4212–4222 *Nucleic Acids Research, 2005, Vol. 33, No. 13*
doi:10.1093/nar/gki724

Cooperative binding of DNA and CBFβ to the Runt domain of the CBFα studied via MD simulations

Bahru Habtemariam, Victor M. Anisimov and Alexander D. MacKerell, Jr*

Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, Baltimore, MD 21201, USA

Received May 27, 2005; Revised and Accepted July 5, 2005

ABSTRACT

The Runt domain (RD) is the DNA-binding region of the Runx genes. A related protein, known as core binding factor β (CBFβ) also binds to the RD to enhance RD–DNA interaction by 6- to 10-fold. Here, we report results from molecular dynamics (MD) simulations of RD alone, as a dimer in complexes with DNA and CBFβ and in a ternary complex with DNA and CBFβ. Consistent with the experimental findings, in the presence of CBFβ the estimated free energy of binding of RD to the DNA is more favorable, which is shown to be due to more favorable intermolecular interactions and desolvation contributions. Also contributing to the enhanced binding are favorable intramolecular interactions between the 'wing' residues (RD residues 139–145) and the 'wing1' residues (RD residues 104–116). The simulation studies also indicate that the RD–CBFβ binding is more favorable in the presence of DNA due to a more favorable RD–CBFβ interaction energy. In addition, it is predicted that long-range interactions involving ionic residues contribute to binding cooperativity. Results from the MD calculations are used to interpret a variety of experimental mutagenesis data. A novel role for RD Glu116 to the RD–CBFβ interaction is predicted.

INTRODUCTION

The Runt domain (RD) is the DNA-binding domain of p53 type transcription factors known as core binding factor α (CBFα) subunits. The CBFαs are expressed by the three related genes Runx1, Runx2 and Runx3 (1). The three Runx genes are important for developmental processes; Runx1 is important for hematopoiesis, Runx2 is essential for skeletal development (osteogenesis) and Runx3 is required for the development of gastric epithelium (2–4). Runx1 and Runx2 also have been shown to be important for vascular development or angiogenesis and their activity is augmented by angiogenic growth factors (5,6). Mutation of the Runx1 gene is associated with human leukemias, whereas mutation of the Runx3 gene is involved with the development and progression of gastric cancers and testicular yolk sac tumor (1,7,8). Mutation of the Runx2 gene is associated with the inherited human skeletal disorder known as cleidocranial dysplasia (9). A related protein, known as CBFβ, binds to the RD of the CBFαs and enhances the DNA-binding affinity of the RD. The RD is the site of both DNA binding and heterodimerization with CBFβ. Because of the involvement of CBFαs in many developmental processes and cancerous diseases, understanding RD–DNA and RD–CBFβ interactions may shed light on the control of these processes. In addition, the involvement of Runx genes with angiogenesis also makes the RD a potential target in the development and treatment of many types of cancers.

The 3D structures of the RD in the free state, bound to DNA, CBFβ and both DNA and CBFβ have been determined using nuclear magnetic resonance (NMR) spectroscopy and/or X-ray crystallography (10–14). Biochemical and biophysical studies have been performed characterizing the binding of RD to DNA in the presence and absence of CBFβ (15,16). In the presence of CBFβ, the dissociation constant (K_D) for RD–DNA decreases 6- to 10-fold (16). Several studies have identified RD residues important for both DNA and CBFβ binding via mutational analysis (17–19). The mutational analysis usually involved mutations to alanine and analysis of its effect on RD–DNA and RD–CBFβ-binding activities, as determined using yeast hybrid assay analysis. In addition, a recent NMR-based study has identified a variety of residues involved in the cooperativity associated with CBFβ binding to RD (20). While these important studies have given insights into the role of individual residues to RD–DNA and RD–CBFβ interactions, the interpretation of the mutational analysis may be complicated by the influence of the mutation on protein stability, protein folding and other localized changes, especially at protein–protein interfaces, while structural and dynamic information from NMR and X-ray studies cannot yield energetic information at an atomic level of detail.

*To whom correspondence should be addressed at 20 Penn Street, Baltimore, MD 21201, USA. Tel: +1 706 410 7442; Fax: +1 410 706 5017; Email: alex@outerbanks.umaryland.edu

© The Author 2005. Published by Oxford University Press. All rights reserved.

The online version of this article has been published under an open access model. Users are entitled to use, reproduce, disseminate, or display the open access version of this article for non-commercial purposes provided that: the original authorship is properly and fully attributed; the Journal and Oxford University Press are attributed as the original place of publication with the correct citation details given; if an article is subsequently reproduced or disseminated not in its entirety but only in part or as a derivative work this must be clearly indicated. For commercial re-use, please contact journals.permissions@oupjournals.org

Computational approaches can yield important details about the interaction of RD with DNA and CBFβ (21). For example, the energetic contribution of individual residues to both DNA and CBFβ binding can be determined. By doing such analysis in systems with all wild-type residues present, the risk of mutation induced conformational change, including changes in protein stability, is avoided. Computational analysis can also account for the contribution of all atoms in a given residue, including the backbone atoms, whereas mutational analysis can only detect the importance of side chain atoms. Thus, computational analysis can fill gaps in understanding the RD–DNA, RD–CBFβ and CBFβ–RD–DNA ternary complex interactions.

In the present work, the RD was subjected to molecular dynamic (MD) simulations in the free state and in several complexes. Complexes include the binary RD–DNA and RD–CBFβ complexes, and the CBFβ–RD–DNA ternary complex. In addition, two DNA simulations and a CBFβ monomer simulation were performed. From the MD simulations, in combination with free energy component analysis (22–24) based on the Generalized-Born model (25,26), overall binding energies, desolvation effects and detailed interaction energies were calculated. These results, along with previous structural analysis and observations, were used to interpret a variety of experimental data involving RD–DNA and RD–CBFβ interactions (16–19,27).

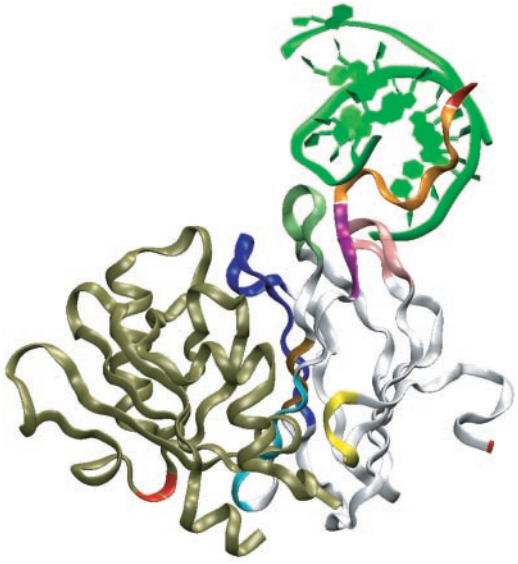

Figure 1. Ribbon image of the CBFβ–RD–DNA after 4.5 ns dynamics simulation including coloring of the RD segments used for analysis (Table 2). DNA (green), CBFβ (tan) and residues surrounding omitted residues 71–79 in CBFβ (red).

METHODS

MD simulations and all calculations were performed using the program CHARMM (28). The all-hydrogen CHARMM22 protein (29) and CHARMM27 nucleic acid (30,31), the TIP3P water model (32), along with published sodium and chloride parameters were used (33). Coordinates from the CBFβ–RD–DNA ternary complex X-ray crystal structure (PDB accession code 1H9D) (13) were used as starting structures of the seven simulation systems (Supplementary Table S1). The only exception was the second DNA simulation, which was initiated with the DNA in the canonical B form (34), as obtained from the program Quanta (Accelrys Inc.). In the crystal structure terminal residues 50–53 and 179–183 in RD are not observed experimentally as are residues 71–79 of CBFβ. Accordingly, in the simulations these residues were omitted. Omission of residues 71–79 of CBFβ, which are spatially removed from the RD in the experimental structure (Figure 1), was performed by creating a peptide bond between residues 70 and 80 and relaxing the structure during the minimization and MD simulation protocols described below. Root-mean-square (RMS) difference analysis of the residues adjacent to the deleted residues in CBFβ (i.e. residues 66–70 and 80–84) showed significant structural changes to occur only in the one or two residues directly adjacent to the deleted residues (Supplementary Table S2).

Each structure to be simulated was generated with CHARMM and the hydrogens added. The dimensions of each complex (solute) were determined, and a box of solvent was prepared to overlay the solute. The edge of the solvent was extended to a minimum of 8 Å beyond the solute in each dimension. Then each solute was overlaid with the solvent water box that also contained either sodium or chloride ions, depending on the total charge of the solute. All water molecules whose oxygen atom was within 2.5 Å of any solute non-hydrogen atom were then deleted and excess ions were deleted to make the systems electrically neutral. Details of the systems are shown in Supplementary Table S1. All subsequent calculations were performed using periodic boundary conditions, with images generated using the CRYSTAL module of CHARMM.

Equilibration of each simulation system was carried out by performing 500 steps of steepest descent energy minimization followed by a 10 ps MD NVT simulation with harmonic restraints of 5 kcal/(mol Å) on the solute non-hydrogen atoms. Additional minimization using the Adopted Basis Newton–Raphson (ABNR) algorithm was performed for 100 steps without any restraints. The systems were then subjected to a 5 ns MD simulation in the NPT ensemble at 300 K with the Leap Frog integrator, and a time step of 0.002 ps. Pressure control was performed with the Langevin Piston algorithm (35). SHAKE was used to constrain covalent bonds involving hydrogens (36). Particle Mesh Ewald (PME) was used to treat long-range electrostatic interactions (37). PME calculations were carried out using real space electrostatic and Lennard–Jones interaction cutoffs at 10 Å, with utilization of heuristic non-bond list update out to 12 Å and smoothing of the LJ interactions from 8 to 10 Å using the force switch method (38). The first 500 ps of the simulations were treated as equilibration. Coordinates were saved every 5 ps, and by the end of the 4.5 ns production MD simulation, 900 sets of structural coordinates were available for analysis. Structural figures were generated using the program VMD (39) and rendered by Raster3D (40).

Calculation of binding free energies via free energy component analysis: binding free energies, ΔG_{bind}, were calculated as follows:

$$\Delta G_{\text{bind}} = G_{\text{complex}} - G_{\text{partnerA}} - G_{\text{partnerB}}, \quad \mathbf{1}$$

where the overall free energy of each complex was described using (24)

$$G = \mathrm{E_{MM}} + G_{\mathrm{Solvation}} - TS_{\mathrm{MM}}. \quad \mathbf{2}$$

In Equation 2 E_{MM} describes the molecular mechanical energy of the system consisting of all components in the CHARMM potential energy function, as shown in the following Equation 3:

$$E_{\mathrm{MM}} = E_{\mathrm{bond}} + E_{\mathrm{angle}} + E_{\mathrm{Urey-Bradley}} + E_{\mathrm{dihedral}} + E_{\mathrm{improper}} + E_{\mathrm{vdW}} + E_{\mathrm{elec}}. \quad \mathbf{3}$$

For complexes, E_{MM}, can be separated into interaction energy, E_{IE} and strain, E_{strain} components:

$$E_{\mathrm{MM,\,complex}} = E_{\mathrm{IE}} + E_{\mathrm{strain}}, \quad \mathbf{4}$$

where E_{IE} is the sum of the vdW and electrostatic interaction energy between individual molecules in a complex and E_{strain} is the sum of E_{MM} terms for each molecule in the complex (i.e. all MM energy contributions for each species ignoring contributions from the other molecules in the complex). The $G_{\mathrm{solvation}}$ term describes the solvation free energy of a system and includes both electrostatic, G_{elec} and non-polar, G_{np}, components.

$$G_{\mathrm{solvation}} = G_{\mathrm{electrostatic}} + G_{\mathrm{np}}. \quad \mathbf{5}$$

The electrostatic portion of the solvation free energy was calculated using the Generalized-Born molecular volume (GBMV) method (25,26). The non-polar portion of the solvation energy (i.e. due to cavity formation and hydrophobicity) was calculated using the Still equation, $E_{\mathrm{np}} = \gamma \mathrm{SA}$, where γ is an empirical atomic solvation parameter, 7.2 cal/Å^2, and SA is the solvent accessible surface area calculated with a solvent probe radius of 1.4 Å (25,41). The final term in Equation 2, $\mathrm{TS_{MM}}$, describes the entropic part of the energy as a sum of translational, rotational and vibrational entropy terms calculated via the harmonic approximation (42) for the temperature 300 K. Vibrational analysis was performed on the solute molecules extracted out of the solvent box. The solute structure was minimized using the Conjugate-Gradient method to an RMS gradient of 10^{-4} kcal/(mol Å) with a distance-dependent dielectric of $4r$ applied to approximate the screening of electrostatic interactions by solvent, where r is the distance between two atoms, as performed previously (22).

To better understand the contributions of the various energetic terms to complex formation, differences in the energy terms upon formation of the complexes were calculated. This typically involved calculation of the difference between a selected term for the complex and those of the individual molecules comprising the complex. In the case of the interaction energies their contribution was simply the sum of the contribution between the individual molecules in each complex (i.e. for the CBFβ–RD dimer the E_{IE} between CBFβ and RD was included). However, in the case of the trimer formation via dimer + monomer complexation the E_{IE} contribution includes those between the monomer being added and the two molecules in the original dimer along with the change in the interaction energy between the molecules in the original dimer upon going to a trimer. For example, $E_{\mathrm{IE,CBF\beta\text{-}RD\,+\,DNA}} = E_{\mathrm{IE,CBF\beta\text{-}DNA}} + E_{\mathrm{IE,RD\text{-}DNA}} - \Delta E_{\mathrm{IE,CBF\beta\text{-}RD(dimer\ to\ trimer)}}$. This term accounts for gains or losses in the interaction energy of the molecules in a dimer when the third molecule binds to form a trimer.

Interaction energies, E_{inter}, and strain energies, E_{strain}, were calculated for each individual species using an infinite cutoff in order to mimic the long-range electrostatic interactions treated via PME in the MD simulations. Energy calculations were performed on the 900 structures from each simulation system with statistical analysis performed by dividing each 4.5 ns simulation into nine 500 ps time blocks. Statistical analysis was performed by obtaining the averages from each time block, with those averages treated as independent data points, allowing calculation of the overall averages, and standard errors as described previously (43). The entropy calculations were performed on 30 structures extracted from the simulation trajectories at 150 ps time intervals, with the statistical analysis derived by dividing the 30 observations into ten 450 ps independent time blocks from which the averages and the standard errors were obtained. Overall RMS differences of all simulated structures versus the crystallographic structure of the trimer are shown in Supplementary Table S3.

RESULTS AND DISCUSSION

MD simulation studies were undertaken on the RD to understand the nature of interactions with the DNA and with the binding partner, CBFβ. Average RMS differences for all non-hydrogen atoms with respect to the crystallographic ternary complex structures were 1.62 ± 0.04, 3.56 ± 0.10, 2.15 ± 0.04, 3.23 ± 0.16, 3.63 ± 0.11 and 2.81 ± 0.05 Å for the DNA, RD, CBFβ, RD–DNA, RD–CBFβ and CBFβ–RD–DNA structures, respectively, indicating that the overall structures were maintained in the simulations. The structure of the trimer after a 4.5 ns simulation is illustrated in Figure 1. Analysis focused on the overall binding as calculated using Equation 1 followed by investigation of the contributions of different components to the overall binding (Equations 2 and 4). This was followed by identifying contributions of different regions of the proteins and DNA to binding, ultimately analyzing the role of individual amino acids. Emphasis was placed on using the present results to interpret the body of experimental data available on RD.

Free energies of complex formation for the dimers and the trimer were calculated using Equation 1 based on the data presented in Supplementary Table S4. The resulting free energies of binding for the various complexes are presented in Table 1 and are included in the thermodynamic cycle shown in Figure 2. Two free energies of binding are reported, ΔG_{bind} and $\Delta G_{\mathrm{bind,TS}}$, that exclude and include the configuration entropy, $\mathrm{TS_{mm}}$, respectively. For ΔG_{bind} the calculated energies are all favorable, although the interactions involving addition of CBFβ are extremely exothermic. In contrast, upon inclusion of the entropy the events that involve DNA binding become unfavorable. In addition, the calculations predict that RD–CBFβ binding is more favorable than RD–DNA binding, in contrast to experimental observations (16). The presence of these discrepancies is not surprising due to a variety of assumptions in the calculations, including the omission of energetic contributions, such as the solvation and entropic contributions of the added counterions (23), limitations in

Table 1. Calculated binding energies and component contributions

Binding step	$\Delta G_{solvation}$	ΔE_{strain}	ΔE_{IE}	ΔG_{bind}[a]	ΔTS_{MM}	$\Delta G_{bind,TS}$[b]
Dimer formation						
RD + DNA	2403.1	109.0	−2513.4	−0.2	−49.4	49.2
CBFβ + RD	262.0	−104.7	−238.7	−81.0	−48.5	−32.5
Trimer formation						
RD–DNA + CBFβ	201.0	75.5	−363.0	−87.0	−52.0	−35.0
RD–CBFβ + DNA	2342.1	289.2	−2637.7	−6.3	−52.9	46.6
Trimer total	2604.1	184.5	−2876.4	−87.3	−101.4	14.1
Δ Trimer–dimer	−61.0	180.2	−124.4	−6.0	−3.5	−2.5

Energies in kcal/mol calculated using Equations 1–4 from the data reported in Supplementary Table S4.
[a]Free energy of binding based on the differences in the free energy of each monomer or complex, *G*, in Supplementary Table S4. It should be noted that the sum of the terms $\Delta G_{solvation}$, ΔE_{strain} and ΔE_{IE} are not exactly equivalent to ΔG_{bind} due to rounding errors. Based on the sums, free energies of binding of −1.3, −81.4, −86.5, −6.4, −87.8 and −5.2 kcal/mol were obtained for the RD + DNA, CBFβ + RD, RD–DNA + CBFβ, RD–CBFβ + DNA, trimer total and Δ trimer–dimer, respectively.
[b]Free energy of binding including the molecular mechanical entropic contribution, ΔTS_{MM}.

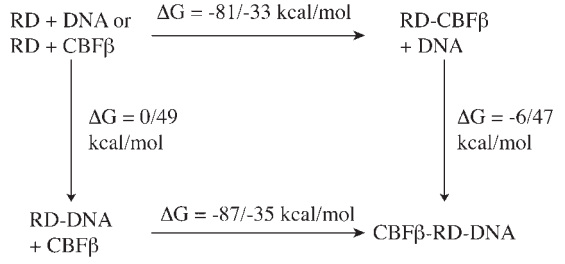

Figure 2. Overall binding pathways RD–DNA, RD–CBFβ and RD–CBFβ–DNA. Free energies before and after the/represent ΔG_{bind} and $\Delta G_{bind,TS}$, as reported in Table 1.

the potential energy function and the force field (44) along with possible limitations in the sampling of conformational space. For example, underestimation of the protein–DNA interaction energies in the force field by <4% would account for the prediction that binding of RD to CBFβ is more favorable than RD binding to DNA. With respect to the TS_{MM}, the use of gas phase minimized structures may be problematic, especially in the case of the polyanionic DNA where the electrostatic screening due to the 4*r* dielectric is expected to be insufficient when compared with the proteins. However, when the two types of binding free energy are compared, the trends are similar. Moreover, in both scenarios, the binding of DNA to the RD–CBFβ dimer to form the trimer is more favorable than the binding of RD and DNA to form the RD–DNA dimer, consistent with the experimental data (16,18). Thus, it appears that the present results, although not in quantitative agreement with experiment due to assumptions and limitations in the theoretical models, are representative of the experimental regimen, allowing for atomic details of events driving binding to be elucidated. In addition, the emphasis in this work on differences in energies and structures between the systems rather than absolute energies should allow for increased confidence in the observations and the conclusions based on those observations. Concerning the convergence of the present results, the energies calculated from the two independent DNA simulations are within the estimated errors (Supplementary Table S4), suggesting that the present calculations have adequately converged.

Analysis of the thermodynamic cycle in Figure 2 predicts that binding of both CBFβ and DNA are cooperative. These results are consistent with previous studies showing the binding of DNA to RD to be enhanced by the binding of CBFβ as well as the binding of CBFβ to RD to be enhanced by DNA based on electrophoretic mobility shift assay (10,16,18). As discussed, such cooperativity may be important for maintaining binding of RD to DNA via stabilization of the ternary complex to insure adequate transcriptional activation. Tighter protein–protein interactions due to DNA binding have been observed in other types of transcription factors (45).

Contributions from solvation, strain and interaction energies to the calculated binding energies are shown in Table 1. As expected, the solvation energies are all unfavorable upon binding, with the unfavorable contributions being significantly larger for the interactions involving the highly solvated polyanionic DNA. The strain energies are all also unfavorable with the exception of formation of the CBFβ–RD dimer. Typically, it would be expected that the protein–protein or protein–DNA interactions would lead to an unfavorable strain contribution upon binding (23). Ongoing studies in our laboratory are investigating the unusual behavior in the CBFβ–RD dimer interaction. Countering the unfavorable solvation and strain contributions are favorable interaction energies for all cases. Clearly, the free energies of binding are driven by the large, opposing forces of solvation and, typically, strain energies versus interaction energies between the macromolecules in the complexes.

Data at the bottom of Table 1 addresses the contributions to the calculated gain in binding energy associated with cooperativity upon going from the dimers to the trimer. Contributing to the cooperativity is a decrease in the unfavorable solvation that occurs upon binding, evidenced by the Δtrimer–dimer value for $\Delta G_{solvation}$ being −61 kcal/mol. In addition, the interaction energies become more favorable by −124 kcal/mol. The more favorable change in interaction energy occurred with both the RD–DNA and the RD–CBFβ interactions in the trimer (Table 1); details of the contributions of different regions of the proteins to this will be discussed below. Countering these contributions are unfavorable increases in the strain energy, consistent with gains in interaction energies, as discussed above. Thus, the cooperativity of DNA and CBFβ binding to RD is predicted to be dominated by gains in the favorable interactions between the macromolecules and with lowering of the solvation penalty to complex formation.

A more detailed understanding of the atomic determinants of cooperativity associated with RD binding to CBFβ and DNA was obtained via interaction energy partitioning analysis between different regions of the RD and both CBFβ and DNA. This was facilitated by partitioning RD into a variety of segments as specified in Table 2 and shown in Figure 1. Several regions of the protein have been previously indicated to contribute to either RD–DNA or RD–CBFβ interactions (13). Residues adjacent to the respective binding partner but not included in a defined segment are referred to as Other-DNA and Other-CBFβ and all residues not previously defined are referred to as non-interacting-CBFβ, allowing for the quantification of their contributions to binding. Of note are the Wing residues, which are primarily important for DNA binding, while residues in Wing1 are proximal to the Wing residues, contacting both CBFβ and the Wing residues. For all these segments of RD, interaction energies were calculated between the respective monomers in the dimers as well as in the trimer.

As discussed above, cooperativity exists between both DNA and CBFβ with respect to their binding to RD. Table 3 includes the change in RD–DNA interaction energies in both the trimer and the dimer complexes along with the change in interaction energy upon going from the RD–DNA dimer to the trimer, with those contributions broken down into the segments presented in Table 2. Upon trimer formation the total RD–DNA interaction energy becomes more favorable, with three of the segments contributing significantly to the more favorable interaction energy in the trimer. These include Wing, βA loop and Other-DNA segments. The largest contribution is from the Wing segment, consistent with its location in the vicinity of both the DNA and the CBFβ. This location allows for interactions of CBFβ to be communicated to Wing via Wing1 (see below for additional details), as discussed previously (13). Interestingly, the Other-DNA region makes a significant contribution to cooperativity by its unfavorable contribution in the RD–DNA dimer becoming less unfavorable in the trimer. Such a change indicates that the binding of CBFβ leads to a conformational change in the core region of RD that favors interactions with the DNA and, due to the Other-DNA region not being in direct contact with the DNA, suggests that long-range interactions contribute to binding cooperativity.

To better understand the role of the Wing segment in facilitation of DNA binding upon trimer formation, the contribution of this region was broken down into a per residue basis (Table 4). Overall, the majority of the favorable interactions are made by Arg139, Arg142 and Lys144, as expected due to their positive charge and the polyanionic nature of DNA. With all three residues their interaction energy contribution becomes more favorable upon trimer formation. In addition, more favorable interactions occur with Ser140, Gly143 and Ser145, with the contribution of Gly143 being relatively large. Thus, specific residues in the Wing contribute to more favorable interactions with the DNA upon formation of the ternary complex. Details of the role of CBFβ and the Wing1 segment to these contributions will be discussed below.

Similar to RD–DNA binding, the RD to CBFβ interaction energy is increased upon trimer formation. Table 5 presents the energy breakdown into the different RD segment contributions. The overall cooperativity gain of −60 kcal/mol is due to the Wing1 and N-terminal segments. The contribution of the Wing segment to CBFβ binding becomes less favorable whereas the contribution of this segment becomes more favorable to DNA binding upon trimerization (Table 3). Such a scenario suggests a subtle balance between interactions of the Wing and Wing1 segments that contribute to the cooperative binding. Interestingly, as with the Other-DNA segment discussed above for RD–DNA binding, the N-terminal segment makes a favorable contribution to the cooperativity of

Table 2. Definitions of regions of the RD interacting with DNA or CBFβ

Segment name	Residues	Color[a]
DNA binding regions		
Wing	139–145	Lime
βA loop	77–84	Pink
βG strand	166–169	Purple
Tail	170–177	Orange
Other-DNA	54–76, 85–138, 146–165, 178	White
CBFβ binding regions		
Wing1	104–116	Blue
Area1/2	156–161	Cyan
N terminus	66–69	Yellow
Other-CBFβ	149–151	Brown
Non-interacting-CBFβ	All remaining RD residues not interacting with CBFβ	White

[a]Color as shown in Figure 1. Red indicates residues adjacent to those omitted from the MD simulations.

Table 3. RD region contributions to the RD–DNA interaction energy

Region	CBFβ–RD–DNA	Complex RD–DNA	ΔIE
Wing	−1158.1 ± 9.9	−1091.1 ± 10.1	−67.0
βA Loop	−899.7 ± 2.5	−888.2 ± 5.3	−11.5
Tail	−206.9 ± 8.7	−220.6 ± 5.8	13.7
βG strand	−418.1 ± 0.8	−416.3 ± 4.8	−1.8
Other-DNA	59.9 ± 10.3	102.8 ± 20.9	−42.9
Total	−2622.9 ± 13.4	−2513.4 ± 11.9	−109.5

Energies in kcal/mol as mean ± SE, ΔIE = CBFβ–RD–DNA − RD–DNA.

Table 4. Interaction energy of RD Wing residues with DNA

RD residues	CBFβ–RD–DNA	Complex RD–DNA	ΔIE
R139	−452.4 ± 3.1	−428.1 ± 6.3	−24.3
S140	−2.2 ± 0.2	0.8 ± 0.6	−3.0
G141	−0.6 ± 0.4	−7.6 ± 1.8	7.0
R142	−432.4 ± 8.6	−404.7 ± 8.8	−27.7
G143	−19.7 ± 0.3	−8.4 ± 1.7	−11.3
K144	−248.8 ± 1.4	−242.6 ± 2.0	−6.2
S145	−2.0 ± 1.1	−0.4 ± 1.1	−1.6

Energies in kcal/mol as mean ± SE, ΔIE = CBFβ–RD–DNA − RD–DNA.

Table 5. RD region contributions to the RD–CBFβ interaction energy

Region	CBFβ–RD–DNA	Complex CBFβ–RD	ΔIE
Wing	0.0 ± 2.2	−12.1 ± 2.6	12.1
Wing1	−114.4 ± 8.9	−70.9 ± 7.8	−43.5
Area2	−65.2 ± 1.4	−66.6 ± 1.3	1.4
N-terminus	−77.5 ± 4.8	−29.4 ± 1.7	−48.1
Other-CBFβ	−12.2 ± 0.4	−13.3 ± 0.2	1.1
Non-interacting CBFβ	−29.5 ± 3.8	−46.4 ± 7.6	16.9
Total	−298.7 ± 0.0	−238.7 ± 2.9	−60.0

Energies in kcal/mol as mean ± SE, ΔIE = CBFβ–RD–DNA − CBFβ–RD.

RD–CBFβ binding although it is spatially remote from the DNA in the ternary complex. This indicates that binding of DNA to form the trimer leads to global conformational changes in RD that favor RD–CBFβ interactions, further indicating the contribution of long-range interactions to cooperativity.

Residue breakdown of the interaction energy contributions of the Wing1 and N-terminal regions to CBFβ–RD interactions is presented in Table 6. The energetic contribution of a number of the residues changes significantly upon going from the CBFβ–RD dimer to the trimer, with both favorable and unfavorable contributions occurring. The most significant contributions, both favorable, occur with Glu116 in Wing1 and Asp66 in the N-terminal segment. The importance of Glu116 for the binding of CBFβ has not previously been described, although perturbed chemical shifts of Glu116 have been observed in NMR studies in the presence of either CBFβ or DNA (13,17,46) and a significant decrease in the order parameter of Glu116 has been observed upon going from the RD–DNA dimer to the trimer (20). The interaction energies between RD Glu116 and CBFβ in the binary and ternary complexes are −20 ± 8 and −67 ± 9 kcal/mol, respectively. In the present study, it was found that RD Glu116 interacts with CBFβ residue Arg33. Experimentally, mutation of CBFβ Arg33 to alanine increased the RD–CBFβ dissociation constant in the ternary complex by 20-fold (18). This finding is consistent with the present result. Interestingly, the experimental change in K_D was not accounted for in terms of any residue(s) in the RD. To understand this omission, distances between the two residues in the binary and ternary complexes as a function of time were obtained (Supplementary Figure S1). It is observed that in the ternary complex during the MD simulation, the sidechains of RD Glu116 and CBFβ Arg33 move to within 2–3 Å, in contrast to the OE2-NH2 distance being 9.4 Å in the crystallographic structure of the ternary complex (13). In contrast, in the binary complex crystal structure the OE2-NH2 distance is 12 Å; this distance initially decreases to shorter values during the MD simulation, but upon further simulation time assumes a stable conformation with a distance of ∼12 Å between the OE2-NH2 atoms.

To better understand the significant changes in the RD Glu116 and CBFβ Arg33 distances for the different systems the experimental structure along with final time frames from the 4.5 ns RD–CBFβ binary and ternary simulations were obtained (Figure 3). In the RD–CBFβ dimer, RD Glu116 and CBFβ Arg33 are relatively far apart (Figure 3A). Upon going to the ternary complex in the MD simulation it is clear that a salt bridge forms between Glu116 and Arg33 (Figure 3B). The formation of this salt bridge is associated with a shift in the position of the Arg33 sidechain towards the RD, which is suggested to be associated with the electrostatic field due to the presence of DNA in the ternary complex as evidenced by the favorable interactions between the CBFβ

Table 6. Interaction energy of selected RD residues with CBFβ

RD residue	CBFβ–RD–DNA	CBFβ–RD	ΔIE
Wing1			
T104	−1.2 ± 0.1	−3.3 ± 0.8	2.1
V105	0.1 ± 0.1	0.7 ± 0.1	−0.6
M106	−6.9 ± 0.6	−9.5 ± 0.2	2.6
A107	−2.1 ± 0.3	−1.2 ± 0.1	−0.9
G108	−0.8 ± 0.3	−1.2 ± 0.1	0.4
N109	−5.4 ± 0.3	−4.7 ± 0.2	−0.7
D110	3.2 ± 0.6	8.2 ± 1.0	−5.0
E111	−3.4 ± 1.3	1.2 ± 0.7	−4.6
N112	−5.2 ± 1.4	−0.7 ± 0.6	−4.5
Y113	−14.1 ± 0.9	−27.6 ± 1.1	13.5
S114	−14.8 ± 1.7	−14.0 ± 0.3	−0.8
A115	2.7 ± 0.9	1.5 ± 0.1	1.2
E116	−66.5 ± 8.8	−20.4 ± 7.8	−46.1
N-Terminus			
D66	−44.9 ± 4.4	1.0 ± 1.1	−45.9
S67	−9.9 ± 0.4	−7.8 ± 0.5	−2.1
P68	−13.5 ± 0.3	−12.8 ± 0.9	−0.7
N69	−9.3 ± 0.2	−9.9 ± 0.3	0.6
Area2			
T154	0.6 ± 0.4	−0.1 ± 0.4	0.7
N155	−4.6 ± 1.1	−5.3 ± 0.4	0.7
P156	−11.1 ± 0.4	−10.1 ± 0.6	−1.0
P157	−12.1 ± 0.2	−12.0 ± 0.1	−0.1
Q168	−9.2 ± 0.5	−10.4 ± 1.0	1.2
V169	−16.3 ± 0.1	−15.7 ± 0.1	−0.6
A160	−4.2 ± 0.1	−4.2 ± 0.1	0.0
T161	−8.3 ± 0.2	−9.0 ± 0.1	0.7

Energies in kcal/mol as mean ± SE, ΔIE = CBFβ–RD–DNA − CBFβ–RD.

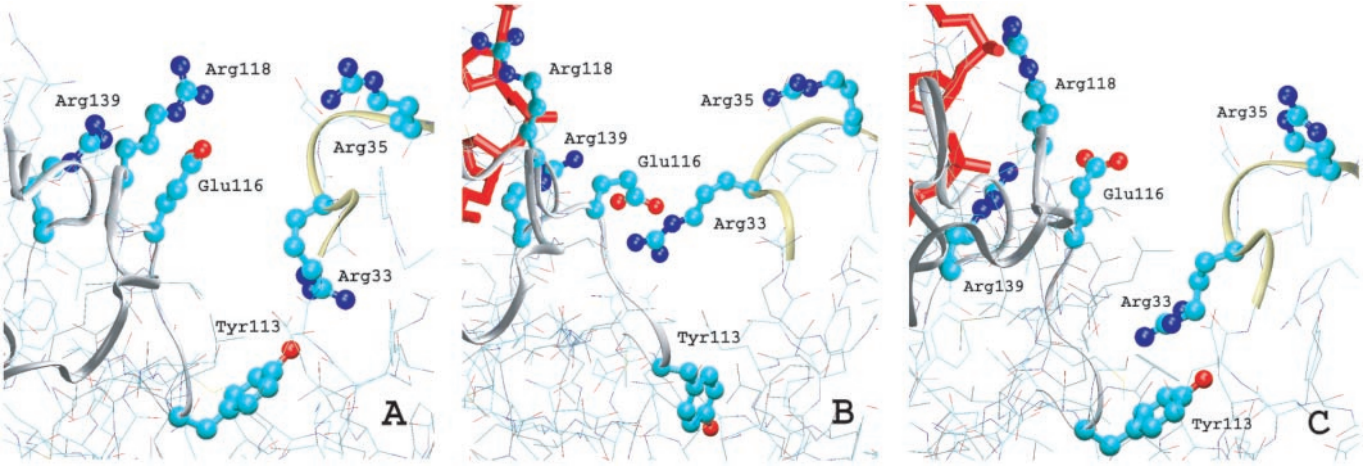

Figure 3. RD–CBFβ interactions involving RD residues Tyr113, Glu116, Arg118 and Arg139, and CBFβ residues Arg33 and Arg35. Images from snapshots at 4.5 ns from the (**A**) RD–CBFβ binary and (**B**) ternary MD simulations and (**C**) from the X-ray crystal structure. DNA is shown in red.

Arg33 and the DNA (−252.2 ± 6.1 kcal/mol). In addition, energy partitioning analysis shows that Arg33 of CBFβ overall disfavors RD–CBFβ binding in the trimer with this effect significantly increased in the RD–CBFβ dimer by ∼40 kcal/mol [Arg33-RD(trimer) = 11.0 ± 4.9 and Arg33-RD(dimer) = 50.3 ± 3.3], further disfavoring the RD Glu116 to CBFβ Arg33 interaction in the dimer. However, the question remains why is the RD Glu116 and CBFβ Arg33 interaction not observed in the ternary complex crystal structure (Figure 3C). Although speculative, it may be due to the presence of higher salt concentrations in the experimental study (200 mM NaCl) (13), which may partially screen the Arg33–DNA interactions. This screening combined with competing interactions involving other residues, such as the RD Tyr113 to CBFβ Arg33 interaction that occurs in the crystal and RD–CBFβ dimer MD structures, may lead to the subtle difference between the two systems. Supporting this possibility is the NMR observed decrease in the Glu116 order parameter in the ternary complex (20), suggesting that in solution and at lower salt concentration the proposed salt bridge may be present. Other experimental results indicate a 5-fold increase in the dissociation constant upon formation of the ternary complex from the RD–DNA dimer (18), which may be related to Tyr113 properly orienting Arg33 for the actual binding event. Regardless, the RD Glu116 to CBFβ Arg33 salt bridge along with the additional interactions leads to the more favorable interaction energy between CBFβ and RD in the ternary versus the binary complex (Table 3). Thus, it is predicted that Glu116 of RD has a significant impact on the binding of CBFβ to RD as well as on the cooperativity of RD binding to DNA and CBFβ. Notably, binding of CBFβ to the RD–DNA dimer has been shown via NMR to decrease the conformational variability of Glu116 (20), supporting the impact of RD interactions with its binding partners on this residue.

The interaction energies for Met106, Ala107, Asn109, Asn112, Tyr113 and Ser114 (Table 6) are consistent with the experimental mutation data. Mutation of these residues decreases RD–CBFβ binding in the experimental study (17,18), consistent with the favorable interaction energies of these residues with CBFβ. However, mutation of Asn109 to alanine decreased the RD–CBFβ binding by 60-fold while the calculated interaction energy is relatively small. This discrepancy may be due to mutation induced structural perturbations of the RD (i.e. partial unfolding), leading to the experimentally observed decrease in binding (18). Recently, it has been shown that polar residues at protein–protein interfaces provide structural rigidity so as to maintain specificity and to reduce the entropic cost upon binding (47). Thus, Asn109 may play an important role in maintenance of the structural integrity of RD with a modest contribution to affinity associated with direct interactions with CBFβ.

The N-terminal segment makes a significant, favorable gain in interaction energy with CBFβ upon going from the RD–CBFβ binary to the ternary complex (Table 5). As shown in Table 6, the contribution was dominated by Asp66. However, experimental mutation of Asp66 to alanine only moderately affected the RD–CBFβ binding strength (19). Visual inspection of the binary MD, ternary MD and crystal structures (Figure 4), along with energy partitioning analysis shows Asp66 to interact favorably, via its backbone carbonyl, with Lys11 of CBFβ, with this interaction becoming a salt bridge involving the sidechains in the trimer. The difference between the RD–CBFβ dimer MD and crystal structures and the trimer MD structure appears to be due to a competitive interaction between Asp66 with His163 in RD (Figure 4A and C, respectively). Consistent with the results discussed above, this suggests that the electrostatic field imposed by the polyanionic DNA in the ternary simulations disrupts the Asp66-His163 interaction seen in the dimer and helps in orienting Asp66 toward Lys11 (Figure 4C), with the presence of 200 mM NaCl in the experimental structure potentially damping this interaction. Alternatively, the protonation state of His163 may influence the results. His163 was treated as neutral in the MD simulations; hypothetically, a different result might be obtained if His163 was treated as being protonated. Presumably this would create a stronger interaction between RD Asp66 and His163, such that this interaction would dominate in both the ternary and the binary MD simulations, with the Asp66 backbone carbonyl oxygen interacting with the Lys11 sidechain amino group in both cases. Such a scenario where the interaction of Asp66 with Lys11 involves the formers backbone atoms could explain why a significant

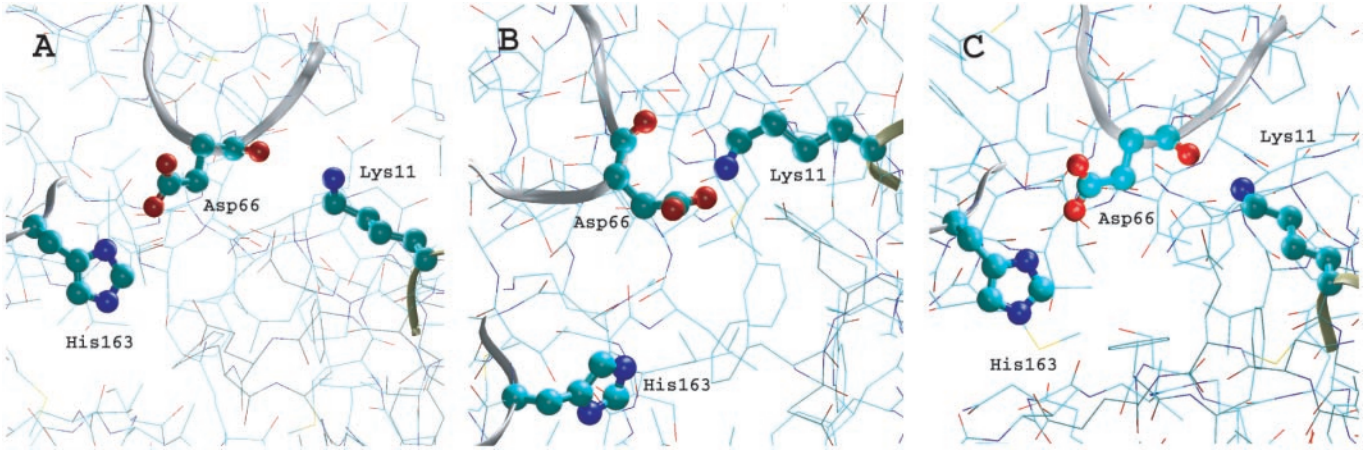

Figure 4. RD–CBFβ interactions involving RD residues Asp66 and His163 and CBFβ residue Lys11. Images from snapshots at 4.5 ns from the (**A**) RD–CBFβ binary and (**B**) ternary MD simulations and (**C**) from the X-ray crystal structure.

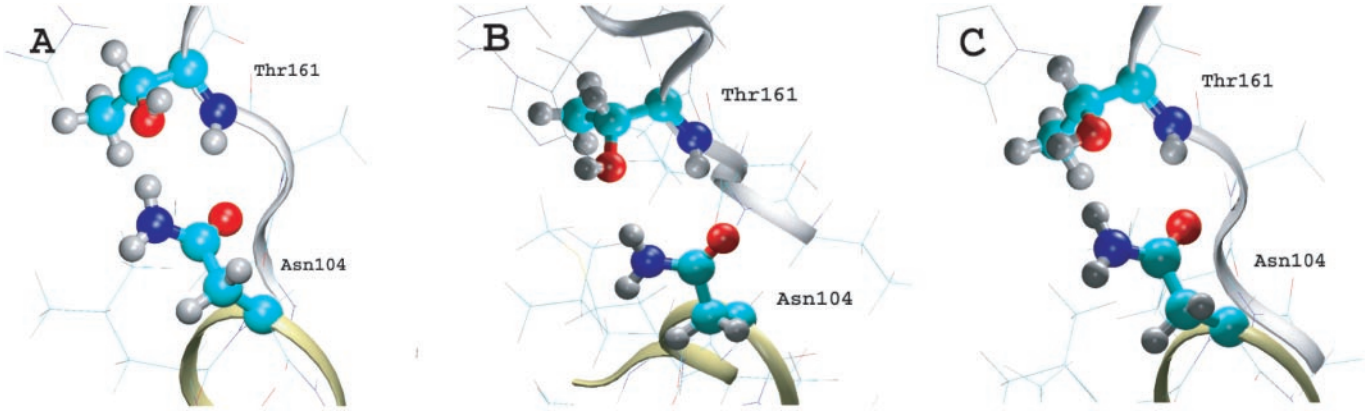

Figure 5. RD–CBFβ interactions involving RD residue Thr161 and CBFβ residue Asn104. Images from snapshots at 4.5 ns from the (**A**) RD–CBFβ binary and (**B**) ternary MD simulations and (**C**) from the X-ray crystal structure.

contribution of this residue to binding was not observed experimentally.

Area2 residues make favorable contributions to the interaction energy between RD and CBFβ (Table 5). Although this region does not contribute to the cooperativity, as evidenced by the similar contribution in the binary and the ternary complexes, the large favorable interaction energy contribution makes it of interest to investigate the role of individual residues in this region to binding (Table 6). The majority of residues in this region make favorable contribution to the interaction energy. Experimental studies found Thr161 to be a 'hot spot' residue (18). The mutation of Thr161 to alanine increased the RD–CBFβ dissociation constant in the ternary complex by 40-fold (18). The importance of Thr161 is thought to be due to its interaction with Asn104 of CBFβ. Indeed, visual inspection of the RD–CBFβ binary MD, ternary MD and crystal structures shows two hydrogen bonds between these residues [Asn(NH_2)-Thr(OG1) and Asn(OD1-Thr(HN)] (Figure 5). However, calculation of the residue–residue interaction energy between RD Thr161 and CBFβ Asn104 yields a value of -4.9 ± 0.4 kcal/mol, including the contribution from the Thr HN atom, which does not correlate well with the change in dissociation constant in the T161A mutation. Alternatively, calculation of the contribution of Thr161 to the strain energy of RD yields values of -21.0 ± 1.8 and -26.5 ± 0.2 kcal/mol in the RD monomer and RD–CBFβ dimer, respectively. Interestingly, it has been reported that the mutation of Thr161 to alanine leads to decreased exchange broadening in RD alone in NMR experiments, indicating a decrease in the conformational flexibility of residue 161 and surrounding residues (18). This observation combined with the present energetic results suggests that Thr161 may play a role in maintaining the conformation of the local region of RD, thereby impacting binding, rather than contributing directly to protein–protein interactions. Such a role of Thr161 is consistent with the idea that conserved polar residues at protein–protein binding sites provide structural rigidity (47). The role of such conserved residues is primarily structural rather than functional.

As discussed previously (13) and as emphasized above, the region of RD that includes the segments Wing and Wing1 appears to be of central importance for binding cooperativity. Presented in Table 7 are the interaction energies between

Table 7. Interaction energies between individual Wing and Wing1 residues in RD alone and in the three complexes

Residue	RD	RD–CBFβ	RD–DNA	CBFβ–RD–DNA
Wing residues to Wing1				
R139	−74.4 ± 1.0	−81.0 ± 1.8	−75.7 ± 1.0	−75.5 ± 0.8
S140	1.5 ± 0.5	−1.2 ± 0.6	1.3 ± 0.0	−2.2 ± 0.2
G141	−2.9 ± 0.5	−7.6 ± 1.0	−2.3 ± 1.0	−8.1 ± 0.4
R142	−68.9 ± 2.3	−118.6 ± 11.1	−61.1 ± 2.0	−65.9 ± 1.3
G143	−3.5 ± 0.2	−6.0 ± 0.2	−0.5 ± 0.0	0.6 ± 0.1
K144	−213.6 ± 5.0	−185.2 ± 2.5	−206.2 ± 5.0	−213.4 ± 3.9
S145	1.9 ± 0.4	0.9 ± 0.5	2.0 ± 0.0	0.7 ± 0.3
Wing1 residues to Wing				
T104	0.4 ± 0.0	0.4 ± 0.0	0.4 ± 0.0	0.4 ± 0.0
V105	0.3 ± 0.0	0.2 ± 0.0	0.2 ± 0.0	0.1 ± 0.0
M106	0.4 ± 0.1	0.5 ± 0.1	0.4 ± 0.1	0.2 ± 0.0
A107	0.8 ± 0.0	0.9 ± 0.0	0.7 ± 0.1	0.8 ± 0.0
G108	0.6 ± 0.1	0.8 ± 0.1	0.5 ± 0.1	0.6 ± 0.0
N109	−3.1 ± 0.7	−6.1 ± 0.5	−0.8 ± 1.7	−6.7 ± 0.2
D110	−153.2 ± 0.9	−158.9 ± 1.4	−149.4 ± 0.5	−141.4 ± 1.8
E111	−142.8 ± 6.6	−168.9 ± 7.4	−126.0 ± 5.6	−149.9 ± 3.4
N112	−2.3 ± 0.7	−5.6 ± 0.7	−2.3 ± 0.7	−5.3 ± 0.2
Y113	−0.2 ± 0.2	−0.1 ± 0.1	−0.5 ± 0.2	0.5 ± 0.3
S114	−0.9 ± 0.0	−1.0 ± 0.1	−1.1 ± 0.1	−0.9 ± 0.1
A115	1.9 ± 0.1	1.2 ± 0.1	2.0 ± 0.1	1.0 ± 0.1
E116	−61.8 ± 0.8	−62.2 ± 0.7	−66.5 ± 1.8	−63.3 ± 1.2
Total[a]	−359.9	−398.7	−342.5	−363.8

Energies in kcal/mol.
[a]Total interaction energy between Wing and Wing1 residues.

Wing and Wing1, including a breakdown of the interaction of Wing residues with all Wing1 residues and vice versa. Upon going from RD to the RD–CBFβ dimer, the Wing–Wing1 total interaction energy becomes significantly more favorable, even though favorable intermolecular interactions occur between Wing1 and CBFβ in the RD–CBFβ binary complex (Table 5). This suggests that the binding of CBFβ to RD, while forming a favorable interaction of CBFβ with Wing1 also induces a conformational change in Wing1 that favors the Wing–Wing1 interaction. Indeed, Figure 6 shows that two salt bridges, Arg142-Glu111 and Lys144-Asp110, are present in the RD–CBFβ dimer (Figure 6B); this represents a gain in interaction energy over the hydrogen bonds between Lys144 and Asp110 and Glu111 in RD alone (Figure 6A). Indeed, these favorable interactions contribute to the gain in favorable strain energy upon formation of the RD–CBFβ dimer

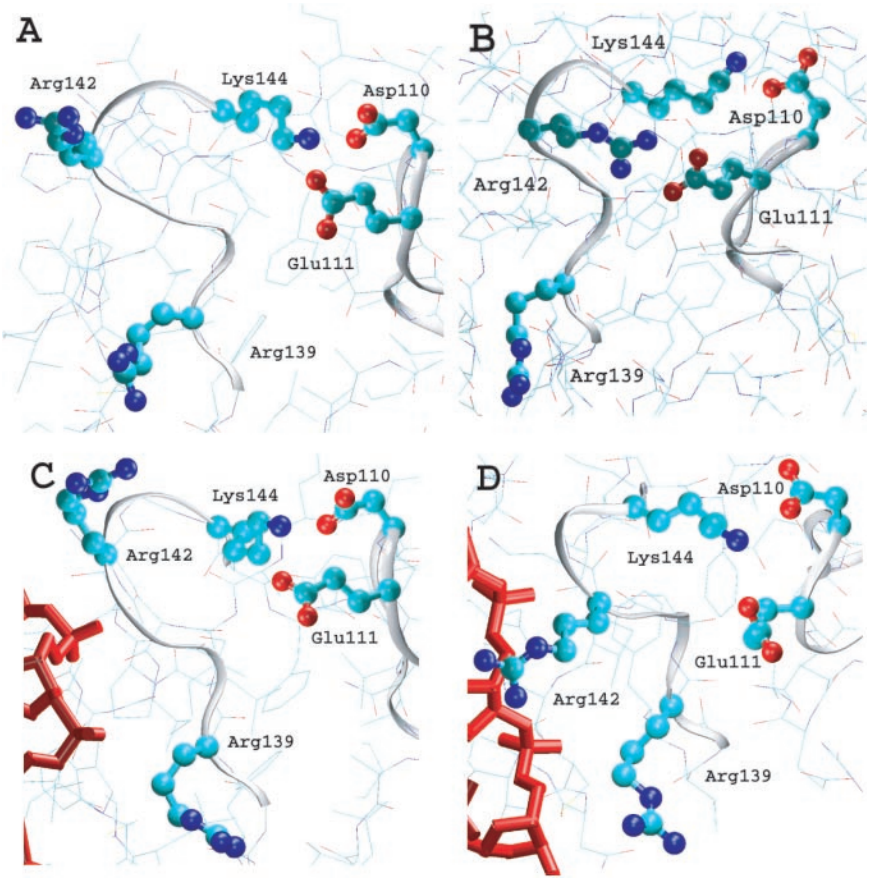

Figure 6. RD Wing–Wing1 interaction in the MD simulation structures at 4.5 ns for (**A**) RD, (**B**) RD–CBFβ, (**C**) RD–DNA and (**D**) the trimer. RD residues Arg139, Arg142, Lys144, Asp110 and Glu111 are shown. DNA is shown in red.

discussed above (Table 1). Moreover, NMR experiments indicate all four of these residues to undergo changes in mobility upon going from the RD–CBFβ dimer to the trimer (20). Thus, these intramolecular interactions make a significant contribution to the binding of RD to CBFβ.

In contrast to the RD–CBFβ interaction, upon going from RD alone to the RD–DNA binary complex, the Wing–Wing1 total interaction energy becomes less favorable. This is consistent with the significantly favorable interaction energy between RD and DNA in the binary complex, such that when the Wing residues interact with the DNA, the Wing–Wing1 interactions are diminished (Table 7). Similarly, the Wing–Wing1 interaction energy in the trimer is not significantly different as that in RD alone. Visual analysis of the RD structure in the monomer (Figure 6A), the RD–DNA complex (Figure 6C) and the ternary complex (Figure 6D), reveals the spatial relationships of residues Asp110, Glu111, Arg142 and Lys144 to be similar, consistent with the interaction energy contributions in these systems (Table 7). In the RD–DNA binary and ternary complexes, it is evident that the negatively charged DNA attracts Arg142 away from the Wing1 residues, contributing to the less favorable interactions. In the RD monomer, the lack of salt bridges between Arg142 and Asp110 or Glu111 may be related to repulsion between the two positively charged residues, Arg142 and Lys144, such that the system gains more stabilization by orienting Arg142 toward the solvent. The balance of attraction and repulsion taking place between closely situated positively and negatively charged residues suggests a subtle balance of interactions, such that their relative orientations are sensitive to the influence of solvent and the presence of DNA and CBFβ.

As stated above, formation of the ternary complex leads to a Wing–Wing1 interaction energy that is similar to that of the RD monomer. Energy partitioning analysis (Table 7) shows significantly less favorable interaction energies involving Arg142, Asp110 and Glu111 as compared with the CBFβ–RD binary complex. The less favorable interactions of Asp110 and Glu111 are consistent with the loss of salt bridges with Arg142, although more favorable interactions with Lys144 occur (Figure 6D). Thus, a subtle balance of intramolecular and intermolecular interactions appear to contribute to the observed cooperativity. Interaction of RD with DNA leads to favorable intermolecular interactions (Table 3) with a simultaneous loss of Wing–Wing1 interactions (Table 7). However, upon formation of the ternary complex, the binding of CBFβ leads to the favorable Wing–Wing1 interactions being regained, which also enhances the Wing1–DNA interactions.

Interestingly, the majority of residues predicted to make the largest contributions to cooperativity are charged (i.e. RD residues Asp66, Asp110, Glu111, Glu116, Arg139, Arg142 and Lys144). Consequently, the long-range character of electrostatic interactions may be an important factor in the observed cooperativity. The results suggest some interesting mutations that may impact cooperativity. For example, Asp to

Glu mutations, or vice versa, of the acidic residues suggested to be important for cooperativity may alter the subtle balance of intermolecular and intramolecular interactions contributing to the cooperativity. In addition, mutation of any of the charged residues discussed above to neutral species would be expected to impact cooperativity. Asp110 and Glu111 may be the most interesting as their impact on CBFβ binding would be expected to be minimal while their impact on cooperativity should be large.

CONCLUSION

MD simulations of RD along with its binding partners CBFβ and DNA, as well as the CBFβ–RD and RD–DNA binary complexes and the CBFβ–RD–DNA ternary complex have been performed to better understand the cooperative nature of the binding of these residues. In accordance with the experiment, the simulation data show cooperative binding of DNA in the presence of CBFβ as well as predict that binding of DNA to RD facilitates RD–CBFβ binding, indicating cooperativity for this interaction. These observations are consistent with the experimental data (16,18), supporting that ability of the applied simulation models to represent the experimental regimen.

Energetic and structural analysis allow for contributions to the binding and cooperativity to be understood in terms of different regions of the RD protein as well as the contribution of individual residues. Of note is the balance of intermolecular and intramolecular interactions that result in the cooperative binding. These interactions are, to a large extent, dominated by charged residues suggesting that long-range interactions may contribute to cooperativity. From these results a number of residues of interest for mutational analysis are suggested.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at NAR Online.

ACKNOWLEDGEMENTS

The National Institutes of Health is acknowledged for financial support (GM 51501 and CA 95350) and the National Science Foundation Partnerships for Advanced Computational Infrastructure Program, Department of Defense Aeronautical Systems Center Major Shared Resource Computing, and the Pittsburgh Supercomputing Center Terascale Computing for computational support. Funding to pay the Open Access publication charges for this article was provided by NIH GM51501.

Conflict of interest statement. None declared.

REFERENCES

1. Ogawa,E., Maruyama,M., Kagoshima,H., Inuzuka,M., Lu,J., Satake,M., Shigesada,K. and Ito,Y. (1993) PEBP2/PEA2 represents a family of transcription factors homologous to the products of the *Drosophila* runt gene and the human AML1 gene. *Proc. Natl Acad. Sci. USA*, **90**, 6859–6863.
2. Okuda,T., van Deursen,J., Hiebert,S.W., Grosveld,G. and Downing,J.R. (1996) AML1, the target of multiple chromosomal translocations in human leukemia, is essential for normal fetal liver hematopoiesis. *Cell*, **84**, 321–330.
3. Otto,F., Kanegane,H. and Mundlos,S. (2002) Mutations in the RUNX2 gene in patients with cleidocranial dysplasia. *Hum. Mutat*., **19**, 209–216.
4. Lund,A.H. and van Lohuizen,M. (2002) RUNX: a trilogy of cancer genes. *Cancer Cell*, **1**, 213–215.
5. Namba,K., Abe,M., Saito,S., Satake,M., Ohmoto,T., Watanabe,T. and Sato,Y. (2000) Indispensable role of the transcription factor PEBP2/CBF in angiogenic activity of a murine endothelial cell MSS31. *Oncogene*, **19**, 106–114.
6. Sun,L., Vitolo,M. and Passaniti,A. (2001) Runt-related gene 2 in endothelial cells: inducible expression and specific regulation of cell migration and invasion. *Cancer Res*., **61**, 4994–5001.
7. Kato,N., Tamura,G., Fukase,M., Shibuya,H. and Motoyama,T. (2003) Hypermethylation of the RUNX3 gene promoter in testicular yolk sac tumor of infants. *Am. J. Pathol*., **163**, 387–391.
8. Li,Q.L., Ito,K., Sakakura,C., Fukamachi,H., Inoue,K., Chi,X.Z., Lee,K.Y., Nomura,S., Lee,C.W., Han,S.B. *et al*. (2002) Causal relationship between the loss of RUNX3 expression and gastric cancer. *Cell*, **109**, 113–124.
9. Goseki-Sone,M., Orimo,H., Watanabe,A., Hamatani,R., Yokozeki,M., Ohyama,K., Kuroda,T., Watanabe,H., Miyazaki,H., Shimada,T. *et al*. (2001) Identification of a novel frameshift mutation (383insT) in the RUNX2 (PEBP2 alpha/CBFA1/AML3) gene in a Japanese patient with cleidocranial dysplasia. *J. Bone Miner. Metab*., **19**, 263–266.
10. Backstrom,S., Wolf-Watz,M., Grundstrom,C., Hard,T., Grundstrom,T. and Sauer,U.H. (2002) The Runx1 Runt domain at 1.25 Å resolution: a structural switch and specifically bound chloride ions modulate DNA binding. *J. Mol. Biol*, **322**, 259–272.
11. Tahirov,T.H., Inoue-Bungo,T., Sasaki,M., Shiina,M., Kimura,K., Sato,K., Kumasaka,T., Yamamoto,M., Kamiya,N. and Ogata,K. (2001) Crystallization and preliminary X-ray analyses of quaternary, ternary and binary protein–DNA complexes with involvement of AML1/Runx-1/CBFalpha Runt domain, CBFβ and the C/EBPβ bZip region. *Acta Crystallogr. D Biol. Crystallogr*., **57**, 850–853.
12. Warren,A.J., Bravo,J., Williams,R.L. and Rabbitts,T.H. (2000) Structural basis for the heterodimeric interaction between the acute leukaemia-associated transcription factors AML1 and CBFβ. *EMBO J*., **19**, 3004–3015.
13. Bravo,J., Li,Z., Speck,N.A. and Warren,A.J. (2001) The Leukemia-associated AML1 (Runx1)–CBRß complex functions as a DNA-induced molecular clamp. *Nature Struct. Biol*., **8**, 371–377.
14. Bartfeld,D., Shimon,L., Couture,G.C., Rabinovich,D., Frolow,F., Levanon,D., Groner,Y. and Shakked,Z. (2002) DNA recognition by the RUNX1 transcription factor is mediated by an allosteric transition in the Runt domain and by DNA bending. *Structure*, **10**, 1395–1407.
15. Crute,B.E., Lewis,A.F., Wu,Z., Bushweller,J.H. and Speck,N.A. (1996) Biochemical and biophysical properties of the core-binding factor alpha2 (AML1) DNA-binding domain. *J. Biol. Chem*., **271**, 26251–26260.
16. Tang,Y.Y., Crute,B.E., Kelley,J.J., Huang,X., Yan,J., Shi,J., Hartman,K.L., Laue,T.M., Speck,N.A. and Bushweller,J.H. (2000) Biophysical characterization of interactions between the core binding factor alpha and beta subunits and DNA. *FEBS Lett*., **470**, 167–172.
17. Tahirov,T.H., Inoue-Bungo,T., Morii,H., Fujikawa,A., Sasaki,M., Kimura,K., Shiina,M., Sato,K., Kumasaka,T., Yamamoto,M. *et al*. (2001) Structural analyses of DNA recognition by the AML1/Runx-1 Runt domain and its allosteric control by CBFβ. *Cell*, **104**, 755–767.
18. Zhang,L., Li,Z., Yan,J., Pradhan,P., Corpora,T., Cheney,M.D., Bravo,J., Warren,A.J., Bushweller,J.H. and Speck,N.A. (2003) Mutagenesis of the runt domain defines two energetic hot spots for heterodimerization with the core binding factor β subunit. *J. Biol. Chem*., **278**, 33097–33104.
19. Li,Z., Yan,J., Matheny,C.J., Corpora,T., Bravo,J., Warren,A.J., Bushweller,J.H. and Speck,N.A. (2003) Energetic contribution of residues in the Runx1 runt domain to DNA binding. *J. Biol. Chem*., **278**, 33088–33096.
20. Yan,J., Liu,Y., Lukasik,S.M., Speck,N.A. and Bushweller,J.H. (2004) CBFβ allosterically regulates the Runx1 Runt domain via a dynamic conformational equilibrium. *Nature Struct. Mol. Biol*., **11**, 901–906.
21. In Becker,O., MacKerell,A.D.Jr, Roux,B. and Watanabe,M. (eds), *Computational Biochemistry and Biophysics*. . Marcel Dekker, Inc., NY.
22. Gohlke,H., Kiel,C. and Case,D.A. (2003) Insights into protein–protein binding by binding free energy calculation and free energy decomposition for the Ras–Raf and Ras–RaIGDS complexes. *J. Mol. Biol*., **330**, 891–913.

23. Jayaram,B., McConnell,K., Dixit,S.B., Das,A. and Beveridge,D.L. (2002) Free-energy component analysis of 40 protein–DNA complexes: a consensus view on the thermodynamics of binding at the molecular level. *J. Comput. Chem.*, **23**, 1–14.
24. Kollman,P.A., Massova,I., Reyes,C., Kuhn,B., Huo,S.H., Chong,L., Lee,M., Lee,T., Duan,Y., Wang,W. *et al.* (2000) Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc. Chem. Res.*, **33**, 889–897.
25. Still,W.C., Tempczyk,A., Hawley,R.C. and Hendrickson,T. (1990) Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.*, **112**, 6127–6129.
26. Lee,M.S., Salsbury,F.R. and Brooks,C.L.I. (2002) Novel generalized Born methods. *J. Chem. Phys.*, **116**, 10606–10614.
27. Tang,Y.Y., Shi,J., Zhang,L., Davis,A., Bravo,J., Warren,A.J., Speck,N.A. and Bushweller,J.H. (2000) Energetic and functional contribution of residues in the core binding factor β (CBFβ) subunit to heterodimerization with CBFalpha. *J. Biol. Chem.*, **275**, 39579–39588.
28. Brooks,B.R., Bruccoleri,R.E., Olafson,B.D., States,D.J., Swaminathan,S. and Karplus,M. (1983) CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.*, **4**, 187–217.
29. MacKerell,A.D.Jr, Bashford,D., Bellott,M., Dunbrack,R.L., Evanseck,J.D., Field,M.J., Fischer,S., Gao,J., Guo,H., Ha,S. *et al.* (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem.*, **B102**, 3586–3616.
30. Foloppe,N. and MacKerell,A.D.Jr (2000) All-atom empirical force field for nucleic acids: 1) Parameter optimization based on small molecule and condensed phase macromolecular target data. *J. Comput. Chem.*, **21**, 86–104.
31. MacKerell,A.D.Jr and Banavali,N. (2000) All-atom empirical force field for nucleic acids: 2) Application to molecular dynamics simulations of DNA and RNA in solution. *J. Comput. Chem.*, **21**, 105–120.
32. Jorgensen,W.L. (1983) Theoretical studies of medium effects on conformational equilibria. *J. Phys. Chem.*, **87**, 5304–5312.
33. Beglov,D. and Roux,B. (1997) An integral equation to describe the solvation of polar molecules in liquid water. *J. Phys. Chem.*, **101**, 7821–7826.
34. Arnott,S. and Hukins,D.W.L. (1973) Refinement of the structure of b-dna and implications for the analysis of x-ray diffraction data from fibers of biopolymers. *J. Mol. Biol.*, **81**, 93–105.
35. Feller,S.E., Zhang,Y.H., Pastor,R.W. and Brooks,B.R. (1995) Constant-pressure molecular-dynamics simulation—the langevin piston method. *J. Chem. Phys.*, **103**, 4613–4621.
36. Ryckaert,J.P., Ciccotti,G. and Berendsen,H.J.C. (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comp. Phys.*, **23**, 327–341.
37. York,D.M., Darden,T.A. and Pedersen,L.G. (1993) The effect of long-range electrostatic interactions in simulations of macromolecular crystals: a comparison of the Ewald and truncated list methods. *J. Chem. Phys.*, **99**, 8345–8348.
38. Steinbach,P.J. and Brooks,B.R. (1994) New spherical-cutoff methods for long-range forces in macromolecular simulation. *J. Comput. Chem.*, **15**, 667–683.
39. Humphrey,W., Dalke,A. and Schulten,K. (1996) VMD: visual molecular dynamics. *J. Mol. Graph.*, **14**, 33–38.
40. Merritt,E.A. and Bacon,D.J. (1997) Raster3D photorealistic molecular graphics. *Methods Enzymol.*, **277**, 505–524.
41. Lee,B. and Richards,F.M. (1971) The interpretation of protein structures: estimation of static accessibility. *J. Mol. Biol.*, **55**, 379–400.
42. Cramer,C.J. (2002) *Essentials of Computational Chemistry: Theories and Models*. J. Wiley, West Sussex, UK.
43. Loncharich,R.J., Brooks,B.R. and Pastor,R.W. (1992) Langevin dynamics of peptides—the frictional dependence of isomerization rates of *N*-acetylalanyl-*N′*-methylamide. *Biopolymers*, **32**, 523–535.
44. MacKerell,A.D.Jr (2004) Empirical force fields for biological macromolecules: overview and issues. *J. Comput. Chem.*, **25**, 1584–1604.
45. Lima,L.M.T.R., Foguel,D. and Silva,J.L. (2000) DNA tightens the dimeric DNA-binding domain of human papillomavirus E2 protein without changes in volume. *Proc. Natl Acad. Sci. USA*, **97**, 14289–14294.
46. Berardi,M.J., Sun,C., Zehr,M., Abildgaard,F., Peng,J., Speck,N.A. and Bushweller,J.H. (1999) The Ig fold of the core binding factor alpha Runt domain is a member of a family of structurally and functionally related Ig-fold DNA-binding domains. *Struct. Fold. Des.*, **7**, 1247–1256.
47. Ma,B., Elkayam,T., Wolfson,H. and Nussinov,R. (2003) Protein–protein interactions: structurally conserved residues distinguish between binding sites and exposed protein surfaces. *Proc. Natl Acad. Sci. USA*, **100**, 5772–5777.