Interaction energy decomposition in protein–protein association: A quantum mechanical study of barnase–barstar complex

Abdessamad Ababou [a,c,*], Arjan van der Vaart [d], Valentin Gogonea [b], Kenneth M. Merz Jr. [e,c]

[a] *Department of Biochemistry and Molecular Biology, University College London, Gower Sreet, WC1E 6BT, UK*
[b] *Department of Chemistry, Cleveland State University, 1983 East 24th Street, Cleveland, Ohio 44115, USA*
[c] *Department of Chemistry, 152 Davey Laboratory, The Pennsylvania State University, University Park, Pennsylvania 16802, USA*
[d] *Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138, USA*
[e] *Department of Chemistry, University of Florida, 2328 New Physics Building, Gainesville, Florida 32611-8435, USA*

Abstract

Protein–protein interactions are very important in the function of a cell. Computational studies of these interactions have been of interest, but often they have utilized classical modelling techniques. In recent years, quantum mechanical (QM) treatment of entire proteins has emerged as a powerful approach to study biomolecular systems. Herein, we apply a semi-empirical divide and conquer (DC) methodology coupled with a dielectric continuum model for the solvent, to explore the contribution of electrostatics, polarization and charge transfer to the interaction energy between barnase and barstar in their complex form. Molecular dynamic (MD) simulation was performed to account for the dynamic behavior of the complex. The results show that electrostatics, charge transfer and polarization favor the formation of the complex. Our study shows that electrostatics dominates the interaction between barnase and barstar ($\sim 73\%$), while charge transfer and polarization are $\sim 21\%$ and $\sim 6\%$, respectively. Close inspection of the polarization and charge-transfer effects on the charge distribution of the complex reveals the existence of two, well localized, regions in barstar. The first region includes the residues between P27 and Y47 and the second region is between N65 and D83. Since no such regions could be detected in barnase clearly suggests that barstar is well optimized for efficiently binding barnase. Furthermore, using our interaction energy decomposition scheme, we were able to identify all residues that have been experimentally determined to be important for the complex formation and to suggest other residues never have been investigated. This suggests that our approach will be useful as an aid in further understanding protein–protein contacts for the ultimate goal to produce successful inhibitors for protein complexes.
© 2006 Elsevier B.V. All rights reserved.

1. Introduction

Protein–protein association is an important process in biological systems. It is involved in many fundamental processes such as cytoskeletal remodeling [1,2], signal transduction [3–5], transcription [6], cell cycle regulation [7,8] and immune response [9–12]. Understanding these biological processes at the molecular level requires an understanding of the protein–protein interactions at the molecular level. The chemical nature of protein–protein interfaces has been described as having a composition which is intermediate between the composition of the interior (hydrophobic) and exterior (polar) of a typical protein [13,14]. Other studies show that the residue compositions of most protein–protein interfaces are more similar to that of protein surfaces than their cores [15,16]. Moreover, it has been reported that these interfaces have a larger proportion of hydrophobic residues [17]. A recent study revealed that, on average, more than half of the interface area is formed by hydrophobic residues [18]. Furthermore, in the case of the T-lymphocyte cell surface antigen, CD2, the ligand-binding surface is highly charged even more than observed for a typical protein surface [19]. Thus, the chemical nature of protein–protein interfaces varies among different families of protein complexes and hence

---

* Corresponding author. Department of Biochemistry and Molecular Biology, University College London, Gower Sreet, WC1E 6BT, UK. Tel.: +44 20 7679 2174; fax: +44 20 7679 7193.

E-mail address: samad@biochem.ucl.ac.uk (A. Ababou).

is not helpful for fully understanding protein–protein associations. The thermodynamic nature of protein–protein interactions is important for our understanding of the association between protein surfaces and potentially provides us with a rational for designing effective ligands that might interfere with protein/protein association. Consequently, many experimental studies have put effort in the determination of the thermodynamic (free energy of binding) and kinetic quantities (rate constant of association/dissociation) associated with protein–protein complex formation [20–28]. Experimental techniques, such as isothermal calorimetry (ITC) or surface plasmon resonance (SPR), while very powerful in providing macroscopic insights do not provide an adequate rationalization for protein–protein association (or interaction) energies in terms of separate energetic contributions, such as electrostatic and/or van der Waals contributions. However, computational methods have the advantage of evaluating each energetic contribution involved in protein–protein association, which allows us to explore these contributions for the entire interface or for individual residues. Several computational studies dealing with prediction of the binding free energy of (ΔG_{bind}) in protein–protein association as well as the contribution ($\Delta \Delta G_{bind}$) of a given residue have appeared [23,24,29–38]. All these methods are based on molecular mechanics (MM) calculations, using conventional MM force fields such as AMBER [39] and CHARMM [40], and a dielectric continuum model for the solvent [41].

Many body effects, such as polarization, can be evaluated with properly parametrized MM force fields [42], and significant effort has been dedicated to develop polarizable protein force fields [43–45]. However, difficulties within polarizable force fields, such as molecular flexibility, are not yet well understood [46] and to our knowledge no work has been reported regarding the inclusion of polarization effects in protein–protein association. Moreover, other quantum mechanical phenomena, such as charge transfer (CT), cannot be handled by classical methods. The importance of CT in the energetics of solvated macromolecular systems has been reported in the case of major cold shock protein A (CspA) [47]. The authors showed that charge transfer from protein to solvent is important in the first solvation layer.

Consequently, to explore the polarization and CT contributions in protein–protein interactions, in this paper, we use QM calculations to evaluate the protein–protein interaction energy in the barnase–barstar complex [48] (see Fig. 1). Using the earlier reported energy decomposition scheme [49], the interaction energy was decomposed into electrostatic, polarization and charge transfer contributions. This energy decomposition scheme was implemented using the divide and conquer (DC) semi-empirical method [49] and is available in the DivCon program [50–52]. The DC method is a linear scaling quantum mechanical approach, which allows the calculation of large molecular systems. By splitting the system into subsystems, the diagonalization of the Fock matrix of the total system is replaced by the diagonalization of several smaller matrixes in the most expensive part of the calculation [50]. Charge flow between subsystems is controlled by subsystem overlap and the Fermi energy. This method, in a modified form, also allows the elimination of polarization or both polarization and charge transfer, from the intermolecular interactions. Analysis of the atomic charges obtained from the separate intermolecular interaction
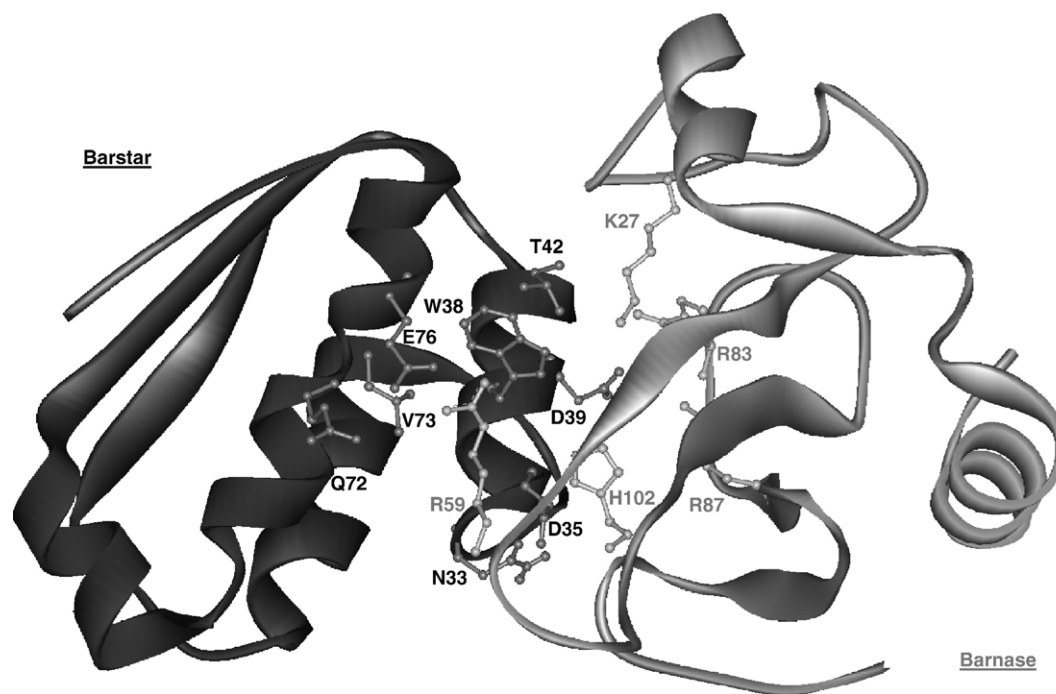

Fig. 1. Ribbon representation of the structure of barnase–barstar complex illustrating barnase (grey), barstar (black) and the important residues in binding at the interface. The atomic coordinates used are derived from the X-ray structure of the complex (Protein Data Bank code 1BRS).

calculation can then be used to rationalize the effect of polarization and charge transfer on the charge distribution of the system.

Here we investigate the energy contributions of electrostatic, polarization and CT in the interaction energy of barnase–barstar complex, and the effect of polarization and CT on the charge distribution of the complex. We discuss our results in term of the importance of each contribution and their effects in the formation of the complex in light of previously reported experimental work. Furthermore, we have attempted to identify whether there are other residues which are *important* (thermodynamically

and/or kinetically) for the formation of barnase–barstar complex and have never been investigated experimentally.

2. Results

2.1. MD simulation

The 500 ps MD trajectory of the barnase–barstar complex was used to generate snapshots (on which our energy decomposition scheme was applied), which account for the dynamics
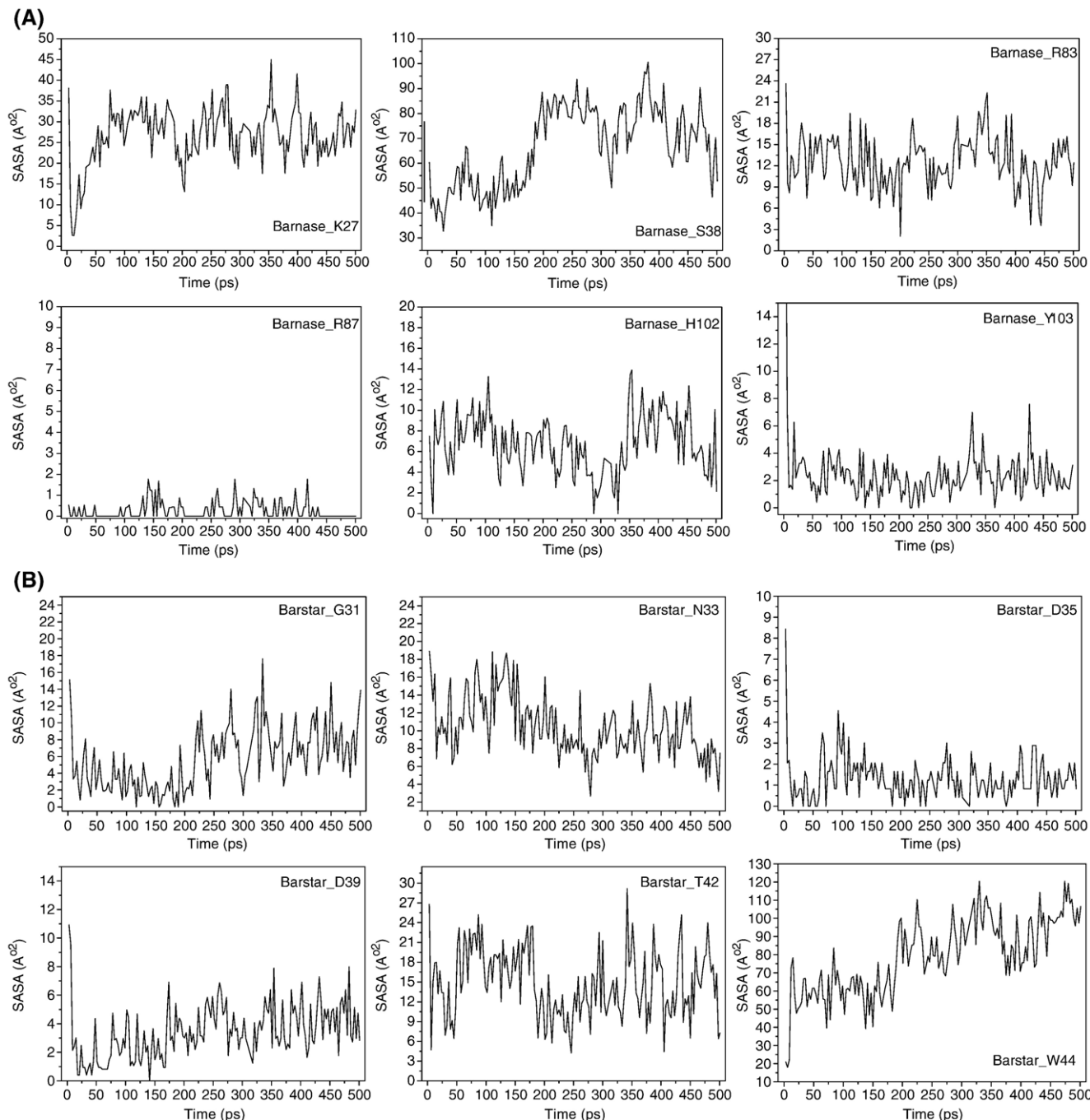

Fig. 2. (A) Plot of solvent accessible surface area (SASA) changes of particular residues from barnase as function of time. (B) Plot of solvent accessible surface area (SASA) changes of particular residues from barstar as function of time.

of the complex, like H-bond formation and/or breaking, as well as the exposure and/or burial of polar and charged side chains.

Using the solvent accessible surface area (SASA) as a criteria for identifying the interface residues of our complex structure (see Methods), we identified in barnase the following interface residues: K27, A37, S38, F56, R59, E60, F82, R83, S85, H102, Y103 and Q104. In barstar, we found Y29, G31, N33, L34, D35, W38, D39, T42, G43, W44 and E46 (which is at the limit of our ΔSASA criteria, i.e. ~22 Å), as interface residues. During the entire trajectory, these interface residues show very low accessibility to the solvent (Fig. 2A and B). Exceptions in barnase are
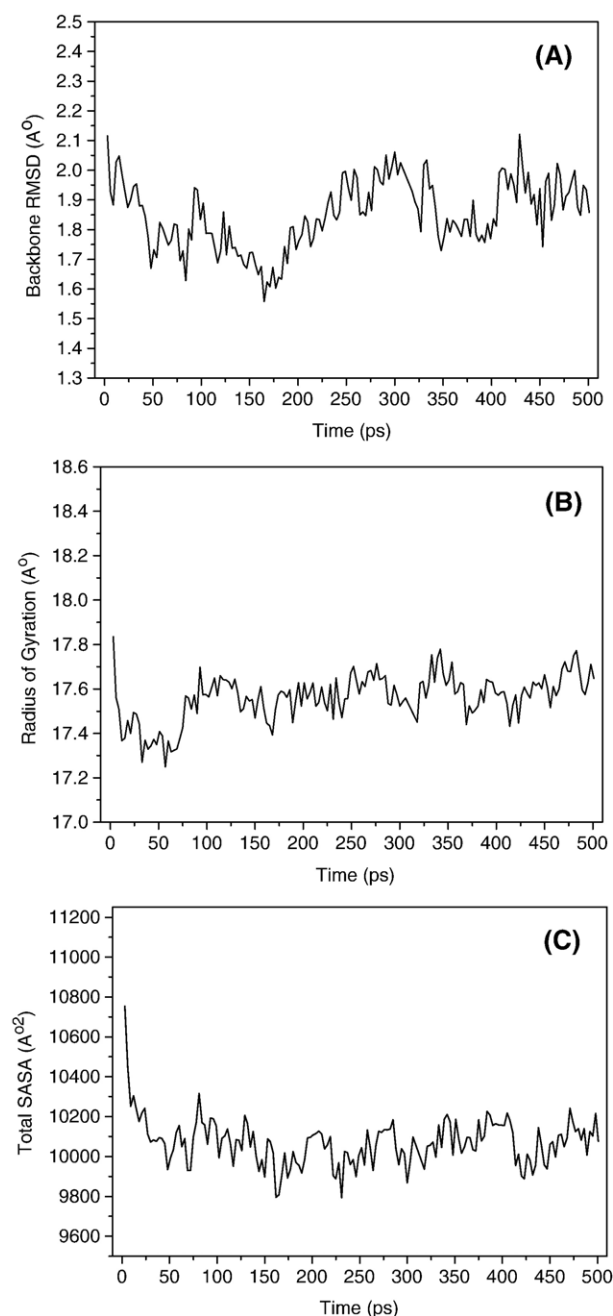

Fig. 3. Global MD trajectory analysis of barnase–barstar complex. (A) Backbone RMSD, (B) radius of gyration and (C) total SASA of the complex.

S38, R59, E60, F82 and Q104, and in barstar Y29, W38, W44 and E46. This result was expected since these residues are located at the external interface ring of the complex between barnase and barstar. Fig. 3 shows almost no significant changes in the radius of gyration and the total SASA of the complex during the simulation. This result suggests that the complex is stable during the simulation. The root-mean square deviation (RMSD) along the trajectory between the complex structures and the starting crystal structure (see Methods) shows some variability. It decreases toward 150 ps and then increases till 300 ps; however, it shows less variation around 455 ps and reaches a final value of 1.9 Å. Fig. 4 shows all H-bonds between barnase and barstar. We used a distance cutoff of 2.5 Å between the donor and the acceptor and an angle cutoff between 120° and 180° at the hydrogen atom. There are eight strong H-bonds (R59 (HH12)–E76(OD2), E60(H)–D35(OD2), R83(HH21)–D39 (OD2), R83(O)–Y29(HH), R87(HH11)–D39(OD2), H102 (O)–N33(HD22), R59(H)–D35(OD1), Y103(HH)–D39 (OD1)) which are present during the entire simulation. Interestingly, each of R59 and R87 has a second H-bond that lasts for more than 60% of time. In the case of E60-L34, two H-bonds exist for 50% of time for each carboxylate oxygen (OD1 and OD2) of E60, due to the rotation of the carboxylate group. However, ignoring the identity of these two H-bonds then E60–L34 has one H-bond, which is present for the entire simulation. The other H-bonds last for 50% of the simulation time or less as shown in Fig. 4.

The existence of such a H-bond network (side chain-side chain and side chain-backbone) and its long survival during the simulation suggests that the binding between barnase and barstar is quite strong, which is in agreement with the extensive work reported on this complex [53–55]. The H-bonds at the interface of the barnase–barstar complex seem to prefer to hydrogen bond with other protein atoms rather than water. For example, E76 in barstar is initially exposed to the solvent, but during the MD simulation its SASA decreased from 90 Å^2 to 45 Å^2 during the first 150 ps and after that remained at around 55 Å^2. The reason for this is the formation of a strong H-bond, which was present throughout the simulation, and another one which was intermediate in duration (Fig. 4). In contrast with our findings, previous work has reported that these H-bonds are weak due to constraints of the interface [56], which imply that E76 should prefer to H-bond to water rather than to another residue at the interface of the complex. The MD simulation provided some insights into the barnase–barstar interface, but it was predominantly generated to carry out a QM energy decomposition analysis, which is described in detail below.

2.2. Interaction energy decomposition

A comparison of the interaction energy decomposition between the complex C–F (C_{C-F}, see Methods) and the complex B–E (C_{B-E}) from the PDB (code 1brs) is summarized in Table 1. In C_{B-E}, three residues, Lys1, Glu64 and Asn65, are absent in barstar, while C_{C-F} lacks the first two residues Ala1 and Gln2 in barnase (see Methods). The purpose of this comparison was to choose the 'best' complex structure to be used in
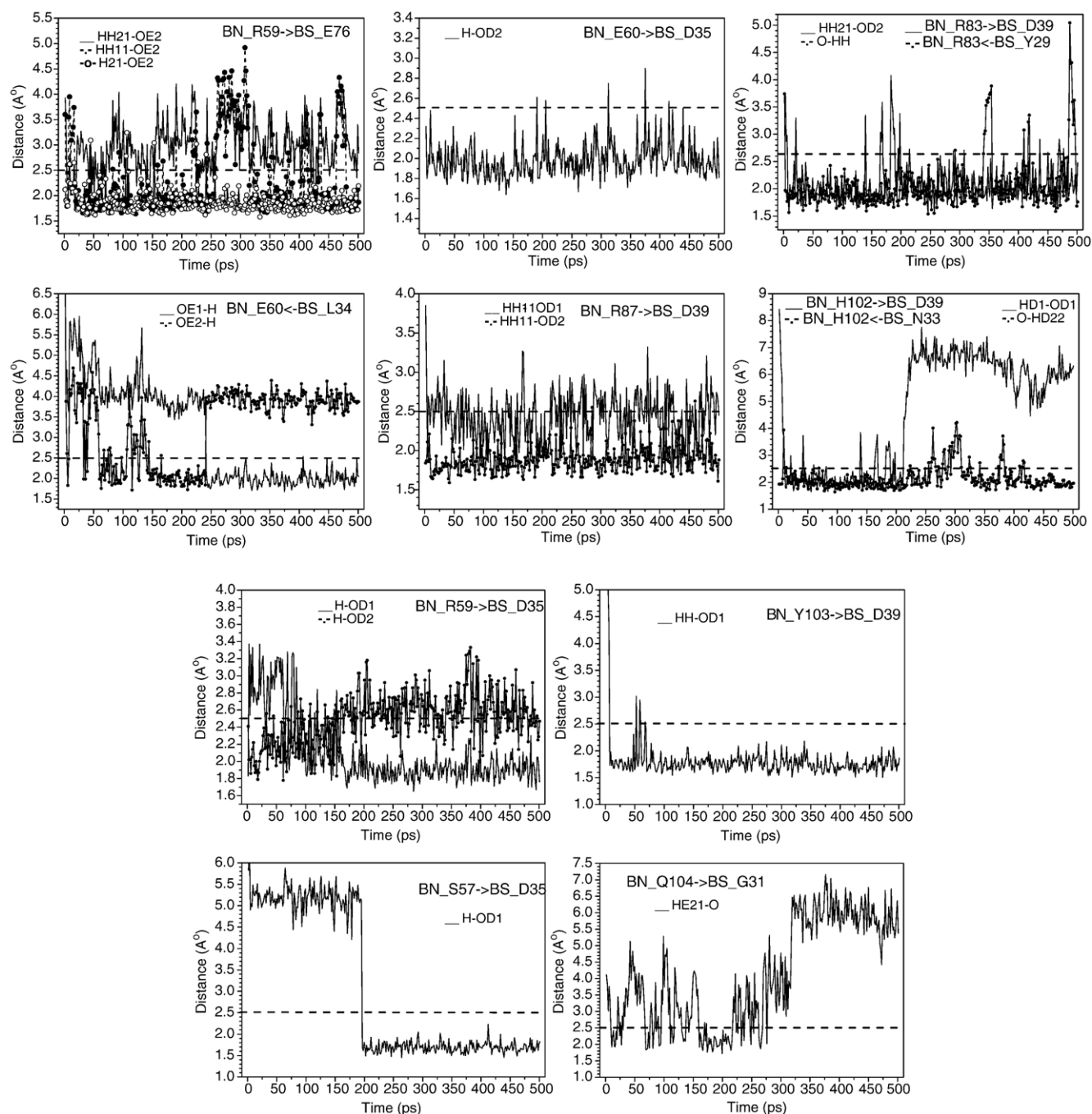
Fig. 4. Plot of hydrogen bond presence during the MD trajectory between donor–acceptor atoms of residues at the interface of the barnase–barstar complex. The dashed line represents the cutoff distance criteria of 2.5 Å for the presence of a hydrogen bond.

this work and to verify if there are any differences between the complexes in the PDB file, since other computational works on this complex have used C_{B-E} [34,57]. In both complexes, the electrostatics, polarization and charge transfer interactions calculated with either AM1 or PM3 Hamiltonian favor the formation of the complex. While the contribution of polarization in E_{INT} is similar for both complexes, the electrostatics is $\sim 14\%$ higher in C_{C-F} than C_{B-E}, and the charge transfer (CT) is $\sim 12\%$ lower in C_{C-F} than C_{B-E}. The CT difference between complexes can be attributed to the absence of some of the H-bonds at the

interface in the case of C_{C-F}. Indeed, the route for a CT interaction has been reported to be via hydrogen bonding [47,58]. Simple geometric criteria for H-bond determination show that C_{B-E} has four more H-bonds than C_{C-F} at the interface, due to the absent residues in the structure. In summary, our interaction energy decomposition for both complexes indicates that electrostatics is the dominant component in the association of the barnase–barstar complex. Although the importance of electrostatic has similar trend in both complexes, as it has been reported for C_{B-E} [34], we suggest that, for any similar type of

Table 1
Comparison of interaction energy decomposition between B–E and C–F complexes of banase–barstar

Complex (1brs)	E_{INT} (kcal/mol)	E_{ES} (kcal/mol)	E_{POL} (kcal/mol)	E_{CT} (kcal/mol)	%E_{ES}[a]	%E_{POL}[a]	%E_{CT}[a]
AM1 Hamiltonian							
B–E	−450.235	−286.481	−38.547	−125.206	63.6	8.6	27.8
Completed C–F	−471.451	−368.193	−28.123	−75.135	78.1	6.0	15.9
PM3 Hamiltonian							
B–E	−468.560	−230.807	−39.529	−198.223	49.3	8.4	42.3
Completed C–F	−489.776	−312.519	−29.105	−148.1512	63.8	5.9	30.3

[a] %E_X = 100|E_X|/(|E_{ET}| + |E_{POL}| + |E_{CT}|), X = ES, POL, CT.

calculations, missed residues or side chain atoms have to be modeled in order to avoid any misinterpretation of electrostatic or energetic contributions in the complex formation.

In the remainder of this article, we will only analyze the $C_{C–F}$ complex. To account for the dynamics of the complex, we have applied the interaction energy decomposition scheme to five snapshots from the MD trajectory (100, 200, 300, 400 and 500 ps). These calculations show that the electrostatic energy is stabilizing for binding and constitutes between ∼ 60% (PM3) and ∼ 73% (AM1) of the total interaction energy between barnase and barstar (Table 2). It is well known that the solvation effect (electrostatic part) opposes protein–protein binding due to the desolvation penalty of polar and charged residues [20,59,60]. Barnase–barstar complex is an extremely tight-binding complex containing many polar and charged groups at the interface (see MD simulation) and has a K_D of 10^{-14} M [53,54], which suggests that electrostatic interactions between the two proteins play a key role in this tight-binding and our energy decomposition results confirm this. Polarization contributes ∼ 6% (AM1 and PM3) to the total interaction energy and charge transfer (CT) ∼ 21% (AM1) to ∼ 34% (PM3). Polarization and CT also have a stabilizing effect on the binding. While the contribution of polarization is similar for both Hamiltonians (AM1 and PM3), there is a difference of 13% in the electrostatics contribution. For PM3, the core–core repulsion term contains non-physical "shoulders" in and around the minimum region which potentially may make this term more repulsive than in AM1 [49,61].

2.3. Polarization effect

Figs. 5A and 6A show the effect of polarization and CT on the charge distribution as the results of complexation calculated using AM1 and PM3, respectively. The bars represent the standard deviation (S.D.) in the partial charges of one residue in the five snapshots. The residues of the complex correlated to those of barnase and barstar are shown in separate plots. Polarization significantly alters the charge distribution of the complex, and the dashed lines in the figures indicate the average boundary of the polarization effect on the residue. For both monomers, the average range is between −0.005 and 0.005e, but for some residues the polarization effect of the total charge of the residue can be as large as ±0.2e. Upon complexation, polarization effects significantly alter the charged and polar residues; however, hydrophobic residues are affected as well, A37(+), G40(+), L42(+), I55(+), A74(+), G81(+), F56(−), G61(−), P64(−) and F82(−) in barnase, and G31(+), L34(+), L37(+), A40(+), L41(+), G43(+), V45(+), F74(+), P27(−), A36(−), W38(−), W44(−), V73(−), A81(−) and G83(−) in barstar. The sign (+) or (−) refers to loss or gain of electrons, respectively.

In general, negatively charged residues gain electrons, while positively charged residues lose electrons upon polarization. However, exceptions to this general rule exist as shown in Figs. 5A and 6A. Polarization is strongly dependent on the local environment [62]. For example, E60 in barnase loses electrons as well as D35 in barstar. Close inspection of the environment of

Table 2
Interaction energy decomposition of solvated barnase–barstar complex

Time (ps)	Hamiltonian	E_{INT} (kcal/mol)	E_{ES} (kcal/mol)	E_{POL} (kcal/mol)	E_{CT} (kcal/mol)	%E_{ES}[a]	%E_{POL}[a]	%E_{CT}[a]
100	AM1	−501.685	−377.631	−30.215	−93.840	75.3	6.0	18.7
100	PM3	−520.011	−321.957	−31.197	−166.857	61.9	6.0	32.1
200	AM1	−472.377	−349.786	−23.664	−98.926	74.0	5.0	20.9
200	PM3	−495.786	−296.936	−24.827	−174.024	59.9	5.0	35.1
300	AM1	−433.836	−322.771	−26.995	−84.069	74.4	6.2	19.4
300	PM3	−459.983	−284.137	−30.097	−145.749	61.8	6.5	31.7
400	AM1	−440.016	−307.809	−27.220	−104.987	70.0	6.2	23.9
400	PM3	−470.431	−264.721	−29.750	−175.961	56.3	6.3	37.4
500	AM1	−477.258	−349.417	−28.838	−99.004	73.2	6.0	20.7
500	PM3	−504.920	−305.880	−32.231	−166.809	60.6	6.4	33.0
Average	AM1	−465.034	−341.483	−27.386	−96.165	73.4	5.9	20.7
S.D.		28.044	27.031	2.457	7.830	2.1	0.5	2.0
Average	PM3	−490.226	−294.726	−29.620	−165.880	60.1	6.1	33.9
S.D.		24.700	21.698	2.852	11.990	2.3	0.6	2.4

[a] %E_X = 100|E_X|/(|E_{ET}| + |E_{POL}| + |E_{CT}|), X = ES, POL, CT.
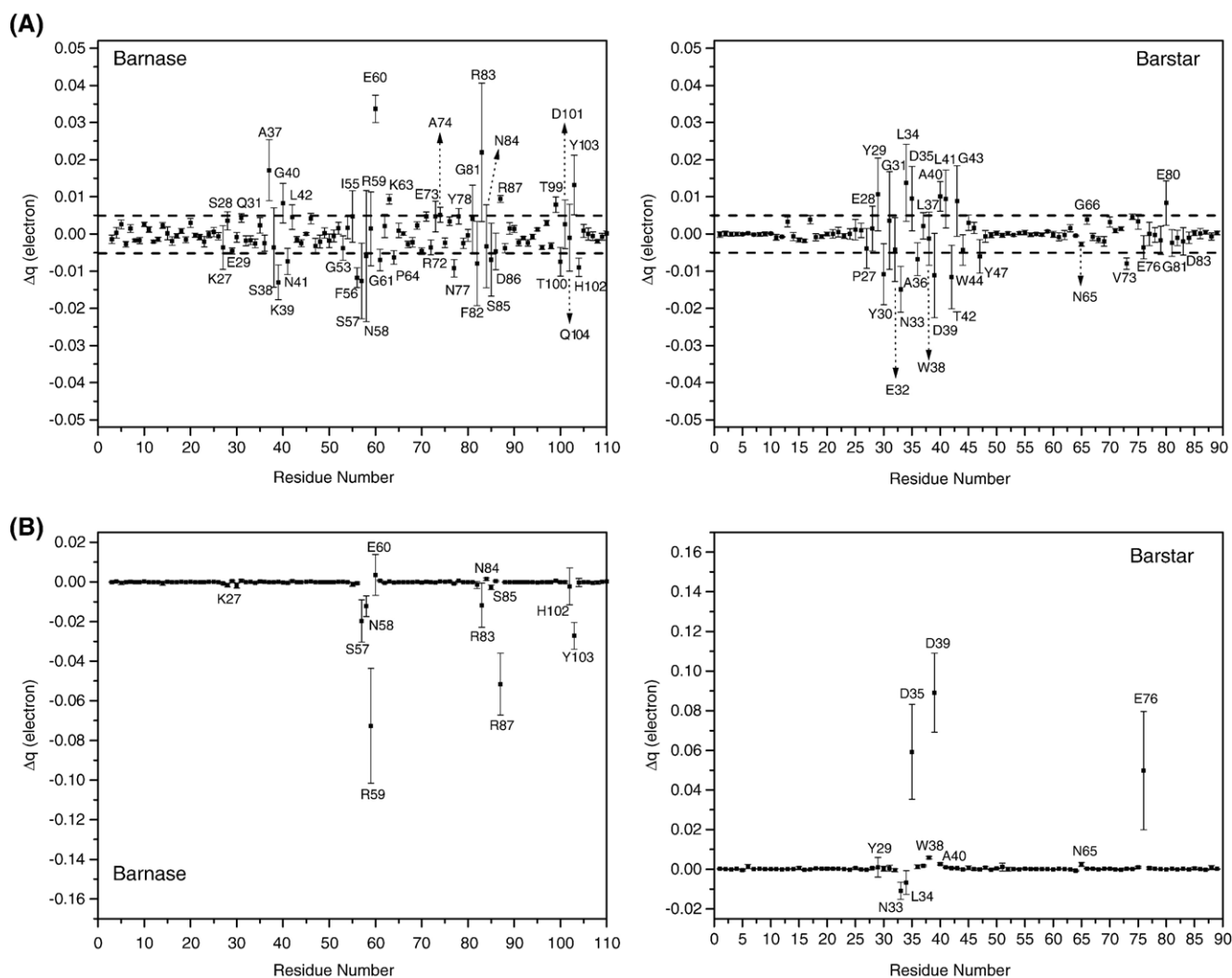
Fig. 5. The effect of polarization (A) and charge transfer (B) on the charge distribution of the barnase–barstar complex with the AM1 Hamiltonian. Shown is the difference in calculated CM2 charges when: (A) polarization is added to electrostatics and (B) charge transfer is added to polarization + electrostatics. Charges are listed by residue numbers. The bars indicate the standard deviations from the five snapshots.

these residues reveals that both residues are at the interface of the complex and, in particular, they are at 5.6 Å (centroid-to-centroid) distance from each other. While the carboxyl groups of these residues do not come closer than 6.0 Å a strong H-bond E60(H)–D35(OD2) is formed which persists during the MD simulation ($d = 2.0$ Å and $\theta = 150.1°$). This is an example of how strongly the environment influences the polarization of certain residues. Interesting examples are F56 or F82 in barnase, which both gain electrons, while F74 in barstar loses electrons. Close inspection of the local environment of the latter shows to be surrounded by hydrophobic residues: V70, L71, A77 and I84 of barstar at distances less than 5.9 Å (centroid-to-centroid). F82 is near G81 (∼4.4 Å) and no other residue at distances less than 6.0 Å. F56 is surrounded by S58 and N59 (less than ∼5.0 Å), W71 (∼5.4 Å) and E73 (∼5.8 Å). This comparison suggests that Phe's in hydrophobic environments lose electrons, while in polar environments they gain electrons. This clearly indicates how the polarization effect is greatly influenced by the characteristics of the local environment.

Figs. 5A and 6A show that the polarization effect on the charge distribution of barnase is evenly distributed among the protein residues with a particularly strong effect on those residues located at the interface of the complex and hence interacting with barstar residues. In the case of barstar, the effect of polarization has a localized effect on certain residues, while on the others there is only small effect. Consequently, two major sequence regions of barstar protein were identified. The first region includes residues between P27 and Y47 and the second region has residues between N65 and D83. Although the X-ray structure of the complex has revealed the important regions of the protein monomers involved in complex formation, with our interaction energy decomposition scheme (in particular the polarization effect), it is possible to clearly identify such regions as shown in Figs. 5A and 6A, especially in the case of barstar. The identification of these two regions supports the reported works showing some "hot spots" residues in the same regions, using a classical continuum electrostatic model [57], as well as hybrid QM/MM strategy [63].
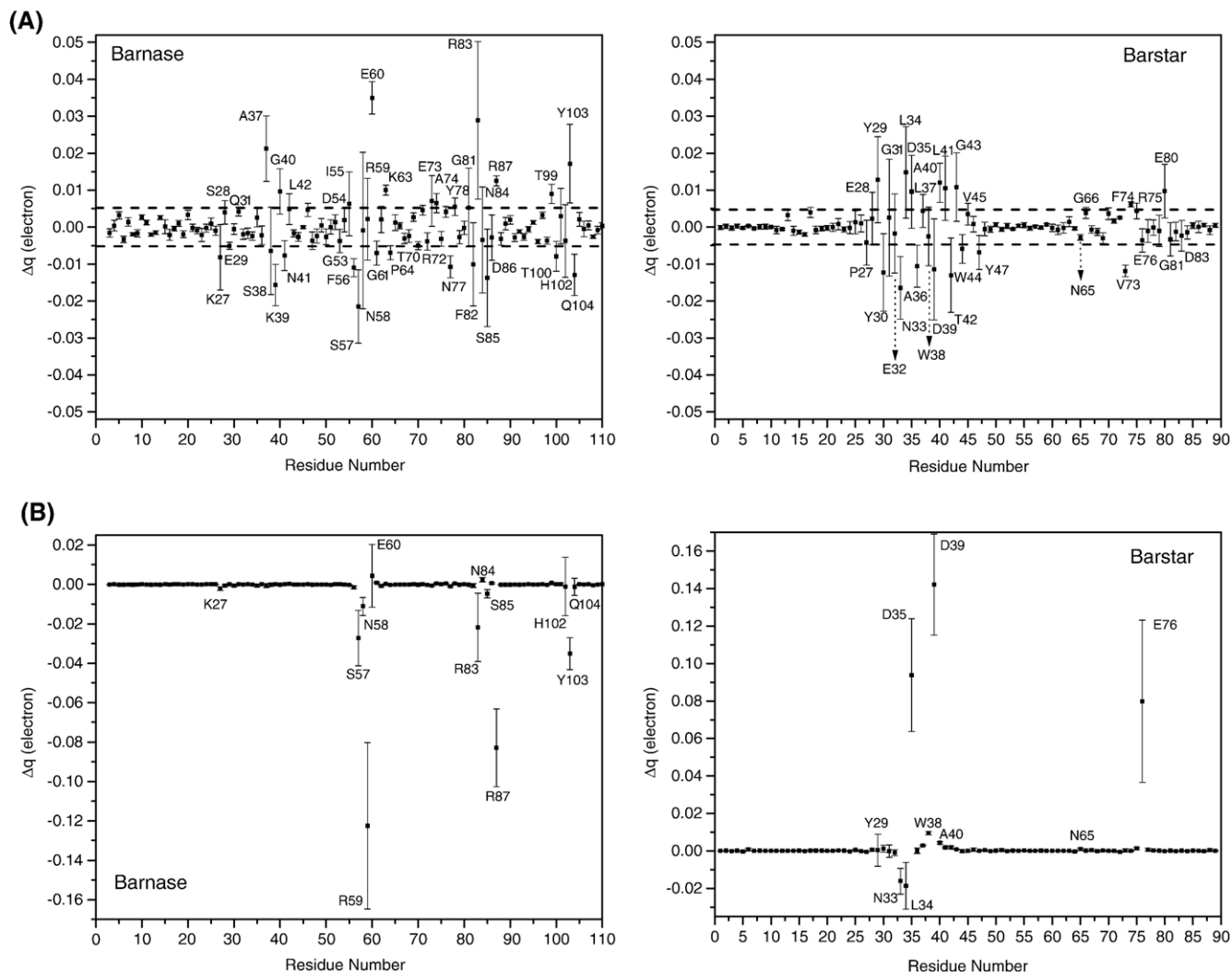
Fig. 6. The effect of polarization (A) and charge transfer (B) on the charge distribution of the barnase–barstar complex with the PM3 Hamiltonian. Shown is the difference in calculated CM2 charges when: (A) polarization is added to electrostatics and (B) charge transfer is added to polarization+electrostatics. Charges are listed by residue numbers. The bars indicate the standard deviations from the five snapshots.

2.4. The charge-transfer effect

Figs. 5B and 6B show the effect of CT on the charge distribution of the complex calculated with AM1 and PM3, respectively. The bars represent the standard deviation (SD) in the partial charges of one residue in the five snapshots. The charge-transfer effect has less impact on the charge distribution of barnase and barstar than polarization. For CT, the boundaries on the charges within which most residues appear are between −0.00075 and 0.00075 (not shown). Positive CT effect means that residues lose electrons and negative CT effect means a gain of electrons. In general for negatively charged residues, positive CT effect occurs, while for positively charged residues the CT effect is negative. Figs. 5B and 6B show that charged residues are the source of the largest CT between monomers. However, about half of the residues are hydrophilic (S57, N58, N84, S85, Y103, Q104) in barnase, and two hydrophilic (Y29, N33) and three hydrophobic (L34, W38, A40) in barstar. The CT effect on the charge distribution of monomers is mainly at the interface,

and close inspection of those residues reveals that they are involved in H-bonds between barnase and barstar except W38, which makes H-bonds with L34 and L41 in barstar, respectively.

2.5. Polarization and charge-transfer effects by residue type

To further analyze the effect of polarization and CT on the charge distribution of the barnase–barstar complex, we calculated their average effects for each residue type as shown in Tables 3 and 4. For barnase, the average polarization shows a strong effect on the negatively charged and hydrophilic groups. Instead, CT affects the positively charged groups as well as hydrophilic groups. For barstar, the polarization, on average, affects the hydrophilic groups, while the CT affects almost exclusively the negatively charged groups. As expected from the charge complementarity at the interface between barnase and barstar, the average flow of charge takes places between negatively and positively charged groups with the further contribution of hydrophilic groups in barnase.

Table 3
Effect of polarization and charge transfer on the charge distribution of barnase

Barnase	Residue[a]	Polarization effect[b]		Charge transfer effect[b]	
		AM1	PM3	AM1	PM3
Negative	D	−0.0013±0.0012	−0.0013±0.0014	0.0000±0.0001	0.0000±0.0001
	E60	0.0337±0.0037	0.0350±0.0044	0.0035±0.0103	0.0044±0.0159
	E	0.0113±0.0021	0.0123±0.0032	0.0012±0.0035	0.0013±0.0052
	All[c]	0.0018±0.0012	0.0021±0.0014	0.0003±0.0009	0.0004±0.0013
Positive	V3(+)	−0.0014±0.0011	−0.0016±0.0011	−0.0001±0.0001	−0.0001±0.0001
	K27	−0.0037±0.0059	−0.0082±0.0088	−0.0008±0.0004	−0.0021±0.0010
	K	−0.0026±0.0006	−0.0036±0.0011	−0.0002±0.0001	−0.0005±0.0001
	R59	0.0014±0.0099	0.0021±0.0110	−0.0726±0.0290	−0.1225±0.0423
	R83	0.0220±0.0186	0.0289±0.0213	−0.0117±0.0112	−0.0219±0.0174
	R87	0.0094±0.0009	0.0125±0.0013	−0.0516±0.0156	−0.0829±0.0198
	R	0.0063±0.0024	0.0085±0.0028	−0.0272±0.0074	−0.0455±0.0102
	All[c]	0.0007±0.0012	0.0008±0.0015	−0.0099±0.0027	−0.0165±0.0037
Hydrophilic	H102	−0.0010±0.0090	−0.0037±0.0098	−0.0022±0.0093	−0.0011±0.0148
	H	−0.0001±0.0044	−0.0014±0.0047	−0.0011±0.0047	−0.0005±0.0074
	N58	0.0060±0.0176	0.0009±0.0212	−0.0122±0.0052	−0.0111±0.0046
	N	−0.0042±0.0021	−0.0036±0.0023	−0.0021±0.0010	−0.0017±0.0009
	Q104	−0.0090±0.0025	−0.0129±0.0056	−0.0003±0.0021	−0.0013±0.0043
	Q	−0.0015±0.0013	−0.0029±0.0023	0.0001±0.0008	−0.0004±0.0015
	S57	−0.0127±0.0102	−0.0215±0.0099	−0.01924±0.0107	−0.0272±0.0140
	S85	−0.0069±0.0098	−0.0137±0.0132	−0.0027±0.0011	−0.0047±0.0021
	S	−0.0031±0.0024	−0.0053±0.0028	−0.0027±0.0012	−0.0037±0.0016
	T	−0.0013±0.0004	−0.0012±0.0004	−0.0001±0.0001	0.0000±0.0000
	Y103	0.0132±0.0080	0.0172±0.0106	−0.0272±0.0067	−0.0351±0.0082
	Y	0.0029±0.0012	0.0037±0.0016	−0.0039±0.0010	−0.0050±0.0012
	R110(±)	0.0002±0.0006	0.0003±0.0009	0.0001±0.0002	0.0001±0.0001
	All[c]	−0.0013±0.0005	−0.0018±0.0005	−0.0018±0.0004	−0.0022±0.0005
Hydrophobic	A37	0.0172±0.0082	0.0212±0.0088	−0.0002±0.0002	−0.0006±0.0010
	A	0.0034±0.0016	0.0042±0.0018	−0.0003±0.0002	−0.0001±0.0002
	F56	−0.0118±0.0026	−0.0110±0.0024	−0.0006±0.0005	−0.0014±0.0007
	F82	−0.0079±0.0113	−0.0101±0.0113	−0.0015±0.0018	−0.0006±0.0010
	F	−0.0045±0.0026	−0.0050±0.0029	−0.0006±0.0005	−0.0005±0.0002
	G	−0.0005±0.0016	−0.0002±0.0018	−0.0001±0.0000	0.0000±0.0000
	I	0.0001±0.0007	0.0005±0.0009	−0.0001±0.0002	0.0000±0.0001
	L	0.0029±0.0007	0.0030±0.0007	−0.0001±0.0002	0.0000±0.0001
	P	−0.0033±0.0006	−0.0036±0.0006	−0.0002±0.0001	−0.0002±0.0001
	V	0.0000±0.0007	0.0001±0.0008	0.0001±0.0000	0.0001±0.0000
	W	0.0021±0.0007	0.0022±0.0006	0.0001±0.0001	−0.0002±0.0002
	All[c]	0.0004±0.0003	0.0006±0.0003	−0.0002±0.0001	−0.0001±0.0001

[a] Amino acids are indicated by their one-letter symbols.

[b] Average CM2 charge (in electron) with ±S.D. Averages are taken over all snapshots for the listed residues. The important residues at the interface are listed. A negative sign means electrons are gained, while a positive sign indicates that electrons are lost.

[c] Average of all residues of the particular type over all snapshots.

In Table 5, the calculated percentage of CT effect shows that the negatively charged groups are the source of most of CT from barstar to barnase, while it is the opposite in the case of barnase to barstar, namely positively charged groups are the source for the CT. Thus, there is a net flow of electrons from barstar to barnase as a result of the CT effect. The total amount of charge transferred from barstar to barnase is ∼0.21 electrons for AM1 and ∼0.31 electrons for PM3 (Table 6). In the case of barnase, the hydrophilic groups have substantial contribution to CT of 11.3% (PM3) and 15% (AM1). The CT effect is localized within the two regions (see above) in barstar, while barnase again shows a distributed CT effect throughout the interface with barstar. This suggests that barstar was highly optimized to specifically inhibit the RNase activity of barnase, using few res-

idues that are ultimately affected by polarization and CT effects while the rest of the protein residues do not (Figs. 5 and 6).

3. Discussion

The interaction energy decomposition scheme applied here to study barnase–barstar complex association revealed a major and important favorable contribution of electrostatics in the formation of this complex (Table 2). However, it has been reported that the electrostatic binding free energy in barnase–barstar complex, based on continuum solvent model, is unfavorable [57], a near zero contribution [64] and a favorable contribution when high protein dielectric constant was used [65]. These contradictory results may be attributed mainly to the lack

Table 4
Effect of polarization and charge transfer on the charge distribution of barstar

Barstar	Residue[a]	Polarization effect[b]		Charge transfer effect[b]	
		AM1	PM3	AM1	PM3
Negative	D35	0.0095±0.0086	0.0096±0.0099	0.0592±0.0240	0.0938±0.0301
	D39	−0.0112±0.0114	−0.0115±0.0137	0.0891±0.0200	0.1421±0.0270
	D	−0.0013±0.0034	−0.0014±0.0039	0.0373±0.0019	0.0590±0.0020
	E76	−0.0036±0.0031	−0.0036±0.0032	0.0497±0.0298	0.0799±0.0433
	E80	0.0084±0.0060	0.0097±0.0072	−0.0000±0.0000	−0.0001±0.0001
	E	0.0001±0.0011	0.0005±0.0013	0.0045±0.0024	0.0072±0.0037
	S89(−)	0.0003±0.0005	0.0005±0.0005	0.0001±0.0002	0.0001±0.0001
	All[c]	−0.0002±0.0014	0.0000±0.0016	0.0124±0.0014	0.0197±0.0022
Positive	K	0.0001±0.0003	0.0001±0.0004	0.0001±0.0001	0.0001±0.0000
	R	0.0009±0.0009	0.0013±0.0010	0.0004±0.0001	0.0005±0.0002
	All[c]	0.0004±0.0005	0.0006±0.0006	0.0002±0.0001	0.0003±0.0001
Hydrophilic	K1(+)	−0.0002±0.0007	−0.0001±0.0006	0.0002±0.0002	0.0000±0.0000
	H	0.0039±0.0011	0.0040±0.0014	−0.0001±0.0001	0.0002±0.0003
	N33	−0.0149±0.0062	−0.0165±0.0085	−0.0109±0.0043	−0.0161±0.0069
	N	−0.0058±0.0021	−0.0063±0.0028	−0.0024±0.0013	−0.0048±0.0024
	Q	0.0000±0.0004	0.0003±0.0004	0.0001±0.0001	0.0000±0.0001
	S	−0.0007±0.0004	−0.0010±0.0004	0.0001±0.0000	0.0001±0.0000
	T42	−0.0116±0.0085	−0.0131±0.0101	0.0005±0.0003	0.0019±0.0007
	T	−0.0026±0.0021	−0.0029±0.0024	0.0003±0.0002	0.0006±0.0002
	Y29	0.0106±0.0098	0.0128±0.0116	0.0010±0.0050	0.0004±0.0086
	Y	−0.0021±0.0044	−0.0021±0.0049	0.0004±0.0020	0.0005±0.0033
	All[c]	−0.0015±0.0008	−0.0016±0.0009	−0.0002±0.0004	−0.0005±0.0007
Hydrophobic	A	0.0001±0.0009	−0.0002±0.0012	0.0007±0.0001	0.0007±0.0002
	A40	0.0101±0.0040	0.0120±0.0053	0.0027±0.0005	0.0043±0.0006
	F	0.0022±0.0005	0.0029±0.0005	0.0001±0.0001	0.0000±0.0001
	G31	0.0036±0.0132	0.0026±0.0158	0.0005±0.0015	−0.0001±0.0033
	G43	0.0089±0.0095	0.0108±0.0093	0.0005±0.0001	0.0008±0.0001
	G	0.0028±0.0041	0.0027±0.0043	0.0003±0.0003	0.0002±0.0006
	I	0.0005±0.0003	0.0004±0.0003	−0.0001±0.0001	−0.0001±0.0001
	L34	0.0138±0.0104	0.0148±0.0124	−0.0068±0.0060	−0.0186±0.0125
	L	0.0020±0.0014	0.0023±0.0016	−0.0002±0.0004	−0.0012±0.0010
	P	−0.0023±0.0027	−0.0024±0.0028	0.0003±0.0003	0.0000±0.0002
	V	−0.0003±0.0005	−0.0008±0.0007	0.0004±0.0003	0.0001±0.0002
	W38	−0.0013±0.0072	−0.0025±0.0079	0.0058±0.0007	0.0094±0.0008
	W44	−0.0043±0.0041	−0.0059±0.0039	0.0000±0.0003	−0.0002±0.0007
	W	−0.0020±0.0028	−0.0030±0.0032	0.0020±0.0003	0.0031±0.0003
	All[c]	0.0008±0.0004	0.0007±0.0005	0.0003±0.0001	0.0001±0.0003

[a] Amino acids are indicated by their one-letter symbols.
[b] Average CM2 charge (in electron) with ± S.D. Averages are taken over all snapshots for the listed residues. The important residues at the interface are listed. A negative sign means electrons are gained, while a positive sign indicates that electrons are lost.
[c] Average of all residues of the particular type over all snapshots.

of inclusion of polarization contribution in standard force fields that use atomic charges as conformation and environment configuration independent. Indeed, recent study, using QM/MM protocol to obtain self-consistent point charge model of the protein electrostatic potential, has shown the importance of polarization contribution in the electrostatic potential determination in proteins [63]. Our result, revealing the importance of favorable electrostatics in the formation of this complex, has been obtained using semi-empirical QM calculation and hence shows the importance in using such methodologies, whenever it is possible, to overcome the problem associated with polarization contribution in standard MM force fields.

The interaction energy decomposition results using AM1 and PM3 Hamiltonians show similar trends; however, the magnitude of CT effects is different between these Hamiltonians. This difference reaches 0.07 electrons for some residues (Table 3 for positive groups and Table 4 for negative groups), but the precise origin of this effect is not clear [47]. However, plausible origin for this effect might be caused by the difference in the hydrogen bonding geometry preference for these Hamiltonians. AM1 favors bifurcated structures with nonlinear ŐH and ŃH hydrogen bonds, while PM3 favors linear hydrogen bonds [66,67].

As reported for the CspA–water system [47], CT occurs between residues that are hydrogen bonded with one another and, hence, introduce a covalent character to this type of interaction. Recently, different groups have established this experimentally [68–71]. However, the establishment of the partial covalent character of the H-bond interaction is due to Compton-scattering experiments on ice crystals [72]. The CT effect is QM in nature and results in the sharing of electrons between a pair of molecules that are hydrogen bonded. Biological systems (proteins, DNA, RNA, etc.) use H-bonds to build up

Table 5
Average charge transfer effects per residue type

Residue type[a]	Barnase		Barstar	
	AM1 (%)[b]	PM3 (%)[b]	AM1 (%)[b]	PM3 (%)[b]
Negative	2.2	1.9	95.0	96.2
Positive	81.6	86.4	1.7	1.3
Hydrophilic	15.0	11.3	1.2	2.2
Hydrophobic	1.2	0.4	2.1	0.3

[a] See Tables 3 and 4 for definition of the residue type.
[b] Percentage of the total CT effect. This percentage is defined as $100 \sum_{i \in \text{type X}} |CT_i| / \left(\sum_{X}^{\text{all types}} \sum_{i \in \text{type X}} |CT_i| \right)$ where CT_i is the charge transfer effect of residue i.

their tertiary structures, to facilitate inter/intramolecular association, to affect catalysis and to interact with the surrounding solvent. Thus, the QM treatment of biological systems is necessary to explore CT effects, which may have important implication for their function. Classical methods (MM and continuum electrostatic model) cannot presently quantify such QM processes.

While the accurate evaluation of binding free energies is important, we have focused on exploring the driving forces in the association of two proteins (the barnase–barstar complex). Classical methods for treating protein–protein interactions take into account only electrostatic and hydrophobic contributions (for a review see [73]). Our QM interaction energy decomposition scheme reveals that polarization and CT have a substantial contribution to the barnase–barstar association energy. These contributions (polarization or CT) have not been previously considered in any interaction energy calculation between proteins. Obviously, this is due to the QM nature of CT and the lack of well polarized force fields in the molecular mechanics methods that have been used. By investigating the polarization and CT effects, we were able to reveal the interfacial residues that are involved in strong polarization and CT interactions. Indeed, we found many of the residues that have been experimentally reported to be important for the complex formation, such as K27, R59, E60, E73, R83, R87 and H102 in barnase, and Y29, D35 W38, D39, T42, W44, E80 and E76 in barstar (Figs. 5 and 6) [53,54,74,75], were affected by polarization and CT interactions. For example, in barstar, N33 was affected by polarization and CT, while V73 only had a significant polarization effect, suggesting they are important in the formation of the complex. These residues were

Table 6
Charge transfer from barstar to barnase

Time (ps)	AM1[a]	PM3[a]
100	0.207	0.317
200	0.185	0.276
300	0.199	0.297
400	0.229	0.334
500	0.223	0.328
Average	0.209	0.311
S.D.	0.016	0.021

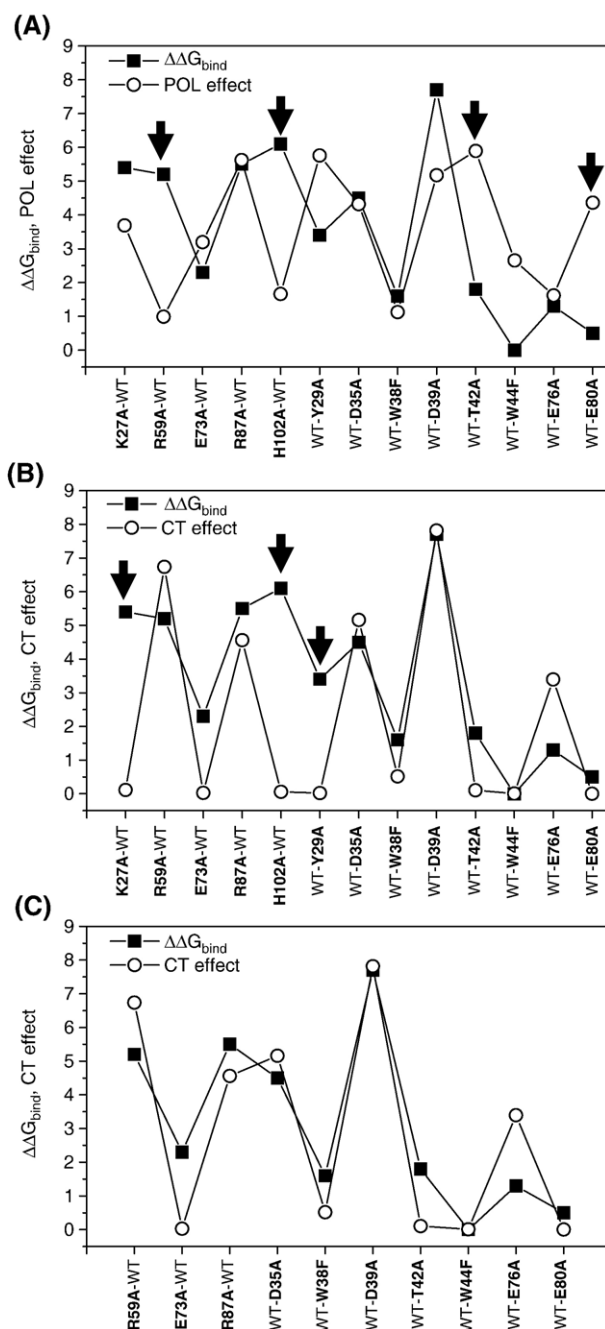[a] CM2 charges in electrons. Barnase has accepted a net positive charge.

Fig. 7. The comparison of experimental binding free energy between the WT barnase–barstar complex and its single mutants ($\Delta\Delta G_{\text{bind}}$) and scaled (A) polarization and (B) CT effects (see Methods for details). The arrows indicate where the comparison is not satisfactory. (C) An example using CT effect, if the unsatisfactory matches have been excluded then very good agreement is achieved. The complexes are named by the mutation in the order of barnase followed by barstar, and the wild type is indicated by 'WT'.

reported to be important for the binding between barnase and barstar, using electrostatic continuum model [57]. Unfortunately, no experimental studies have been reported for these two residues in this complex and hence no comparison with our results is possible yet. Furthermore, to investigate if there is any correlation between our QM findings and the destabilizing effect of these known residues on the experimental binding free

energies ($\Delta\Delta G_{bind}$), we have compared $\Delta\Delta G_{bind}$ with the polarization and CT effects on these residues. While $\Delta\Delta G_{bind}$ and polarization or CT are two different physico-chemical proprieties (kcal/mol and electron unit, respectively), such comparison reveals an interesting aspect of our QM methodology. Fig. 7 shows this comparison, using PM3 Hamiltonian as an example, where the polarization and CT effects have been scaled to match the range of $\Delta\Delta G_{bind}$ (see details in Methods). Except for few residues (see arrows in Fig. 7), the comparison shows very good agreement (Fig. 7C for the CT effect), which suggests that our QM methodology not only can predict the important residues in protein–protein formation, but to some extent it may be able to predict the extent of the impact of these residues in the protein complexes formation. Note that if polarization and CT effects are combined the impact of these residues on the protein complex formation can be well captured. Fig. 7A and B show that what is well captured with polarization effect is not with CT effect and vice versa.

Our findings demonstrate the usefulness of this interaction energy decomposition scheme in revealing key residues in the interaction between barnase and barstar. Consequently, we can predict that other residues, never investigated experimentally, may be important for binding (thermodynamically and/or kinetically). Thus, other residues, which may have an important role to play in the association of these proteins through both polarization and CT effects are S57, N58, S85 and Y103 in barnase, and N33, L34 and A40 in barstar. On the basis of only the polarization effect, we suggest the following residues: A37, K39, G40, N41, F56, K63, N77, F82, T99 and T100 in barnase, and Y30, G31, E32, A36, L37, L41, G43, Y47, V73, G81 and D83 in barstar.

In this work, we did not include in our QM calculations the water molecules at the interface between barnase and barstar as it has been reported in their X-ray structure (pdb code 1brs). We used MD snapshots in our analysis of protein–protein interactions in order to take into account structural fluctuations [76], and the effect of water molecules, that pass through the interface during the MD simulation, on residues located at the interface between the proteins. However, to test the impact of water molecules located at the interface on the interaction energy and polarization and CT effects, we have included six deeply buried water molecules that are bridging the two proteins via a H-bonding network
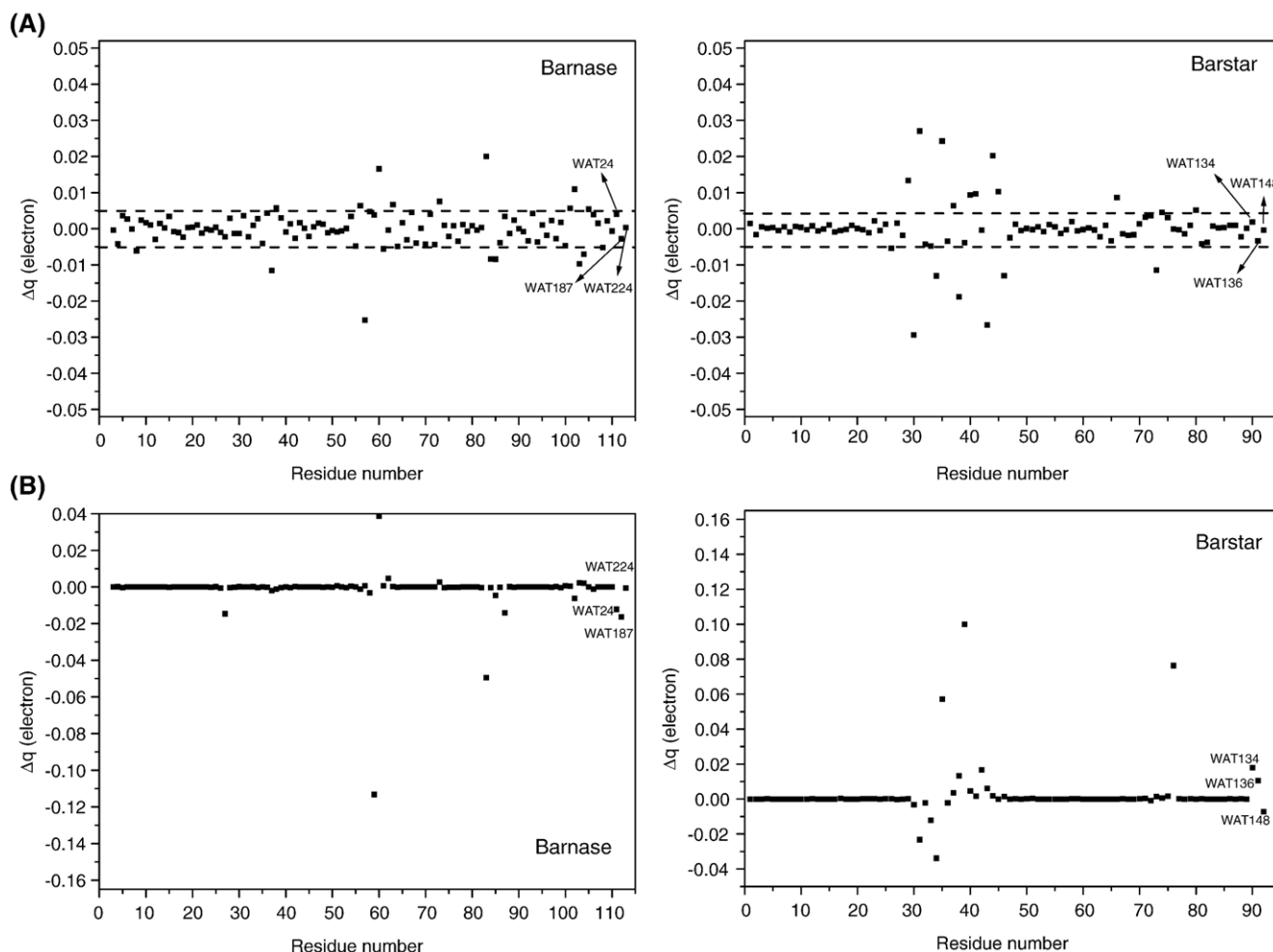

Fig. 8. The effect of polarization (A) and charge transfer (B) on the charge distribution of the barnase–barstar complex using PM3 Hamiltonian and including six water molecules with their original pdb numbering (see Methods for details). The figure shows the difference in calculated CM2 charges when: (A) polarization is added to electrostatics and (B) charge transfer is added to polarization+electrostatics. Charges are displayed by residue numbers.

(for details see Methods). The interaction energy decomposition revealed a similar trend in the energy contributions: $E_{ES} \sim 73\%$ (AM1) and $\sim 60\%$ (PM3), $E_{POL} \sim 5.3\%$ (AM1) and 5.1% (PM3), and $E_{CT} \sim 21.7\%$ (AM1) and $\sim 34.9\%$ (PM3). Note that for the interaction energy a small increase was reported, E_{INT} − 519 kcal/mol (AM1) and −544 kcal/mol (PM3) when compared with E_{INT} from the MD simulation (Table 2). The polarization and CT effects on the charge distribution of the complex are summarized in Fig. 8A and B, respectively, using PM3 Hamiltonian as an example. Fig. 8 shows a similar trend in polarization and CT effects for the residues at the interface (Figs. 5 and 6), with small changes in CT effect for some residues, such as K27. However, the polarization effect is small for these water molecules, while the CT effect is slightly higher as these water molecules are making a H-bonding network between the proteins at the interface. This clearly suggests that water molecules at the interface contribute little to the interaction energy between the proteins as revealed with our QM methodology, but with no substantial effect on interaction energy contributions as well as the polarization and CT effects on the charge distribution of the system. Instead, for protein–small ligand complexes, we expect that the water molecules located at the interface will have a higher impact.

4. Conclusions

Our MD simulation of barnase–barstar shows that the complex is stable along the calculated trajectory and that the majority of the H-bonds at the interface of the complex are strong (>90% present during the entire trajectory). Our QM calculations show that electrostatics dominates and stabilizes the interaction between barnase and barstar. Charge transfer is important in this interaction because it stabilizes the complex and affects the charge distribution of the system at the protein–protein interface (residues involved in the binding). The charge transfer takes place between residues connected through hydrogen bonds. However, while polarization contributes less to the interaction energy between barnase and barstar its effect on the charge distribution of the complex is large. Furthermore, our calculations reveal the existence of two regions in barstar, which were strongly affected by polarization and charge transfer, while the rest of the protein was virtually unaffected. For barnase, this is not the case, because these effects were distributed all over the protein, especially for polarization effect. The existence of these two regions in barstar implies that the latter is well optimized to specifically inhibit barnase activity, since specific and few residues play major role in the interaction with barnase.

The interaction energy decomposition scheme used in this work identifies most of the experimentally reported residues that are important for the binding between barnase and barstar. The key aspect of the methodology is not only detecting the important residues but the separation of the effects of the polarization and charge transfer on the charge distribution of the complex. This type of analysis, applied to the protein–protein complex (barnase–barstar complex as an example), is presented here for the first time. The analysis clearly shows the importance of these effects in protein–protein associations and, hence, for other biological systems as well (i.e., DNA–protein and RNA–protein).

The possibility of analyzing separately these effects is a step forward towards a better understanding of these interactions. Our QM methodology has the potential of quantifying the impact of the residues at the interface on the binding free energy of protein complexes using polarization and CT effects. Furthermore, our QM methodology shows that the electronic structure of water molecules located at the interface of barnase–barstar complex does not affect dramatically the polarization and CT effects due to protein–protein interaction; however, we expect that the effect of these water molecules on the interaction of proteins with small ligands to be more substantial.

Finally, our calculations have another important aspect that may be helpful in drug design of small organic compounds and peptides aimed at inhibiting protein/protein associations. The ultimate goal of the investigation of protein–protein interaction energies is not only limited to the calculation of accurate binding free energies, but rather to the exploration of either inhibiting or enhancing protein–protein interactions (i.e., using in silico mutagenesis to study mutation effects on complexation). Hence, via the use of this energy decomposition scheme to map the individual energy contribution to the interactions at the protein–protein interface and precisely identify the key residues in the complex formation may allow the design of novel protein or small molecule inhibitors that take advantage of *hot spots* at the protein–protein interface.

5. Methods

5.1. Theory

For completeness and clarity, we first briefly review the DC semi-empirical scheme for the decomposition of the interaction energy [49]. In the DC method, the system of interest is divided into subsystems. Since density matrix elements between atoms of two subsystems are only different from zero when these subsystems overlap, charge flow between subsystems only occurs when the subsystems overlap. For the case of two proteins A and B (AB is the complex formed by A and B), placed in different subsystems, the overlap between the A and B subsystems can be written as (A:AB, B:AB), which indicates that the density matrix elements between subsystems are not zero. The total number of electrons is constrained by the Fermi energy ε_F. Thus, the interaction energy between A and B is given by

$$E_{INT} = E[\varepsilon_F, (A:AB, B:AB)] \\ -(E[\varepsilon_F^A, (A:A)] + E[\varepsilon_F^B, (B:B)]) \tag{1}$$

The second term can be obtained with one calculation, by infinitely separating A and B and employing two separate Fermi energies for each system

$$E_{INT} = E[\varepsilon_F, r, P(r), (A:AB, B:AB)] \\ -E[(\varepsilon_F^A, \varepsilon_F^B), \infty, P(\infty), (A:A, B:B)] \tag{2}$$

Here (A:A, B:B) indicates that the two subsystems do not overlap, so that the density matrix elements between the two subsystems are zero. Energies are now shown as explicit

functions of the A–B separation (r) and the density matrix at this separation ($P(r)$). The electrostatic energy can be obtained by bringing the infinitely separated solute and solvent subsystems to the equilibrium distance, without relaxation of the density matrix:

$$E_{ES} = E[(\varepsilon_F^A, \varepsilon_F^B), r, P(\infty), (A{:}A, B{:}B)] \\ - E[(\varepsilon_F^A, \varepsilon_F^B), \infty, P(\infty), (A{:}A, B{:}B)] \tag{3}$$

By allowing charge flow within the A and B subsystems only, the polarization energy is obtained:

$$E_{POL} = E[(\varepsilon_F^A, \varepsilon_F^B), r, P(r), (A{:}A, B{:}B)] \\ - E[(\varepsilon_F^A, \varepsilon_F^B), r, P(\infty), (A{:}A, B{:}B)] \tag{4}$$

Finally, the charge-transfer energy is obtained by allowing electrons to flow between A and B:

$$E_{CT} = E[\varepsilon_F, r, P(r), (A{:}AB, B{:}AB)] \\ - E[(\varepsilon_F^A, \varepsilon_F^B), r, P(r), (A{:}A, B{:}B)] \tag{5}$$

Thus, the interaction energy between the two proteins in the complex is:

$$E_{INT} = E_{ES} + E_{POL} + E_{CT} \tag{6}$$

5.2. MD simulation

We have used the complex C–F of barnase–barstar from the protein databank (1brs) [48] where all residues were present except the two first residues of barnase. The missing atoms in residues K19, D22, E29, Q31, K39, V45, K49, S67 and R110 of barnase, and K22, E28, E46, E64, N65 and S89 of barstar were added using SPA. SPA inserts the best rotamer of the residue of interest (scored by an energy function similar to the AMBER force field) by keeping the best energies for side chain-backbone and side chain–side chain interactions [77]. Then the complex was minimized for the newly restored residues while keeping the rest of the residues fixed (using a harmonic potential constraint with a constant force of 50 kcal/mol/Å^2) in order to insure that the optimized structure does not diverge significantly from the crystal structure. The MD simulation of the barnase–barstar complex was performed with TIP3P [78] water using the AMBER5 program [79]. The time step was 1.5 fs with bonds fixed by SHAKE [80], the formal charge of the complex of $-4e$ was neutralized by adding Na$^+$ ions, the temperature was 300 K, van der Waals interactions were truncated at 10.0 Å, while electrostatic interactions were fully calculated with the Particle Mesh Ewald method [81]. After performing energy minimization only on the water molecules, 50 ps of solvent equilibration at 300 K were calculated after including the Na$^+$ ions. Following equilibration, a trajectory of 500 ps was calculated.

5.3. Interaction energy decomposition in barnase–barstar complex

Snapshots of 100 ps, 200 ps, 300 ps, 400 ps and 500 ps were obtained from the MD simulation of barnase–barstar complex.

Density matrixes $P(\infty)$ and $P(r)$ were constructed from the density matrix of the separate proteins and from the density matrix of the complex in a continuum (with dielectric constant of 80), followed by one SCF calculation. All calculations were performed with the semi-empirical AM1 [82] or PM3 [83,84] Hamiltonian as implemented in the DivCon program [51]. The dielectric continuum calculations were performed with a modification of the DivCon program [85], which includes the Delphi program [41] using CM2 charges [86]. A cutoff of 8.0 Å was used for the off-diagonal elements of the Fock, 1-electron and density matrixes.

5.4. Charge distribution analysis

The polarization effect was calculated from the difference between partial charges obtained from the calculation of $E[(\varepsilon_F^A, \varepsilon_F^B), r, P(r), (A{:}A, B{:}B)]$ and those obtained from the calculation of $E[(\varepsilon_F^A, \varepsilon_F^B), r, P(\infty), (A{:}A, B{:}B)]$. The charge-transfer effect was obtained from the difference between partial charges from the calculation of $E[\varepsilon_F, r, P(r), (A{:}AB, B{:}AB)]$ and $E[(\varepsilon_F^A, \varepsilon_F^B), r, P(r), (A{:}A, B{:}B)]$. Negative polarization and charge-transfer effects indicate gain of electrons, and positive polarization and charge-transfer effects indicate loss of electrons. CM2 [86] charges rather than Mulliken charges were used, since CM2 charges have been shown to provide a better description of the dipole moments for small molecules [86].

5.5. Water molecules treatment and polarization and charge-transfer effects quantification

To test the influence of water molecules at the interface between barnase and barstar on our QM interaction energy decomposition scheme, we have explicitly included six water molecules from the complex CF (pdb code 1brs). The choice of these water molecules is based on simple criteria; water molecules with no solvent accessible surface area are bridging both proteins with H-bonds, or indirectly with another water molecule. Since our QM treats a system with only two subsystems, A and B, we included in the QM calculations three water molecules with each protein (two as directly bridging both proteins and one as indirectly bridging one of the proteins and a water molecule). To compare the difference in the experimental binding free energy ($\Delta\Delta G_{bind}$) between the WT barnase–barstar complex and its single mutants [74,75] and the polarization and CT effects using the PM3 Hamiltonian as an example (a similar trend is found with AM1), we have proceeded as follows: because all calculated $\Delta\Delta G_{bind}$ were positive, we used the absolute values of either polarization or CT effect and scaled them (by a simple multiplicative factor of 455 and 55 respectively) to be able to compare them to $\Delta\Delta G_{bind}$.

References

[1] R.O. Hynes, Integrins: versatility, modulation, and signalling in cell adhesion, Cell 69 (1992) 11–25.

[2] M.D. Welch, The world according to Arp: regulation of actin nucleation by the Arp2/3 complex, Trends Cell Biol. 9 (1999) 423–427.

[3] P.A. Boriack-Sjodin, S.M. Margarit, D. Bar-Sagi, J. Kuriyan, The structural basis of the activation of Ras by Sos, Nature 394 (1998) 337–343.

[4] T. Hunter, Oncoprotein networks, Cell 88 (1997) 333–346.

[5] M. Berridge, Inositol trisphosphate and calcium signalling, Nature 361 (1993) 124–315.

[6] F.3. Andel, A.G. Ladurner, C. Inouye, R. Tjian, E. Nogales, Three-dimensional structure of the human TFIID-IIA-IIB complex, Science 286 (1999) 2153–2156.

[7] M. Enari, H. Sakahira, H. Yokoyama, K. Okawa, A. Iwamatsu, S. Nagata, A caspase-activated DNAase that degrades DNA during apoptosis, and its inhibitor ICAD, Nature 391 (1998) 43–50.

[8] H. Neurath, Proteolytic processing and physiological regulation, Trends Biochem. Sci. 14 (1989) 268–271.

[9] D.R. Madden, D.N. Garboczi, D.C. Wiley, The antigenic identity of peptide-MHC complexes: a comparison of conformations of five viral peptides presented by HLA-A2, Cell 75 (1993) 693–708.

[10] S.-J. Tech, N. Killeen, A. Tarakhovsky, D.R. Littman, H.-S. Teh, Blood 89 (1997) 1308–1318.

[11] S. Koyasu, T. Lawton, D. Novick, M.A. Recny, R.F. Siliciano, B.P. Wallner, E.L. Reinherz, Proc. Natl. Acad. Sci. U. S. A. 87 (1990) 2603–2607.

[12] R.D. Campbell, J. Trowsdale, Map of the human MHC, Immunol. Today 14 (1993) 349–352.

[13] P. Argos, An investigation of protein subunit and domain interfaces, Protein Eng. 2 (1988) 101–113.

[14] J. Janin, S. Miller, C. Chothia, Surface, subunit interfaces and interior of oligomeric proteins, J. Mol. Biol. 204 (1988) 155–164.

[15] S. Jones, J.M. Thornton, Principles of protein–protein interactions, Proc. Natl. Acad. Sci. U. S. A. 93 (1996) 13–20.

[16] C.J. Tsai, S.L. Lin, H.J. Wolfson, R. Nussinov, Studies of protein–protein interfaces: a statistical analysis of the hydrophobic effect, Protein Sci. 6 (1997) 53–64.

[17] T.A. Larsen, A.J. Olson, D.S. Goodsell, Morphology of protein–protein interfaces, Structure 6 (1998) 421–427.

[18] L.L. Conte, C. Chothia, J. Janin, The atomic structure of protein–protein recognition sites, J. Mol. Biol. 285 (1999) 2177–2198.

[19] S.J. Davis, E.A. Davies, M.G. Tucknott, E.Y. Jones, P.A. van der Merwe, The role of charged residues mediating low affinity protein–protein recognition at the cell surface by CD2, Proc. Natl. Acad. Sci. U. S. A. 95 (1998) 5490–5494.

[20] D. Xu, S.L. Lin, R. Nussinov, Protein binding versus protein folding: the role of hydrophilic bridges in protein associations, J. Mol. Biol. 265 (1997) 68–84.

[21] R.R. Gabdoulline, R.C. Wade, Simulation of the diffusional association of barnase and barstar, Biophys. J. 72 (1997) 1917–1929.

[22] M. Vijayakumar, K.Y. Wong, G. Schreiber, A.R. Fersht, A. Szabo, H.X. Zhou, Electrostatic enhancement of diffusion-controlled protein–protein association: comparison of theory and experiment on barnase and barstar, J. Mol. Biol. 278 (1998) 1015–1024.

[23] M.A. Olson, L.T. Reinke, Modeling implicit reorganization in continuum descriptions of protein–protein interactions, Proteins 38 (2000) 115–119.

[24] C.J. Camacho, Z. Weng, S. Vajda, C. DeLisi, Free energy landscapes of encounter complexes in protein–protein association, Biophys. J. 76 (1999) 1166–1178.

[25] C.J. Camacho, S.R. Kimura, C. DeLisi, S. Vajda, Kinetics of desolvation-mediated protein–protein binding, Biophys. J. 78 (2000) 1094–1105.

[26] J. Janin, The kinetics of protein–protein recognition, Proteins 28 (1997) 153–161.

[27] S. Albeck, R. Unger, G. Schreiber, Evaluation of direct and cooperative contributions towards the strength of buried hydrogen bonds and salt bridges, J. Mol. Biol. 298 (2000) 503–520.

[28] A.H. Elcock, J.A. McCammon, Calculation of weak protein–protein interactions: the pH dependence of the second virial coefficient, Biophys. J. 80 (2001) 613–625.

[29] S. Vajda, Z. Weng, R. Rosenfeld, C. DeLisi, Effect of conformational flexibility and solvation on receptor-ligand binding free energies, Biochemistry 33 (1994) 13977–13988.

[30] J. Novotny, R.E. Bruccoleri, M. Davis, K.A. Sharp, Empirical free energy calculations: a blind test and further improvements to the method, J. Mol. Biol. 268 (1997) 401–411.

[31] J. Novotny, R.E. Bruccoleri, F.A. Saul, On the attribution of binding energy in antigen–antibody complexes McPC 603, D1.3, and HyHEL-5, Biochemistry 28 (1989) 4735–4749.

[32] I. Muegge, T. Schweins, A. Warshel, Electrostatic contributions to protein–protein binding affinities: application to Rap/Raf interaction, Proteins 30 (1998) 407–423.

[33] M. Schapira, M. Totrov, R. Abagyan, Prediction of the binding energy for small molecules, peptides and proteins, J. Mol. Recognit. 12 (1999) 177–190.

[34] L.P. Lee, B. Tidor, Optimization of binding electrostatics: charge complementarity in the barnase–barstar protein complex, Protein Sci. 10 (2001) 362–377.

[35] I. Massova, P.A. Kollman, Computational alanine scanning to probe protein–protein interactions: a novel approach to evaluate binding free energies, J. Am. Chem. Soc. 121 (1999) 8133–8143.

[36] S.Y. Noskov, C. Lim, Free energy decomposition of protein–protein interactions, Biophys. J. 81 (2001) 737–750.

[37] F. Dong, M. Vijayakumar, H.-X. Zhou, Comparison of calculation and experiment implicates significant electrostatic contributions to the binding stability of barnase and barstar, Biophys. J. 85 (2003) 49–60.

[38] T. Wang, S. Tomic, R.R. Gabdoulline, R.C. Wade, How optimal are the binding energetics of barnase and barstar? Biophys. J. 87 (2004) 1618–1630.

[39] W.D. Cornell, P. Cieplak, C.I. Bayly, I.R. Gould, K.M. Merz Jr., D.M. Ferguson, D.C. Spellmeyer, T. Fox, J.W. Caldwell, P.A. Kollman, J. Am. Chem. Soc. 117 (1995) 5179–5197.

[40] B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swami-nathan, M. Karplus, CHARMM: a program for macromolecular energy minimization and dynamics calculations, J. Comput. Chem. 4 (1983) 187–217.

[41] B. Honig, A. Nicholls, Classical electrostatics in biology and chemistry, Science 268 (1995) 1144–1149.

[42] R.C. Rizzo, W.L. Jorgensen, J. Am. Chem. Soc. 121 (1999) 4827–4836.

[43] K. Palmo, B. Mannfors, N.G. Mirkin, S. Krimm, Potential energy functions: from consistent force fields to spectroscopically determined polarizable force fields, Biopolymers 68 (2003) 383–394.

[44] B.E. Mannfors, N.G. Mirkin, K. Palmo, S. Krimm, Analysis of the pyramidalization of the peptide group nitrogen: implications for molecular mechanics energy functions, J. Phys. Chem., A 107 (2003) 1825–1832.

[45] J.R. Maple, Y. Cao, W. Damm, T. Halgren, G. Kaminski, L. Zhang, R. Freisner, A polarizable force field and continuum solvation methodology for modeling of protein–ligand interactions, J. Chem. Theory Comput. 1 (2005) 694–715.

[46] P. Ren, J.W. Ponder, Polarizable atomic multipole water model for molecular mechanics simulation, J. Phys. Chem., B 107 (2003) 5933–5947.

[47] A. van der Vaart, K.M. Merz Jr., The role of polarization and charge transfer in the solvation of biomolecules, J. Am. Chem. Soc. 121 (1999) 9182–9190.

[48] A.M. Buckle, G. Schreiber, A.R. Fersht, Protein–protein recognition: crystal structural analysis of a barnase–barstar complex at 2.0-A resolution, Biochemistry 33 (1994) 8878–8889.

[49] A. van der Vaart, K.M. Merz Jr., Divide and conquer interaction energy decomposition, J. Phys. Chem., A 103 (1999) 3321–3329.

[50] W. Yang, T.-S. Lee, A density-matrix divide-and-conquer approach for electronic structure calculations of large molecules, J. Chem. Phys. 103 (1995) 5674–5678.

[51] S.L. Dixon, K.M. Merz Jr., Semiempirical molecular orbital calculations with linear system size scaling, J. Chem. Phys. 104 (1996) 6643–6649.

[52] S.L. Dixon, K.M. Merz Jr., Fast, accurate semiempirical molecular orbital calculations for macromolecules, J. Chem. Phys. 107 (1997) 879–893.

[53] G. Schreiber, A.R. Fersht, Interaction of barnase with its polypeptide inhibitor barstar studied by protein engineering, Biochemistry 32 (1993) 5145–5150.

[54] R.W. Hartley, Directed mutagenesis and barnase–barstar recognition, Biochemistry 32 (1993) 5978–5984.

[55] G. Schreiber, A.R. Fersht, Rapid, electrostatically assisted association of proteins, Nat. Struct. Biol. 3 (1996) 427–431.

[56] D. Xu, C.J. Tsai, R. Nussinov, Hydrogen bonds and salt bridges across protein–protein interfaces, Protein Eng. 10 (1997) 999–1012.

[57] L.P. Lee, B. Tidor, Barstar is electrostatically optimized for tight binding to barnase, Nat. Struct. Biol. 8 (2001) 73–76.

[58] H. Umeyama, K. Morokuma, J. Am. Chem. Soc. 99 (1977) 1316–1332.

[59] D.M. York, T.S. Lee, W. Yang, Quantum mechanical treatment of biological macromolecules in solution using linear-scaling electronic structure methods, Phys. Rev. Lett. 80 (1998) 5011–5014.

[60] N. Froloff, A. Windemuth, B. Honig, On the calculation of binding free energies using continuum methods: Application to MHC class I protein–peptide interactions, Protein Sci. 6 (1997) 1293–1301.

[61] G.I. Csonka, J.G. Angyan, The origin of the problem with the PM3 core repulsion function, J. Mol. Struct. 393 (1997) 31–38.

[62] A. van der Vaart, K.M. Merz Jr., Are many-body effects important in protein folding? J. Phys. Chem., B 104 (1999) 9554–9563.

[63] J.A. Gascon, S.S.F. Leung, E.R. Batista, V.S. Batista, A self-consistent space-domain decomposition method for QM/MM computations of protein electrostatic potentials, J. Chem. Theory Comput. 2 (2006) 175–186.

[64] F.B. Sheinerman, B. Honig, On the role of electrostatic interactions in the design of protein–protein interfaces, J. Mol. Biol. 318 (2002) 161–177.

[65] F. Dong, M. Vijayakumar, H.X. Zhou, Comparison of calculation and experiment implicates significant electrostatic contributions to the binding stability of barnase and barstar, Biophys. J. 85 (2003) 49–60.

[66] Y.J. Zheng, K.M. Merz Jr., J. Comput. Chem. 13 (1992) 1151–1169.

[67] J.J. Dannenberg, Hydrogen bonds: a comparison of semiempirical and ab initio treatments, J. Mol. Struct. 401 (1997) 279–286.

[68] A.J. Dingley, S. Grzesiek, Direct observation of hydrogen bonds in nucleic acid base pairs by internucleotide 2JNN couplings, J. Am. Chem. Soc. 120 (1998) 8293–8297.

[69] N.S. Golubev, I.G. Shenderovich, S.N. Smirnov, G.S. Denisov, H.-H. Limbach, Chem. Eur. J. 5 (1999) 492–497.

[70] F. Cordier, S. Grzesiek, Direct observation of hydrogen bonds in proteins by interresidue 3hJNC′ scalar couplings, J. Am. Chem. Soc. 121 (1999) 1601–1602.

[71] G. Cornilescu, J.-S. Hu, A. Bax, Identification of the hydrogen bonding network in a protein by scalar couplings, J. Am. Chem. Soc. 121 (1999) 2949–2950.

[72] E.D. Isaacs, A. Shukla, P.M. Platzman, D.R. Hamann, B. Barbiellini, C.A. Tulk, Phys. Rev. Lett. 82 (1999) 600–603.

[73] A.H. Elcock, D. Sept, J.A. McCammon, Computer simulation of protein–protein interactions, J. Phys. Chem., B 105 (2001) 1504–1518.

[74] G. Schreiber, A.R. Fersht, Energetics of protein–protein interactions: analysis of the barnase–barstar interface by single mutations and double mutant cycles, J. Mol. Biol. 248 (1995) 478–486.

[75] C. Frisch, G. Schreiber, C.M. Johnson, A.R. Fersht, Thermodynamics of the interaction of barnase and barstar: changes in free energy versus changes in enthalpy on mutation, J. Mol. Biol. 267 (1997) 696–706.

[76] V. Gogonea, D. Suares, A. van der Vaart, K.M. Merz Jr., New developments in applying quantum mechanics to proteins, Curr. Opin. Struck. Biol. 11 (2001) 217–223.

[77] K. Raha, A.M. Wollacott, M.J. Italia, J.R. Desjarlais, Prediction of amino acid sequence from structure, Protein Sci. 9 (2000) 1106–1119.

[78] W.L. Jorgensen, J. Chandrasekhar, J. Madura, R.W. Impey, M.L. Klein, J. Chem. Phys. 79 (1983) 926–935.

[79] D.A. Case, D.A. Pearlman, J.W. Caldwell, T.E. Cheatham III, W.S. Ross, C.L. Simmerling, T.A. Darden, K.M. Merz, R.V. Stanton, A.L. Cheng, J.J. Vincent, M. Crowley, D.M. Ferguson, R.J. Radmer, G.L. Seibel, U.C. Singh, P.K. Weiner, P.A. Kollman, AMBER, vol. 5, University of California, San Francisco, 1997.

[80] J. Ryckaert, G. Ciccotti, H. Berendsen, Numerical integration of the Cartesian equations of motion for a system with constraints: molecular dynamics of *n*-alkanes, J. Comput. Phys. 23 (1977) 327–341.

[81] T.A. Darden, D. York, L. Pedersen, Particle Mesh Ewald: an *N* log(*N*) method for Ewald sums in large systems, J. Chem. Phys. 98 (1993) 10089.

[82] M.J.S. Dewar, E.G. Zoebisch, E.F. Healy, J.J.P. Stewart, J. Am. Chem. Soc. 107 (1985) 3902–3909.

[83] J.J.P. Stewart, J. Comput. Chem. 10 (1989) 221–264.

[84] J.J.P. Stewart, J. Comput. Chem. 10 (1989) 209–220.

[85] V. Gogonea, K.M. Merz Jr., Fully quantum mechanical description of proteins in solution. Combining linear scaling quantum mechanical methodologies with the Poisson–Boltzmann equation, J. Phys. Chem., A 103 (1999) 5171–5188.

[86] J. Li, T. Zhu, C.J. Cramer, D.G. Truhlar, New class IV charge model for extracting accurate partial charges from wave functions, J. Phys. Chem., A 102 (1998) 1820–1831.