_{vdw}	−38.13	−35.56	−32.44	−33.14
ΔE_{int}	0.00	0.00	0.00	0.00
ΔE_{gas}	−296.74	−292.37	−322.01	−323.18
ΔG_{sol_np}	−5.99	−5.95	−6.02	−6.43
$\Delta G_{sol\_polar, PB}$	232.51	228.42	257.13	258.94
$\Delta G_{sol, PB}$	226.52	222.47	251.11	252.51
$\Delta G_{polar, PB}$	−26.11	−28.39	−32.44	−31.10
ΔH_{PB}	−70.22	−69.90	−70.90	−70.67
$\Delta G_{sol\_polar, GB}$	247.39	244.01	272.40	273.86
$\Delta G_{sol, GB}$	241.40	238.05	266.37	267.43
$\Delta G_{polar, GB}$	−11.22	−12.80	−17.18	−16.18
ΔH_{GB}	−55.34	−54.32	−55.64	−55.76
-TΔS	30.27	31.11	32.31	30.25
$\Delta G_{bind, PB}$	−39.95	−38.79	−38.59	−40.42
$\Delta G_{bind, GB}$	−25.07	−23.21	−23.33	−25.51

F22 and F23 systems. Differently, Arg174 forms the only one with O1P of C5 nucleotide in Ori system, but forms several H-bonds with different atoms in the other three systems. Of course, there are also other residues forming H-bond, such as His159 and Lys169. Similar to Ori system, almost all the involved residues locate in the α3 domain.

To show the interfacial feature of hZα_{ADAR1}/Z-DNA complex intuitively, we analyzed the surface charge distributions of hZα_{ADAR1} complexed with four kinds of Z-DNAs (Fig. 6). As shown in Fig. 6, the surface of hZα_{ADAR1} close to Z-DNA is rich in positive charges, indicating that hZα_{ADAR1} fits the Z-DNA both in terms of shape and electrostatic match. Besides, the similar charge distributions on hZα_{ADAR1} surface in four systems further imply their similar binding modes.

As we know, the common feature of the CG-repeat and non-CG-repeat Z-DNAs is that both Z-DNAs share the common "zig-zag" conformation. Although the CG-repeat Z-DNA and non-CG-repeat Z-DNA have completely different sequences, both can be recognized by hZα_{ADAR1}, showing that the hZα_{ADAR1} identifying the Z-DNA does not only rely on their sequences. By our analysis and the reported experimental results, we can conclude that the Z-DNA recognition by hZα_{ADAR1} is conformation-specific.

Conformational variation of hZα_{ADAR1} upon binding to Z-DNA

In order to further explore the conformational changes of hZα_{ADAR1} induced by Z-DNA, additional MD simulations for isolated hZα_{ADAR1} and Z-DNA extracted from their complex were then performed. Using the obtained trajectories, the root-mean-square fluctuation (RMSF) values of per-residue backbone atoms were first calculated to reveal the changes of structural flexibilities of hZα_{ADAR1} upon binding to Z-DNA. As shown in Fig. 7, the fluctuation trends of RMSF values of hZα_{ADAR1} protein in free state have obvious differences from that in Z-DNA-bound state. For the key binding domain, helix α3, it can be seen from Fig. 7 that almost all the residues in this domain have the lower RMSF values in Z-DNA-bound state than that in free state, suggesting that Z-DNA's binding makes the flexibility of helix α3 reduced. The RMSFs of wing domain also weakly decrease after binding to Z-DNA in Ori, F22 and F23 systems, but largely increase in F21 system. By analyzing the binding modes of four complexes, we found that this phenomenon was associated with the different A2 nucleotide of F21 system from the G2 nucleotide of other three systems. Relative to the G2 nucleotide, A2 nucleotide is close to Thr191 (the important residue of wing domain), and they can form many non-polar interactions. The participation of a favorable interaction leads to the obvious increasing of flexibility of residue Thr191 and even other residues in wing domain. Additionally, in Fig. 7, we observed that the RMSFs of loop between α1 and α2 of the bound hZα_{ADAR1} in Ori system were obviously larger than that without Z-DNA, whereas the RMSFs of this domain in the other three systems altered very little. By further analysis, we found that the big RMSF difference of α1-α2 loop was mainly derived from the structural changes of residue Gly151-Lys154, especially Glu152 and Lys154. In apo-hZα_{ADAR1} protein, the long side chains of residue Glu152 and Lys154 stretch in two opposite directions. However, upon binding to Z-DNA, their side chains convert into the same direction. It is likely that the binding of Z-DNA leads to the increase of the long-range electrostatic interactions of Glu and Lys. Meanwhile, the large conformational fluctuations of these two residues prompt the structures of their adjacent residues (residue Gly151 and Gly153) to change during the simulation. Therefore, we consider that the binding of standard Z-DNA (d(CGCGCG)$_2$) has influence not only on the conformation of hZα_{ADAR1} site but also on the α1-α2 loop away from the site. Correspondingly, the flexibilities of Z-DNA also alter upon the binding of hZα_{ADAR1}. That is, RMSFs of six nucleotides are notably different from each other, instead of the similar values in the presence of hZα_{ADAR1}.

To reveal the detailed structural changes of hZα_{ADAR1} upon Z-DNA's binding, a structural comparison of apo-hZα_{ADAR1} (without Z-DNA) and Z-DNA-bound hZα_{ADAR1} (with Z-DNA) was next carried out. The representative structures were obtained by clustering analysis [44] with the last 10 ns trajectory. Figure 8 shows the superposition of the apo-hZα_{ADAR1} and Z-DNA-bound hZα_{ADAR1}. In Fig. 8, we observed that the conformational changes of Lys169 and Asn173 in four systems were similar, and the conformational changes

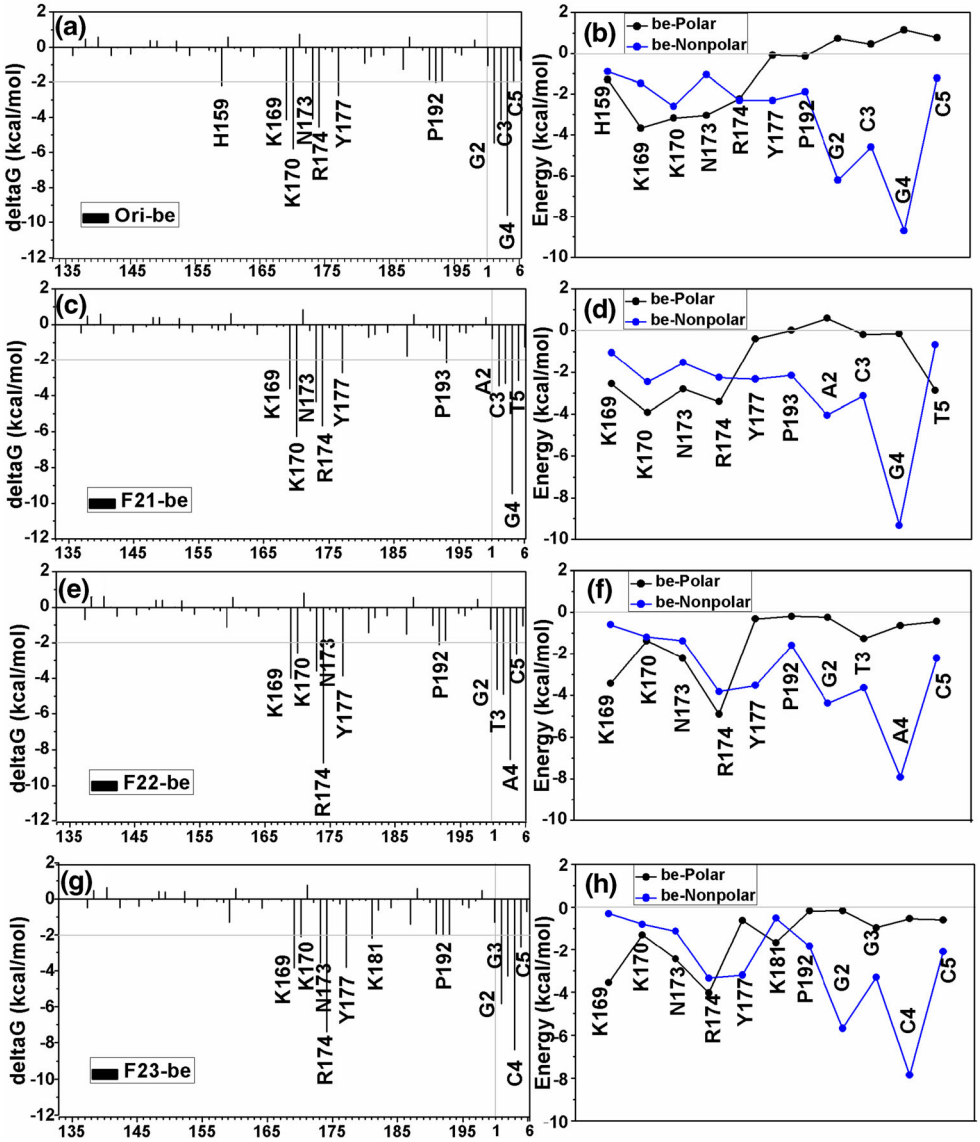

Fig. 5 **A-H** The DNA-residue interaction spectra between Z-DNA and hZα_{ADAR1}, and the polar/nonpolar contributions of hot-spot residues in **A**/**B** Ori, **C**/**D** F21, **E**/**F** F22 and **G**/**H** F23 system. Residues with energy contribution under -2.0 kcal mol^{-1} were labeled

of Arg174 in Ori and F23 systems were smaller than that in F21 and F22 systems. Due to these residues belonging to helix α3, their conformational changes after binding to Z-DNA favorably produce more polar interactions with Z-DNA, as reflected from H-bonds of Fig. 8. As for Thr191, Pro192 and Pro193 of wing domain, which make great contributions to the hZα_{ADAR1}/Z-DNA binding by van der Waals interactions, also have obvious conformational changes during the binding process. Compared with apo-hZα_{ADAR1}, the side chains of Tyr177 in Z-DNA-bound hZα_{ADAR1} of four systems twist by a larger margin around the Cβ atom. Actually, the twist of this side chain is not only favorable to the hydrogen bond formation between Tyr177 and the third nucleotide of Z-DNA, but also favorable to the direct CH-π interaction between the benzene ring of Tyr177 and carbon-8 atom of the *syn* deoxyguanosine or deoxyadenine at the fourth position. Generally, pyrimidine nucleotide preferentially applies the *anti* conformation and purine prefers the *syn* conformation in Z-DNA [9, 10, 20]. However, in F23 system, although the deoxycytidine of C4 nucleotide adopts the *syn* conformation, it is still stabilized by hZα_{ADAR1} through CH-π interactions between the benzene ring of Tyr177 and carbon-5/carbon-6 atoms of C4 (Fig. 8d), which agrees with the experimental result [20].

Furthermore, due to helix α3 and wing domains as the key binding domains, we care more about their conformational changes during the binding process. Thus here, a principal component analysis (PCA) of Cα atoms of helix α3 and wing domains for apo- and Z-DNA-bound hZα_{ADAR1} was performed. The conformations of hZα_{ADAR1} extracted from the last 10 ns trajectory were used for PCA. In PCA, over 75 % of these motions can be accounted for by the first two principal components in each system. To visualize the conformational spaces, the first two principal components were plotted in

Table 2 The occupied percentages of the formed hydrogen bonds between $hZ\alpha_{ADAR1}$ and Z-DNA in Ori system. Atom1, atom2 and atom3 are indicative of acceptor atom, donor atom and hydrogen atom of the donor, respectively

Acceptor Nucleotide (atom 1)	Donor Residue (atom 2-atom 3)	Distance (Å)	Angle (°)	Percentage of occupation (%)
G4 (O1P)	Asn173 (ND2-HD22)	2.87	152.14	95.89
C5 (O1P)	Arg174 (NH1-HH11)	2.83	156.87	69.35
C3 (O2P)	Tyr177 (OH-HH)	2.78	165.26	68.31
G2 (O2P)	His159 (NE2-HE2)	2.90	151.16	66.62
C3 (O1P)	Tyr177 (OH-HH)	2.85	150.83	46.15
C3 (P)	Tyr177 (OH-HH)	3.37	145.43	40.07
G4 (O5)	Asn173 (ND2-HD22)	3.26	141.45	38.62
G2 (O1P)	His159 (NE2-HE2)	3.01	145.87	38.06
G2 (P)	His159 (NE2-HE2)	3.37	143.11	32.37
G4 (O1P)	Lys169 (NZ-HZ2)	2.77	153.26	31.28
G4 (O2P)	Lys170 (NZ-HZ2)	2.79	158.80	28.85
C5 (O2P)	Lys170 (NZ-HZ1)	2.82	160.12	28.84
G4 (O1P)	Lys169 (NZ-HZ3)	2.77	152.69	25.34
G4 (O1P)	Lys169 (NZ-HZ1)	2.77	152.05	24.74
G4 (O2P)	Lys170 (NZ-HZ3)	2.79	158.61	24.00
C5 (O2P)	Lys170 (NZ-HZ2)	2.82	160.44	22.54
G4 (O2P)	Lys170 (NZ-HZ1)	2.79	158.99	20.34

Fig. 9. As can be seen from Fig. 9, it is notable that the distribution of conformational spaces for Z-DNA-bound $hZ\alpha_{ADAR1}$ is narrower than that for apo-$hZ\alpha_{ADAR1}$, suggesting that helix α3 and wing domains of $hZ\alpha_{ADAR1}$ become more stable upon the binding of Z-DNA.

In addition to the ADAR1 studied in this work, Zα domain is also present in three other Z-DNA-binding proteins: E3L protein [9], DLM-1 protein [10] and PKR-like kinase (PKZ) [48]. Experimental studies showed that they recognized the Z-DNA solely based on the feature of its "zig-zag" conformation as well [48–50]. Thus, a general mechanism of Z-DNA recognition is that the Z-DNA-binding proteins recognize the CG-repeat and non-CG-repeat Z-DNAs through their common "zig-zag" conformation. Helix α3 and wing domains of

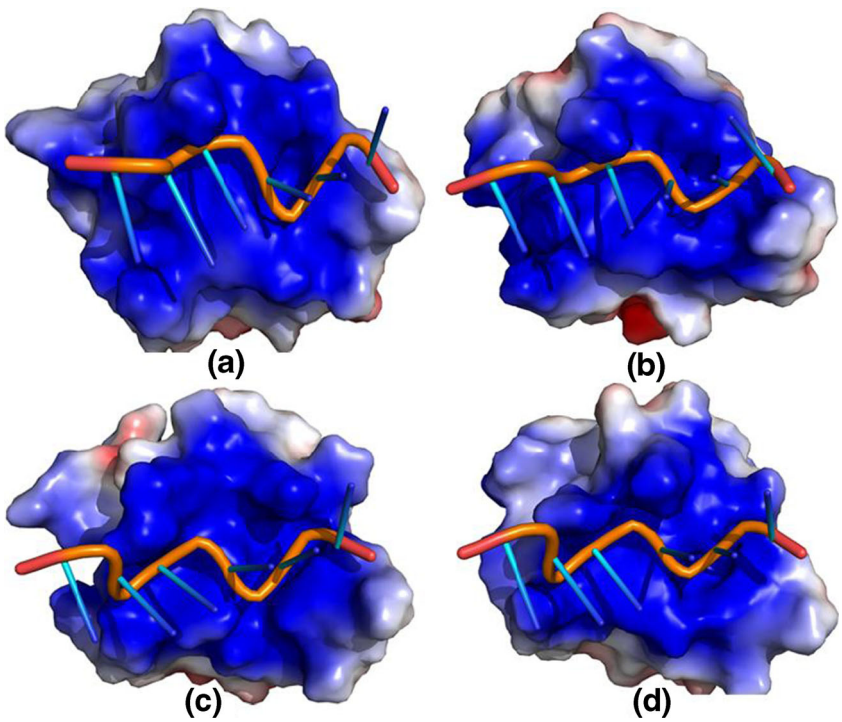

Fig. 6 **A-D** Surface charge distributions of $hZ\alpha_{ADAR1}$ in (**a**) Ori, (**b**) F21, (**c**) F22 and (**d**) F23. Positive (blue) and negative (red) surface potentials are indicated

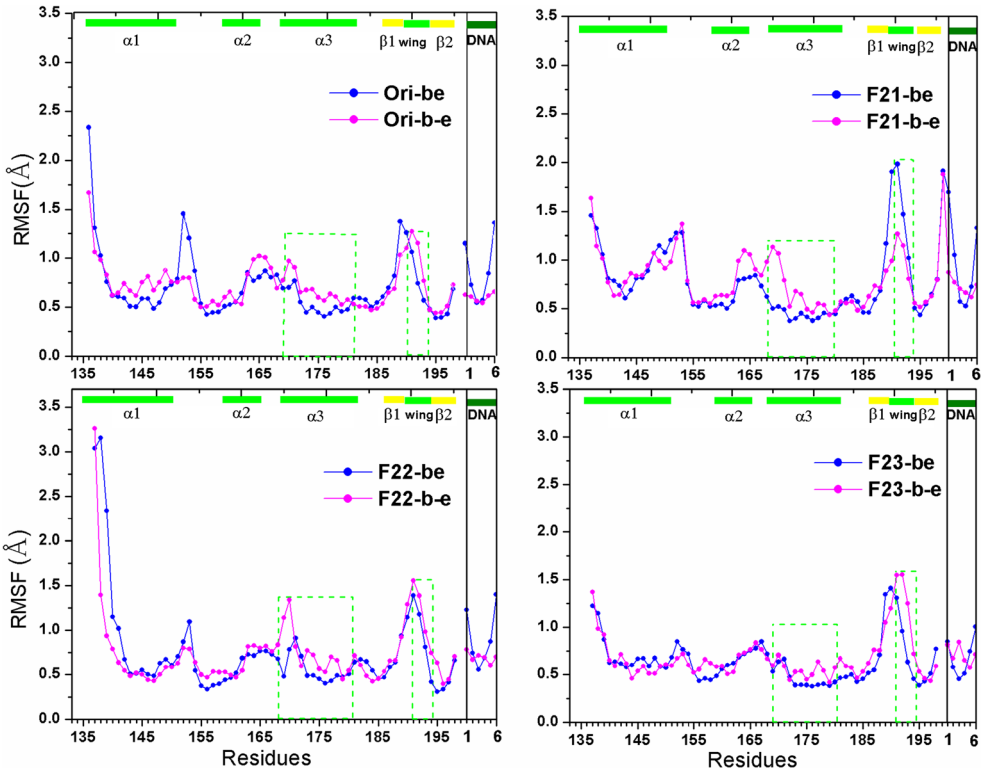

Fig. 7 RMSF values of backbone atoms as a function of residue/nucleotide for hZα$_{ADAR1}$ and Z-DNA in both free and bound states. Helix α3 and wing domains were highlighted by green dashed lines. Herein, we used "Ori-**be**", "F21-**be**", "F22-**be**" and "F23-**be**" to represent hZα$_{ADAR1}$/Z-DNA complex of different systems in bound state, and "Ori-**b-e**", "F21-**b-e**", "F22-**b-e**" and "F23-**b-e**" to represent the combination of the hZα$_{ADAR1}$ and the Z-DNA of different systems in free state

Z-DNA-binding motif are the key interacting domains with Z-DNA, and nonpolar interaction provides the driving force during the recognition process.

The conformation-specific recognition mechanism of Z-DNA and Z-DNA-binding proteins relies on the specific "zig-zag" conformation of Z-DNA. For non-zigzag DNA, most of them are right-handed and have no "zig-zag" conformation like Z-DNA. The recognition of this kind of DNA by proteins is mainly sequence-specific, and takes place primarily in the major groove by the readout mechanism that involves the formation of a series of amino-acid- and base-specific hydrogen bonds [51]. The analysis for protein-DNA

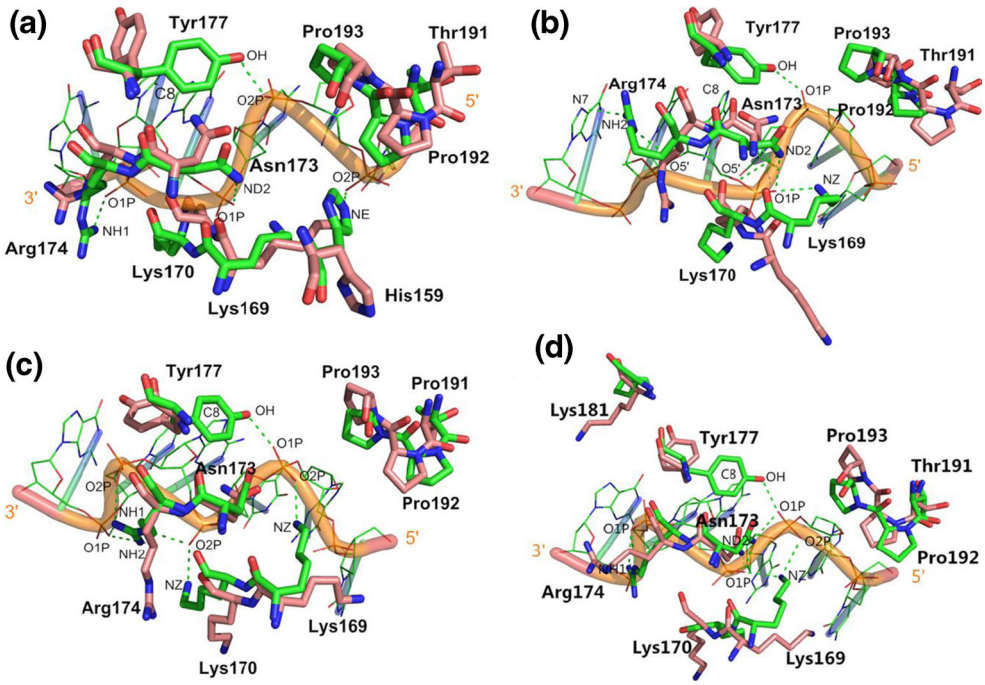

Fig. 8 **A-D** The superposition of apo-hZα$_{ADAR1}$ and Z-DNA-bound hZα$_{ADAR1}$, and the interaction mode hZα$_{ADAR1}$/Z-DNA complex in (**a**) Ori, (**b**) F21, (**c**) F22 and (**d**) F23 system. Pink and green cartoons are indicative of apo-hZα$_{ADAR1}$ and Z-DNA-bound hZα$_{ADAR1}$, respectively. Green dashed lines represent the formed hydrogen bonds of hZα$_{ADAR1}$/Z-DNA complex

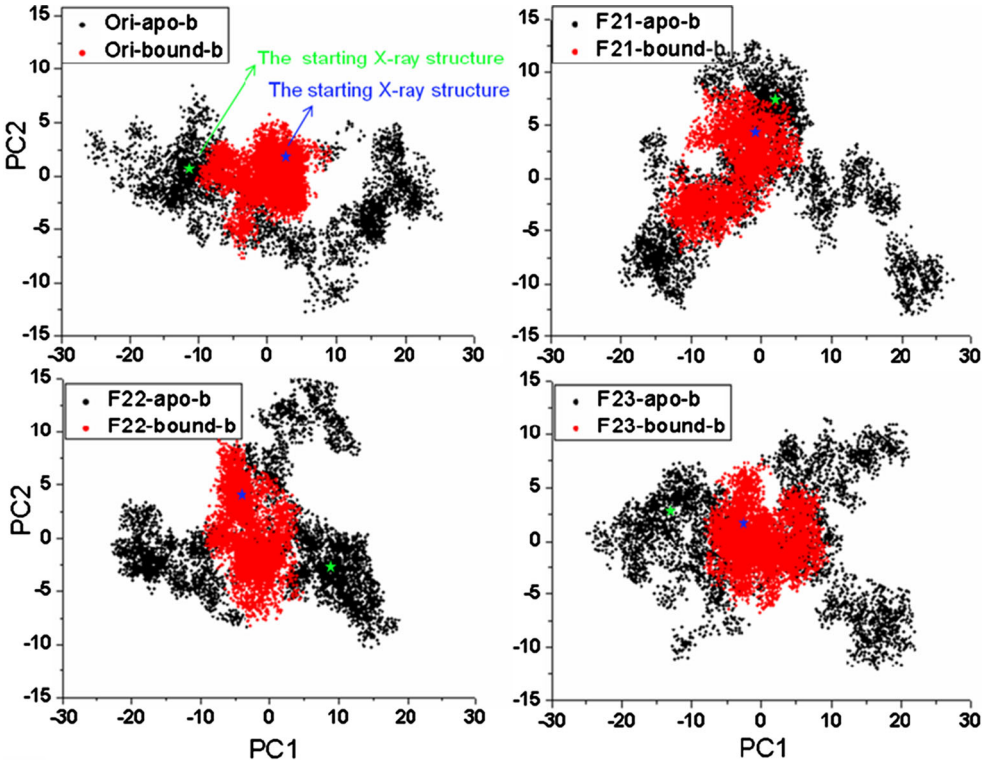

Fig. 9 The principal component analysis of helix α3 and wing domains of hZα$_{ADAR1}$ in free and Z-DNA-bound states. The green and blue stars correspond to the starting structures for the apo- and Z-DNA-bound systems, respectively

complexes shows that the binding of arginine residues to narrow minor grooves is a widely used mode for the protein-DNA recognition, and the ability to detect local variations in DNA shape and electrostatic potential is a general mechanism that enables proteins to use information in the minor groove [52].

Conclusions

In this study, to explore the recognition mechanisms between hZα$_{ADAR1}$ and different Z-DNAs (d(CGCGCG)$_2$, d(CACG TG)$_2$, d (CGTACG)$_2$, and d(CGGCCG)$_2$), three molecular dynamics simulations in each system were performed for apo-hZα$_{ADAR1}$, apo-Z-DNA and hZα$_{ADAR1}$/Z-DNA complex, respectively. The results obtained from MD data show that hZα$_{ADAR1}$ recognized Z-DNAs in different sequence contexts via the common mechanism. By the binding free energy calculation together with energy decomposition, we found that the nonpolar interaction provided the driving force during the binding process and five commonly important residues (Lys169, Lys170, Asn173, Arg174 and Tyr177) were further identified. As the key domains, helix α3 and wing domains interact with Z-DNA mainly by the polar and nonpolar interactions, respectively. Furthermore, conformational analysis and principal component analysis indicate that helix α3 and wing domains of hZα$_{ADAR1}$ were more stable upon the binding of Z-DNA. In general, although the d(CACG TG)$_2$, d (CGTACG)$_2$, and d(CGGCCG)$_2$ sequences are neither the same as the standard Z-DNA nor rich in CG-repeats, the recognitions between hZα$_{ADAR1}$ and Z-DNAs are not influenced and even are conserved. In other words, their recognition mechanism is not sequence-specific but rather conformation-specific. Our work can give a deeper insight into the recognition mechanisms between Z-DNA-binding proteins and various Z-DNAs, and will be useful for better understanding the regulation of biological process.

Acknowledgments This work was supported by the National Natural Science Foundation of China (Grant No 21175063) and the Natural Science Foundation of Gansu Province, China (Grant No: 1208RJYA034).

References

1. Rich A, Nordheim A, Wang AH (1984) The chemistry and biology of left-handed Z-DNA. Annu Rev Biochem 53:791–846
2. Ho PS (1994) The non-B-DNA structure of d (CA/TG)n does not differ from that of Z-DNA. Proc Natl Acad Sci U S A 91:9549–9553
3. Wang G, Vasquez KM (2007) Z-DNA, an active element in the genome. Front Biosci 12:4424–4438
4. Rothenburg S, Koch-Nolte F, Rich A, Haag F (2001) A polymorphic dinucleotide repeat in the rat nucleolin gene forms Z-DNA and inhibits promoter activity. Proc Natl Acad Sci U S A 98:8985–8990
5. Liu H, Mulholland N, Fu H, Zhao K (2006) Cooperative activity of BRG1 and Z-DNA formation in chromatin remodeling. Mol Cell Biol 26:2550–2559
6. Herbert A, Rich A (1996) The biology of left-handed Z-DNA. J Biol Chem 271:11595–11598

7. Herbert A, Rich A (1999) Left-handed Z-DNA: structure and function. Genetica 106:37–47
8. Ha SC, Kim D, Hwang HY, Rich A, Kim YG, Kim KK (2008) The crystal structure of the second Z-DNA binding domain of human DAI (ZBP1) in complex with Z-DNA reveals an unusual binding mode to Z-DNA. Proc Natl Acad Sci U S A 105:20671–20676
9. Ha SC (2004) A poxvirus protein forms a complex with left-handed Z-DNA: crystal structure of a Yatapoxvirus Zalpha bound to DNA. Proc Natl Acad Sci U S A 101:14367–14372
10. Schwartz T, Lowenhaupt K, Heinemann U, Rich A (2001) Structure of the DLM-1–Z-DNA complex reveals a conserved family of Z-DNA-binding proteins. Nat Struct Biol 8:761–765
11. Schade M, Turner CJ, Kuhne R, Schmieder P, Lowenhaupt K, Herbert A, Rich A, Oschkinat H (1999) The solution structure of the Zalpha domain of the human RNA editing enzyme ADAR1 reveals a prepositioned binding surface for Z-DNA. Proc Natl Acad Sci U S A 96:12465–12470
12. Schwartz T (1999) Crystal structure of the Z domain of the human editing enzyme ADAR1 bound to left-handed Z-DNA. Science 284: 1841–1845
13. Herbert A, Schade M, Lowenhaupt K, Alfken J, Schwartz T, Shlyakhtenko LS, Lyubchenko YL, Rich A (1998) The Zalpha domain from human ADAR1 binds to the Z-DNA conformer of many different sequences. Nucleic Acids Res 26:3486–3493
14. Herbert AG, Rich A (1993) A method to identify and characterize Z-DNA binding proteins using a linear oligodeoxynucleotide. Nucleic Acids Res 21:2669–2672
15. Herbert A, Alfken J, Kim YG, Mian IS, Nishikura K, Rich A (1997) A Z-DNA binding domain present in the human editing enzyme, double-stranded RNA adenosine deaminase. Proc Natl Acad Sci U S A 94:8421–8426
16. Kim YG, Lowenhaupt K, Maas S, Herbert A, Schwartz T, Rich A (2000) The zab domain of the human RNA editing enzyme ADAR1 recognizes Z-DNA when surrounded by B-DNA. J Biol Chem 275: 26828–26833
17. Kang YM, Bang J, Lee EH, Ahn HC, Seo YJ, Kim KK, Kim YG, Choi BS, Lee JH (2009) NMR spectroscopic elucidation of the B-Z transition of a DNA double helix induced by the Zalpha domain of human ADAR1. J Am Chem Soc 131:11485–11491
18. Lee YM, Kim HE, Park CJ, Lee AR, Ahn HC, Cho SJ, Choi KH, Choi BS, Lee JH (2012) NMR study on the B-Z junction formation of DNA duplexes induced by Z-DNA binding domain of human ADAR1. J Am Chem Soc 134:5276–5283
19. Lee EH, Seo YJ, Ahn HC, Kang YM, Kim HE, Lee YM, Choi BS, Lee JH (2010) NMR study of hydrogen exchange during the B-Z transition of a DNA duplex induced by the Zalpha domains of yatapoxvirus E3L. FEBS Lett 584:4453–4457
20. Ha SC, Choi J, Hwang HY, Rich A, Kim YG, Kim KK (2008) The structures of non-CG-repeat Z-DNAs co-crystallized with the Z-DNA-binding domain, hZ ADAR1. Nucleic Acids Res 37:629–637
21. Seo YJ, Ahn HC, Lee EH, Bang J, Kang YM, Kim HE, Lee YM, Kim K, Choi BS, Lee JH (2010) Sequence discrimination of the Zalpha domain of human ADAR1 during B-Z transition of DNA duplexes. FEBS Lett 584:4344–4350
22. Pan Y, Nussinov R (2006) Structural Basis for p53 binding-induced DNA bending. J Biol Chem 282:691–699
23. Wan H, Hu JP, Li KS, Tian XH, Chang S (2013) Molecular dynamics simulations of DNA-free and DNA-bound TAL effectors. PLoS One 8:e76045
24. Habtemariam B, Anisimov VM, MacKerell AD Jr (2005) Cooperative binding of DNA and CBFbeta to the Runt domain of the CBFalpha studied via MD simulations. Nucleic Acids Res 33: 4212–4222
25. Case DA, Cheatham TE 3rd, Darden T, Gohlke H, Luo R, Merz KM Jr, Onufriev A, Simmerling C, Wang B, Woods RJ (2005) The Amber biomolecular simulation programs. J Comput Chem 26:1668–1688
26. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. J Chem Phys 79:926–935
27. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters. Proteins 65:712–725
28. Ryckaert JP, Ciccotti G, Berendsen HJC (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J Comput Phys 23:327–341
29. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG (1995) A smooth particle mesh Ewald method. J Chem Phys 103: 8577–8593
30. Shen M, Zhou S, Li Y, Pan P, Zhang L, Hou T (2013) Discovery and optimization of triazine derivatives as ROCK1 inhibitors: molecular docking, molecular dynamics simulations and free energy calculations. Mol BioSyst 9:361–374
31. Hou T, Li N, Li Y, Wang W (2012) Characterization of domain-peptide interaction interface: prediction of SH3 domain-mediated protein-protein interaction network in yeast by generic structure-based models. J Proteome Res 11:2982–2995
32. Liu H, Yao X, Wang C, Han J (2010) In silico identification of the potential drug resistance sites over 2009 influenza A (H1N1) virus neuraminidase. Mol Pharm 7:894–904
33. Yang Y, Liu H, Yao X (2012) Understanding the molecular basis of MK2–p38α signaling complex assembly: insights into protein–protein interaction by molecular dynamics and free energy studies. Mol BioSyst 8:2106
34. Xue W, Qi J, Yang Y, Jin X, Liu H, Yao X (2012) Understanding the effect of drug-resistant mutations of HIV-1 intasome on raltegravir action through molecular modeling study. Mol BioSyst 8:2135
35. Tsui V, Case DA (2000) Theory and applications of the generalized Born solvation model in macromolecular simulations. Biopolymers 56:275–291
36. Hou T, Wang J, Li Y, Wang W (2011) Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. J Chem Inf Model 51:69–82
37. Hou T, Wang J, Li Y, Wang W (2011) Assessing the performance of the molecular mechanics/Poisson Boltzmann surface area and molecular mechanics/generalized Born surface area methods. II. The accuracy of ranking poses generated from docking. J Comput Chem 32:866–877
38. Srinivasan J, Cheatham TE, Cieplak P, Kollman PA, Case DA (1998) Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate−DNA helices. J Am Chem Soc 120:9401–9409
39. Xu L, Sun H, Li Y, Wang J, Hou T (2013) Assessing the performance of MM/PBSA and MM/GBSA methods. 3. The impact of force fields and ligand charge models. J Phys Chem B 117:8408–8421
40. Sitkoff D, Sharp KA, Honig B (1994) Accurate calculation of hydration free energies using macroscopic solvent models. J Phys Chem 98:1978–1988
41. Cui Q, Sulea T, Schrag JD, Munger C, Hung MN, Naim M, Cygler M, Purisima EO (2008) Molecular dynamics-solvated interaction energy studies of protein-protein interactions: the MP1-p14 scaffolding complex. J Mol Biol 379:787–802
42. Gohlke H, Kiel C, Case DA (2003) Insights into protein-protein binding by binding free energy calculation and free energy decomposition for the Ras-Raf and Ras-RalGDS complexes. J Mol Biol 330:891–913
43. Chong LT, Duan Y, Wang L, Massova I, Kollman PA (1999) Molecular dynamics and free-energy calculations applied to affinity maturation in antibody 48G7. Proc Natl Acad Sci U S A 96:14330–14335

44. Shao J, Tanner SW, Thompson N, Cheatham TE (2007) Clustering molecular dynamics trajectories. 1. Characterizing the performance of different clustering algorithms. J Chem Theory Comput 3:2312–2334
45. Lou H, Cukier RI (2006) Molecular dynamics of apo-adenylate kinase: a principal component analysis. J Phys Chem B 110: 12796–12808
46. Prompers JJ, Bruschweiler R (2002) Dynamic and structural analysis of isotropically distributed molecular ensembles. Proteins 46:177–189
47. Wang X, Xu X, Zhu S, Xiao Z, Ma Z, Li Y, Wang Y (2012) Molecular dynamics simulation of conformational heterogeneity in transportin 1. Proteins 80:382–397
48. Hu C, Zhang Y, Huang G, Zhang Q, Gui J (2004) Molecular cloning and characterization of a fish PKR-like gene from cultured CAB cells induced by UV-inactivated virus. Fish Shellfish Immunol 17:353–366
49. Kim YG, Muralinath M, Brandt T, Pearcy M, Hauns K, Lowenhaupt K, Jacobs RL, Rich A (2003) A role for Z-DNA binding in vaccinia virus pathogenesis. Proc Natl Acad Sci U S A 10:6974–6979
50. Kim YG, Lowenhaupt K, Oh DB, Kim KK, Rich A (2004) Evidence that vaccinia virulence factor E3L binds to Z-DNA in vivo: implications for development of a therapy for poxvirus infection. Proc Natl Acad Sci U S A 101:1514–1518
51. Garvie CW, Wolberger C (2001) Recognition of specific DNA sequences. Mol Cell 8:937–946
52. Rohs R, West SM, Sosinsky A, Liu P, Mann RS, Honig B (2009) The role of DNA shape in protein-DNA recognition. Nature 461:1248–1253