PROTEINS: Structure, Function, and Genetics 48:117–125 (2002)

The Nature of Intermolecular Interactions Between Aromatic Amino Acid Residues

Francesco Luigi Gervasio, Riccardo Chelli, Piero Procacci,* and Vincenzo Schettino
Dipartimento di Chimica, Università di Firenze, Firenze, Italy and European Laboratory for Nonlinear Spectroscopy (LENS), Firenze, Italy

ABSTRACT **The nature of intermolecular interactions between aromatic amino acid residues has been investigated by a combination of molecular dynamics and ab initio methods. The potential energy surface of various interacting pairs, including tryptophan, phenilalanine, and tyrosine, was scanned for determining all the relevant local minima by a combined molecular dynamics and conjugate gradient methodology with the AMBER force field. For each of these minima, single-point correlated ab initio calculations of the binding energy were performed. The agreement between empirical force field and ab initio binding energies of the minimum energy structures is excellent. Aromatic–aromatic interactions can be rationalized on the basis of electrostatic and van der Waals interactions, whereas charge transfer or polarization phenomena are small for all intermolecular complexes and, particularly, for stacked structures. Proteins 2002;48:117–125.** © 2002 Wiley-Liss, Inc.

Key words: molecular dynamics; ab initio methods; aromatic–aromatic interactions

INTRODUCTION

The interaction between aromatic residues plays a key role in many chemical and biological processes. These interactions, usually termed as π–π interactions,[1] can influence the stereochemistry of organic reactions[2] and the binding affinities in host–guest chemistry.[3] In biological molecular systems, aromatic residues can engage in specific energetically favorable interactions. For example, base stacking determines the sequence-dependent structure and properties of DNA as well as recognition of DNA by drugs and regulatory proteins.[4–6] Similarly, the interactions between planar aromatic residues are important in stabilizing the tertiary structure of proteins.[7–10]

π–π interactions have been extensively investigated both from the experimental and theoretical-computational standpoints (see Ref. 11 and references therein); nonetheless, important questions regarding their nature and directionality still remain open. As to the nature of aromatic amino acid residues interactions, for example, a controversial point is whether and to what extent inductive and charge transfer effects are important in shaping the potential energy surface (PES). This issue is in turn crucial for the modeling of aromatic-aromatic interactions and, more specifically, for the reliability of empirical models based on electrostatic and van der Waals interactions routinely used in standard molecular dynamics (MD) and molecular mechanics (MM) simulations of biological systems. Base stacking and nonclassical hydrogen-ring hydrogen bonds[12,13] have been often cited as signature of these electronic effects.

All the noncovalent interactions between two amino acid residues involve several different effects that can be classified as (i) van der Waals interactions, (ii) electrostatic interactions, and (iii) polarization and charge transfer effects (higher order effects due to the nonstatic nature of the molecular electron clouds). There is a clear experimental and theoretical evidence that van der Waals and electrostatic interactions play an important role in the stability of aromatic-aromatic complexes.[11,14] Furthermore, solvation/desolvation must be taken into account to predict the correct intermolecular geometries in proteins and in solution.[15] On the other hand, there are no convincing experimental indications that induction effects are important.[11] In fact, there is experimental evidence that charge transfer and electron donor-acceptor effects in interacting ground state aromatic molecules are negligible compared to electrostatics.[16]

From the structural point of view, aromatic–aromatic interactions are characterized by a competition between the stacked and T-shaped geometry, the only exception being phenylalanine (toluene) for which only the stacked minimum exists in vacuum.[17] The T-shaped structure, with the planes of the ring perpendicular to each other, is known to be the preferred arrangement of the benzene dimer in the gas-phase,[18] and a T-shaped arrangement is also found in crystalline benzene.[19] The stacked "edge-displaced"[18] (i.e., with the ring planes parallely displaced) and "sandwich" structures (i.e., with superimposed planes) are also local minima for the benzene dimer, stabilized by concurrent dispersive interactions. In aromatic compounds with a polar moiety, the energetics of these structures, characterized by weak and favorable dispersive

Grant sponsor: Italian Ministero dell'Istruzione dell'Università e della Ricerca (MIUR); Grant sponsor: Progetto finalizzato biotecnologie; the Consiglio Nazionale delle Ricerche (CNR); Grant sponsor: European Union; Grant number: HPRI-CT-1999-00111.

*Correspondence to: Piero Procacci, Dipartimento di Chimica, Università di Firenze, Via della Lastruccia 3, 50019 Sesto Fiorentino, Italy. E-mail: procacci@chim.unifi.it

Received 19 October 2001; Accepted 23 October 2001

interactions, is strongly affected by electrostatic interactions and possibly hydrogen bonds.[14,20] On the other hand, the geometry of the aromatic–aromatic complex is also determined by the nature of the interactions with the surrounding media. In this respect, a recent MD study of the tryptophan-histidine (TRP-HIS) pair in various solvents has shown that in water the stacked conformation is favored, whereas carbon tetrachloride favors the T-shaped one.[15] Statistical analysis from the protein data bank of the preferential geometries of interacting pairs of aromatic residues[7,10,21–23] pointed out that parallel-displaced geometries are more common than expected from a purely random orientation. It is, however, still unclear whether the large fraction of stacked conformations observed in proteins is due to a precise directionality of the aromatic-aromatic interaction or to the interaction with the surrounding residues and/or solvent.

Finally, there is some debate on the magnitude of the interaction. Simple electrostatic models predict an interaction in the range of 6 kJ/mol, whereas correlated ab initio and classical MM calculations, in accord with the finding for the DNA base pairs,[24] predict for these aromatic residues a much deeper minimum, around 20–30 kJ/mol.[14,17,25] In this respect, the observation of clustering around particular interplanar angles in proteins[26,27] seems to support the existence of a deep and well-defined energy minimum.

The present study addresses all the issues raised in the preceding discussion and is the natural extension of a series of studies that our group has done on the interactions of aromatic residues. These studies started with the case of TRP-HIS pair in gas phase[14] and in various solvents,[15] continued with the toluene (PHE) homodimer,[17] and has been extended in the present study to all the other relevant aromatic amino acid residues pairs: TRP-PHE, TRP-TYR, PHE-TYR, and TYR-TYR (where TYR stands for tyrosine), excluding only the very uncommon TRP-TRP pair[†] and the pairs involving HIS, which is generally considered as a special aromatic residue.[22] The PES for each pair is determined with a well-established[14,17,20,28–31] mixed classical/ab initio method, in which the PES is scanned by using classical MD simulation and minimum geometries are searched by quenching many randomly selected and uncorrelated phase space points taken from the MD trajectory. The binding energies at the quenched minimum structures are determined by using MP2 calculations corrected by the basis set superposition error (BSSE).[32] Informations on the magnitude and directionality of the interactions in the gas phase can be derived from the PES as determined by MM, whereas the nature of the interactions and the presence and importance of polarization and charge transfer effects are investigated by examining the difference between the electron density of the dimer and that corresponding to the sum of the electron densities of the monomers.

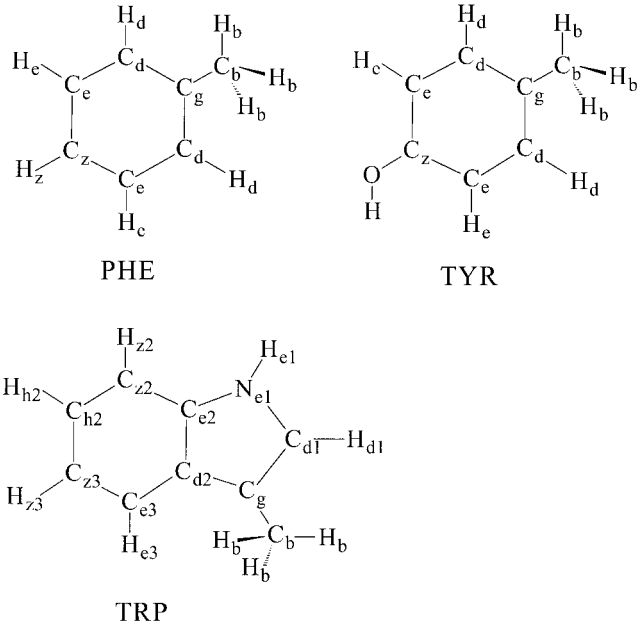

Fig. 1. Models used in the calculations for the aromatic residues: ϵ protonated 3-methyl-indole (TRP), toluene (PHE), and p-cresol (TYR).

The article is organized as follows. In Materials and Methods, we discuss the method used in determining the minima of the PES and the choice of the ab initio method as well as of the basis set used to calculate the binding energies and the differences in electron density. In Results, we discuss the geometries and population of the minima, as well as the ab initio binding energies and the polarization of the residues in the complexes. The final section gives conclusions.

MATERIALS AND METHODS

Molecular Mechanics Potential Model and Determination of Potential Energy Surface

The molecular models used for the aromatic residues are the ε protonated 3-methyl-indole (TRP), toluene (PHE), and p-cresol (TYR) (see Fig. 1). The methyl substitution is introduced to mimic the effect of the steric hindrance of the C_β in proteins.[14,15,26,33]

Following a well-established procedure,[14,17,20,28–31] all possible minima in the intermolecular PES of the TRP-PHE, TRP-TYR, PHE-TYR, and TYR-TYR complexes were determined by quenching 4000 structures regularly sampled from a MD simulation (4 ns for each pair) at 300 K of the two molecules in a cubic box of 25 Å side length with minimum image convention and interacting through the AMBER force field.[34] The conjugate gradient method[35] was used for minimization. All calculations were performed with the program ORAC.[36] The atomic point charges were obtained by a restrained electrostatic potential fit (RESP)[37] of ab initio (MP2/6-31G*(0.25) level of theory) derived electrostatic potentials (Table I). In the case of TYR, two coordinate sets were used for the RESP

[†]The contacts between aromatic residues (centroid–centroid distance < 8.0 Å) in a database of 2396 nonhomologous proteins extracted from the Brookhaven Protein Data Bank were 37756. Of these, only 1199 (3%) were TRP-TRP pairs; for comparison, PHE-TYR contacts were 11288 (30%) and PHE-PHE contacts were 9006 (24%).

TABLE I. RESP Charges and AMBER Atomic Types for TRP, PHE and TYR

Atom	Type	Charge
TRP		
C_b	ct	−0.183594
H_b	hc	0.065385
C_g	c*	−0.026471
C_{d1}	cw	−0.197710
H_{d1}	h4	0.203753
N_{e1}	na	−0.547696
H_{e1}	h	0.418016
C_{e2}	cn	0.266392
C_{z2}	ca	−0.252725
H_{z2}	ha	0.143997
C_{h2}	ca	−0.151528
H_{h2}	ha	0.131647
C_{z3}	ca	−0.164798
H_{z3}	ha	0.131409
C_{e3}	ca	−0.244101
H_{e3}	ha	0.149729
C_{d2}	ca	0.127525
PHE		
C_b	ct	−0.330422
H_b	hc	0.096465
C_g	ca	0.143625
C_d	ca	−0.115342
H_d	ha	0.109987
C_e	ca	−0.210331
H_e	ha	0.139573
C_z	ca	−0.056016
H_z	ha	0.105644
TYR		
C_b	ct	−0.225814
H_b	hc	0.071317
C_g	ca	0.142797
C_d	ca	−0.211325
H_d	ha	0.146975
C_e	ca	−0.241223
H_e	ha	0.152908
C_z	c	0.377992
O	oh	−0.583439
H	ho	0.379843

Atom labels refer to those reported in Figure 1. The intra- and intermolecular potentials are obtained from the atom type (see *type* column) of the AMBER force field.[34] The atomic charges (in fractions of electron) are computed as described in the text.

fit: the first with the hydroxy hydrogen in the ring plane and the second with the hydroxy hydrogen perpendicular to the ring plane. To further validate the AMBER empirical force-field search of low-energy structures, we did an ab initio minimization starting from selected AMBER minima to see how much the resulting dimer structure deviates from the MM counterpart. To this end, we use a density functional theory (DFT)-based method with plane waves. Plane wave DFT calculations on molecular complexes have the great advantage that are not affected by BSSE and, moreover, when using a large cutoff for the wave vector (as we did), check of the basis set convergence is much easier with respect to standard gaussian based methods. The program used for DFT calculations was CPMD,[38] with a B-LYP[39,40] functional, Martins Troullier[41] pseudopotential and a cutoff of 80 Rydberg.

Ab Initio Calculations

For each of the MM minima, ab initio single-point calculations of the binding energy were performed. The binding energy of the complexes was calculated as the sum of the Hartree Fock and the electronic correlation contributions, that is, $E_b = E_b^{(HF)} + E_b^{(COR)}$, where $E_b^{(COR)}$ was calculated at the MP2 level. Both $E_b^{(HF)}$ and $E_b^{(COR)}$ were obtained by subtracting the sum of the energies of the monomers from that of the complex. To eliminate the BSSE, the full function counterpoise method of Boys and Bernardi[32] was applied to both HF and MP2 energy components. The geometry used for the three calculations (two for the isolated monomers with "ghost" partners and that for the complex) was identical to that of the MD quenched structures. To estimate the magnitude of the polarization and charge transfer, the difference between the electron density of the dimer and that of the monomers with the full dimer basis set were calculated, that is, a BSSE corrected differential electron density that measures nonlinear effects due to density reorganization or polarization phenomena. All MP2 calculations were performed by using the program NWChem version 4.0.1.[42]

Basis Set Selection

The selection of the basis set (and of the post Hartree-Fock method) in binding energy computations is a controversial point. Calculations of binding energies with chemical accuracy are very expensive from a computational point of view and can be performed nowadays only for a selected number of minima and medium/small molecular dimers. Accurate binding energies have been computed, for example, for the benzene dimer[18] with a coupled cluster calculation with double and triple excitations [CCSD(T)] and several medium to large basis sets. That set of computations showed that MP2 calculation with a rather small 6-31G*(0.25) basis set reproduces quite accurately the binding energies of the stacked and T-shaped benzene dimer. On the other hand, using the MP2 method with significantly larger basis set in aromatic systems is not recommended[43] because it would lead to an overestimation of the correlation energy, especially for stacked structures where the dispersive contribution is larger. Calculations of the binding energy of DNA stacked and hydrogen bonded dimers** and on dimers of aromatic residues[14,17,25] showed that the 6-31G*(0.25) basis set, possibly because of error cancellations, reproduces (compared to larger basis sets and/or more accurate ab initio methods) the binding energies of stacked aromatic dimers and slightly underestimates the binding energy in the case of hydrogen bonded complexes. Because the focus of this article is on the trends

**Quality of MP2/6-31G*(0.25) hydrogen bonding and stacking energies of DNA base pairs can be deduced from recently published benchmark calculations: hydrogen bonded systems, Ref. 44; stacked systems, Ref. 45.

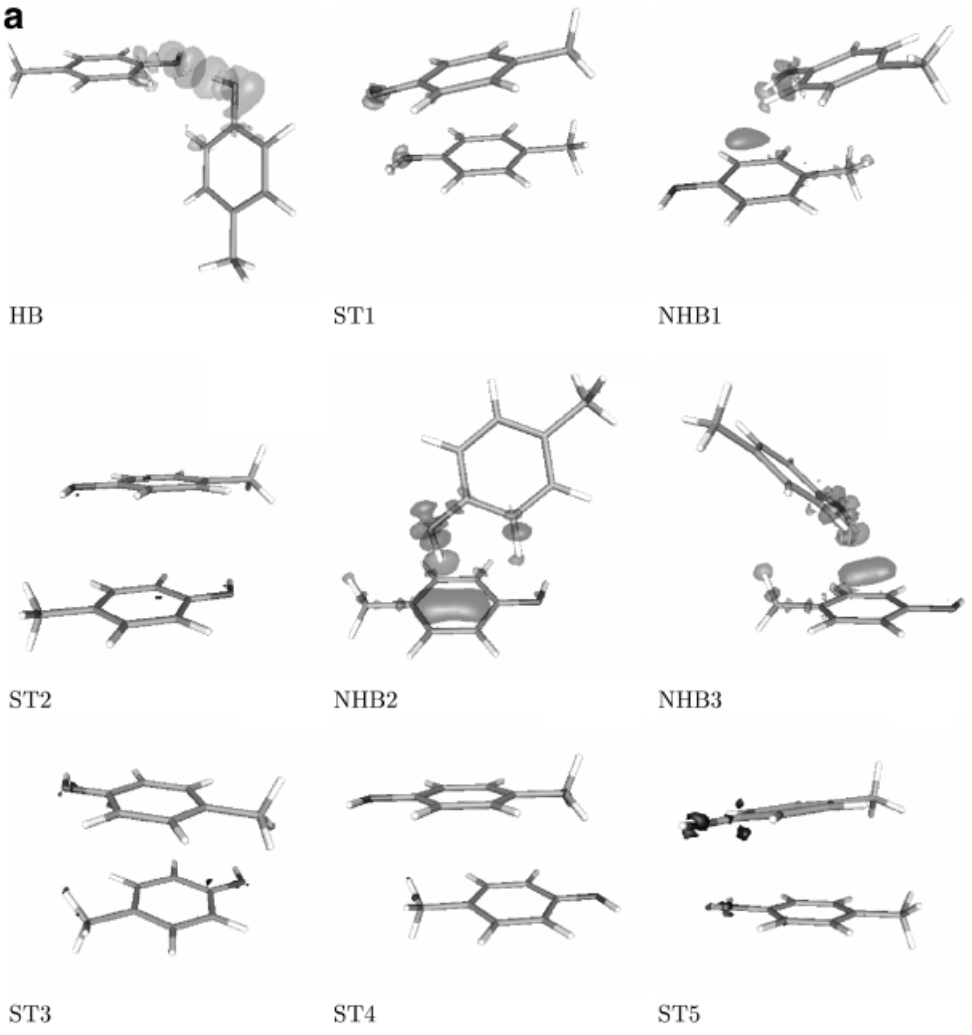

Fig. 2. Geometry of the minima for the aromatic residue pairs. The various structures are reported in the following order: **a:** TYR-TYR; **b:** from top to bottom, PHE-PHE, PHE-TYR, and TRP-PHE; **c:** from top to bottom, TRP-PHE and TRP-TYR. The surfaces refer to the absolute value of the difference between the dimer electron density and that corresponding to the sum of monomer electron densities. The chosen cutoff is 10^{-3} electrons.

of the binding energy rather than on quantitatively accurate calculations of binding energies, the MP2 with 6-31G*(0.25) basis set was used. A rough idea of the error on these calculations can be inferred from the results of Refs. 14 and 18, where the binding energies given by this basis set were compared with those obtained from a larger basis set. In general, the agreement between the two basis sets was found to be of the order of 0.5 kcal/mol or less.

As to the errors in the computation of differential electron density, the error due to the small size of the basis set and low level post Hartree Fock method used should be small: (i) in past studies the polarizability of aromatic compounds calculated using MP2/6-31G*(0.25) was only about 20% less than that calculated with much larger basis set at a CCSD(T) level[43]; (ii) in the present investigation, for the case of selected hydrogen bonded complexes (see Comparison of the Molecular Mechanics Geometries with Ab Initio Minima), the difference of the polarization of the complex compared to that of the two monomers was estimated also with plane wave DFT calculations and was found in fair agreement (<20% difference) with that calculated at the MP2/6-31*(0.25) level.

RESULTS

Molecular Mechanics Geometries

The structures of all principal minima are collected in Figure 2. For each of these structures, ab initio (see next section) and MM calculated binding energies, population and structural data (centroid-centroid distance and interplanar angle) are reported in Table II. The population refers to the fraction (in percentage) of configurations sampled from the MD trajectory that, after conjugate gradient minimization, yield the corresponding minimum and is, therefore, proportional to the size of the associated energy basin. The structures labeled "STn" are stacked, those labeled "HBn" are hydrogen bonded, and "NHBn" are structures with a nonclassical hydrogen bond, where the hydrogen bond acceptor is the

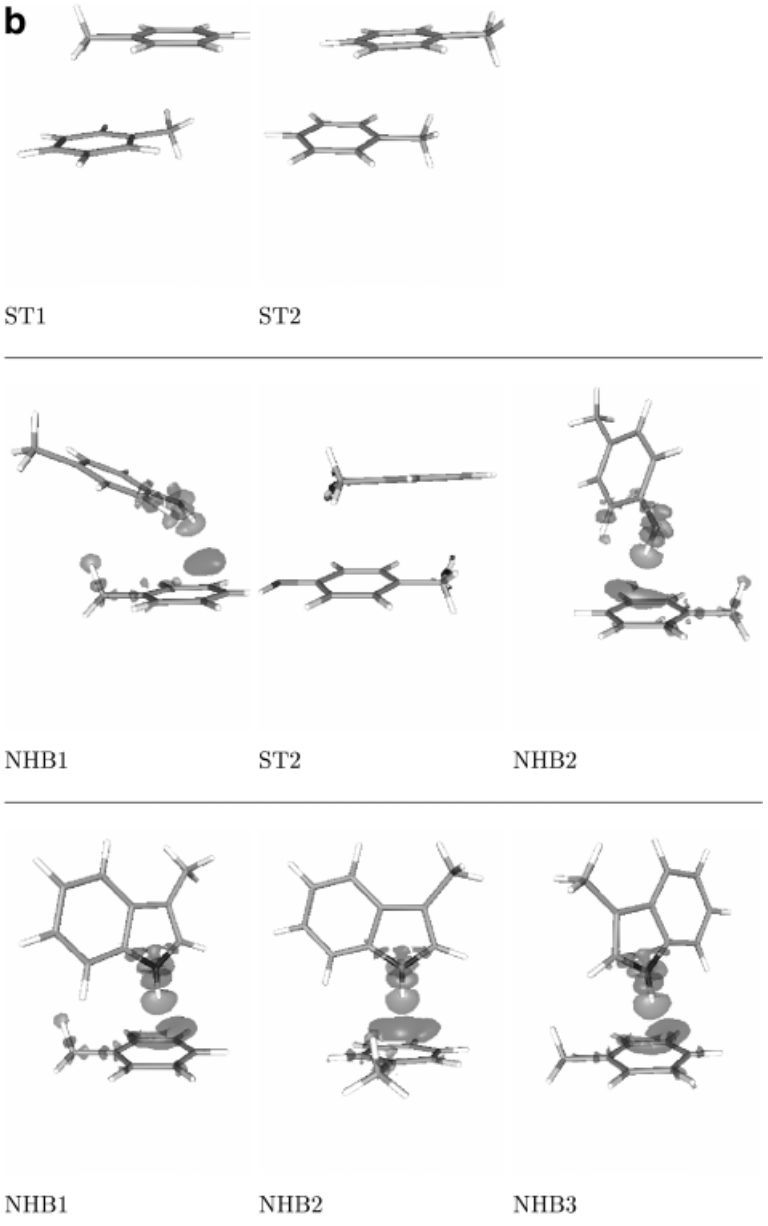

Figure 2. (Continued.)

aromatic ring.[12,13,46] The MM binding energy of the complexes (Table II) is defined as $E_b = E_c - E_{\text{monomer1}} - E_{\text{monomer2}}$, where E_c is the potential energy of the complex, and the energies of the two aromatic residues, E_{monomer1} and E_{monomer2}, are calculated by using the coordinates of the residues in the complex. With this choice for the binding energy definition, the deformation contribution to the energy of the complex is deliberately excluded. All the stacked complexes are stabilized by dispersive Lennard–Jones interactions with a very small, if not positive, electrostatic contribution to the binding energy. On the contrary the HB minima are stabilized mainly by electrostatic interactions. The NHB structures are somewhat in between HB and ST, being stabilized by both electrostatic and dispersive interactions. All the stacked structures, with the exception of the ST1 TYR-TYR minimum, are characterized by parallely displaced geometry. ST1 TYR-TYR has a nearly stacked sandwich-like structure where the stabilization is partly due to a nonlinear OH $\cdots$ O hydrogen bond that forces the stacking of the rings. When conventional hydrogen bonding is possible (TYR-TYR and TYR-TRP), the pure HB structure is the most favored from the entropic point of view (i.e., the most populated). This is most likely due to the local nature and stronger directionality of the hydrogen bond compared to the lack of, in the case of ST, or weak directionality. Hence, strong directionality implies a larger basin of attraction. According to our results, the NHB minima are both deep and populated.

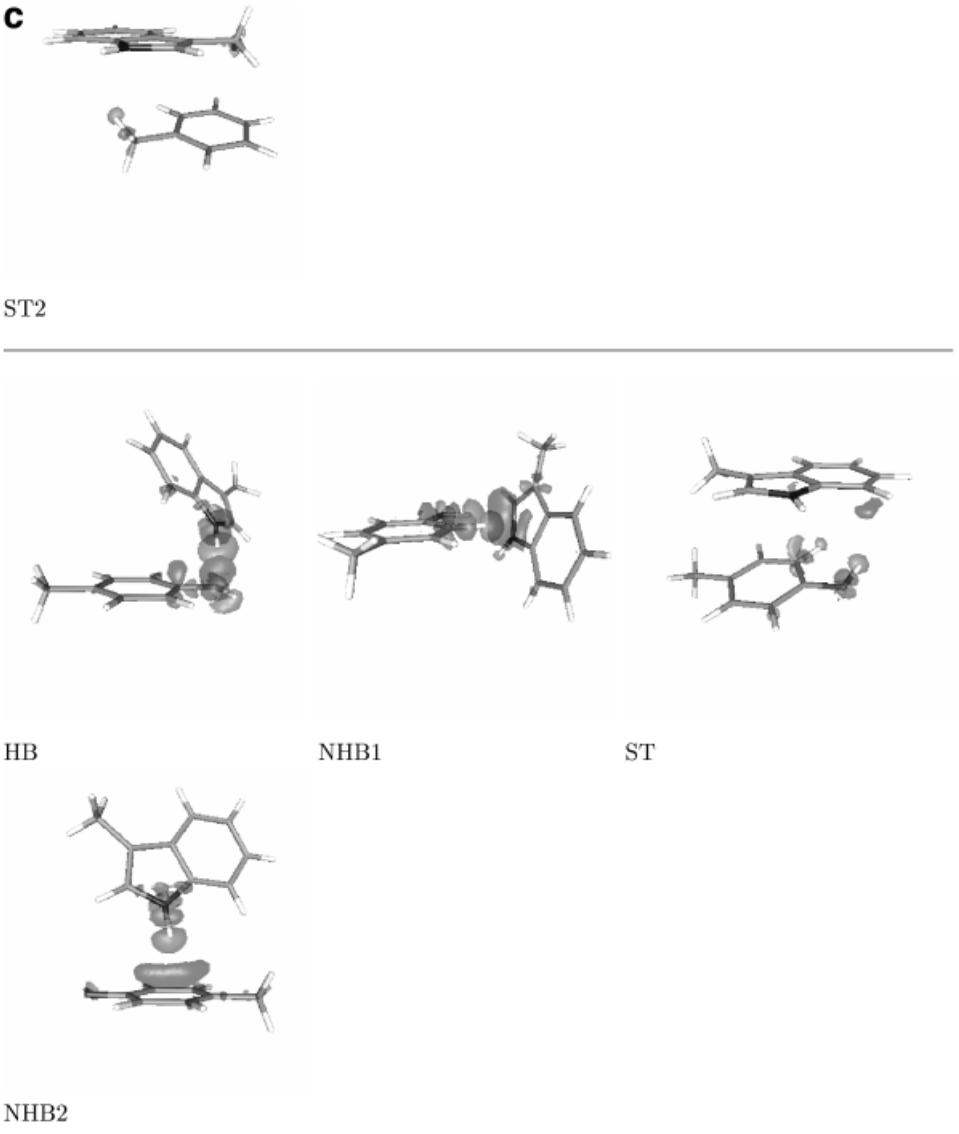

Figure 2. (Continued.)

Comparison of the Molecular Mechanics Geometries With *Ab Initio* Minima

As stated in Basis Set Selection, we performed plane waves DFT geometry optimizations to test the deviation between the ab initio optimized geometry and the MM minima. Given that stacked structures are known to be poorly reproduced by current DFT functionals,[24] because these functionals underestimate the attractive van der Waals forces, we chose the MM structures labeled TRP-PHE NHB1 and TRP-TYR HB (see Fig. 2) that are both stabilized mostly by electrostatics. Moreover, we chose pairs with TRP to estimate the error in the MM geometry for the worst case scenario: in fact, for TRP-TYR and TRP-PHE, MM energies and MP2 ab initio energies (see next section) differ the most. Starting from these MM structures, we did a DIIS[47] minimization and we found the resulting ab initio minima to be very similar to the original MM minima (the root-mean-square displacement between the MM and ab initio structures for TRP-PHE NHB1 and TRP-TYR HB is 0.8 and 0.5 Å, respectively). The main difference between the MM and the ab initio minimum energy structure is found for the H(TRP)-ring(PHE/TYR) distance which, in the ab initio minimized dimer, is shorter than in the corresponding starting MM structure. These findings confirm a widely observed result:[14,17,20,28–31] MM methods with an AMBER-like force field, with RESP atomic charges determined on the monomeric units, is a quite reliable and computationally convenient recipe to explore the PES of aromatic-aromatic interactions. Furthermore, the low cost of the methodology step enables an exhaustive search of the minima, thus avoiding the ambiguities and uncertainties of a partial scanning of the conformational space.[17]

TABLE II. Binding energy (E_b, in kcal/mol) for the structures reported in Figure 2 computed by using ab initio and MM (see text for computational details)

Structure		MP2 E_b	MM E_b	MM E_{el}	Population	r	θ
PHE-PHE	ST1	−4.3	−4.5	−0.2	91	4.5	8.8
	ST2	−3.5	−3.6	0.5	9	4.0	5.1
PHE-TYR	NHB1	−5.9	−5.7	−2.4	64.7	4.6	31.7
	ST	−4.1	−4.6	0.04	31.4	4.3	7.3
	NHB2	−5.1	−5.2	−2.9	3.8	5.2	74.2
TRP-PHE	NHB1	−5.2	−7.1	−3.9	26.0	4.7	80.8
	NHB2	−5.1	−7.0	−4.0	37.7	4.7	77.1
	NHB3	−4.7	−6.9	−4.0	15.6	4.7	82.1
	ST	−5.5	−5.4	−0.1	20.6	4.5	10.8
TYR-TYR	HB	−5.8	−6.6	−5.8	26.1	5.9	69.8
	ST1	−6.3	−6.5	−1.5	15.3	3.7	8.9
	NHB1	−5.9	−5.4	−1.7	13.8	4.5	21.4
	ST2	−5.4	−5.5	−0.6	10.3	3.8	0.02
	NHB2	−5.9	−5.5	−2.7	9.6	4.9	53.4
	NHB3	−5.9	−5.4	−2.6	6.4	4.9	44.0
	ST3	−4.3	−4.9	0.8	5.6	3.6	0.3
	ST4	−4.6	−4.9	0.6	4.8	3.8	2.8
	ST5	−5.6	−5.5	−0.2	3.0	3.7	8.5
TRP-TYR	HB	−5.2	−7.6	−5.8	23.8	5.2	57.2
	NHB1	−6.2	−7.0	−4.5	12.4	5.0	74.0
	ST	−7.8	−7.1	−1.3	11.7	3.9	13.2
	NHB2	−5.7	−6.9	−3.5	18.4	4.7	76.7

The electrostatic contribution to the MM E_b energy is also reported (MM E_{el}). In the last three columns, the percentage population, the centroid–centroid distance (R, in Å) and the interplanar angle (θ, in degrees) are reported.

Binding Energies and Electron Density Differences

The agreement between ab initio and MM binding energies with a mean difference of only 0.5 kcal/mol (see Table II) is excellent and within the error of the ab initio calculations.[6,43] Remarkably, the agreement of the binding energies is extremely good for the case of ST structures, it is quite good in the case of NHB structures (with the noticeable exception of TRP case), and is satisfactory in the case of HB structures. We do believe that this agreement, indeed beyond expectations, between MM and ab initio binding energies is not coincidental but has instead a quite convincing physical explanation. The AMBER force field represents the intermolecular potential by a combination of RESP charges (i.e., atomic charges that reproduce the ab initio electrostatic potential given by the ground state electron density and the nuclear charges of the monomeric units) and of atom–atom Lennard–Jones centers. The underlying assumption of the AMBER approach is that, for interacting molecules, nonlinear effects in the dispersive and electrostatic contributions due to charge reorganization or charge transfer phenomena are negligible. This assumption can be checked by evaluating the differential electron density of the dimer and the sum of the two electron density of the two monomers. The absolute value of the differential electron density is represented as three-dimensional isodensity surfaces (the chosen value was 10^{-3} electrons) in Fig. 2. Because the value for the isodensity surface is the same for all the complexes of Fig. 2, they can be directly compared such that larger surfaces correspond to larger polarization and possibly charge transfer. The differential electron density revealed that, in the case of stacked structures, there is no evidence whatsoever of important polarization or charge transfer effects. No wonder that the best agreement between MM and MP2 binding energies occurs for the ST structures. As expected, "nonclassical" effects are more important for HB and NHB complexes. However, also in these cases, the isodensity surfaces do not indicate important charge transfer or induction phenomena because they refer to only 10^{-3} electrons. This finding is also confirmed by analyzing the amount of electron involved in induction effects, calculated by integrating the absolute value of the differential electron density on the whole space. Results are reported in Table III. The extent of electron density reorganization is indeed quite small also for hydrogen bond-stabilized structures. A similar conclusion was drawn also for the TRP-HIS pair.[14] Experimental evidences[16] in the case of substituted benzenes confirm the hypothesis of weak polarization effects in the aromatic-aromatic interactions.

If the total electrostatic energy in the aromatic-aromatic dimers can be safely calculated by simply considering the classical electrostatic interactions of the two static charge distribution of monomers, we conclude that the only source of error in the MM parameterization must come from the Lennard–Jones part, as the RESP charges are designed so as to reproduce the correct [i.e., MP2/6-31Gd(0.25)] electrostatic energy of the dimer. For improving the agreement between the MM and ab initio binding energies, it would then be possible to adjust only the Lennard–Jones part of the AMBER parameterization while leaving unchanged the RESP charges. In Table II, for example, the MM and ab initio binding energies of the TRP-PHE and TRP-TYR

TABLE III. Integrated differential electron densities (in 10^{-2} electrons) for the structures reported in Figure 2

Structure		Charge reorganization
PHE-PHE	ST1	0.01
	ST2	0.01
PHE-TYR	NHB1	3.93
	ST	0.16
	NHB2	3.99
TRP-PHE	NHB1	4.94
	NHB2	5.14
	NHB3	4.91
	ST	0.46
TYR-TYR	HB	9.93
	ST1	0.72
	NHB1	2.01
	ST2	0.08
	NHB2	4.46
	NHB3	4.54
	ST3	0.17
	ST4	0.08
	ST5	0.33
TRP-TYR	HB	7.63
	NHB1	7.86
	ST	1.45
	NHB2	4.65

The computation was made as follows: for the dimer and the two monomers the electron density was computed on a grid of $41 \times 41 \times 41$ points. The differential electronic density, $\Delta\rho$, (dimer − sum of the monomers) was calculated for each point on the grid. Finally, the integral of $|\Delta\rho|$ inside all the isosurfaces with a value of 10^{-3} electrons was calculated.

pairs differ in average by 1.4 kcal/mol, a rather large deviation with respect to the mean 0.3 kcal/mol of all other pairs. In plane wave DFT minimization for the TRP-PHE and TRP-TYR pairs (see Comparison of the Molecular Mechanics Geometries With Ab Initio Minima), we noticed that the H(TRP)-ring(PHE/TYR) distance in the ab initio minimized dimer was shorter than in the corresponding starting MM structure. This indicates that van der Waals potential parameter σ in the standard AMBER potential for the H_{11} hydrogen of TRP could be too small. Therefore, we reoptimized the AMBER geometry of the dimer for different values of this σ parameter. For each of the new geometries corresponding to increased σ values, we calculated the single point ab initio BSSE corrected MP2/6-31Gd(0.25) binding energy. The MM and ab initio binding energies are found to be coincident when $\sigma = 1.1$ Å, a value between that of polar hydrogen of HIS and that of the H3 hydrogen bound to an aliphatic carbon with three electron withdrawing groups. This result is consistent with the lower acidity of the polar hydrogen of tryptophan with respect to that of histidine.

CONCLUSIONS

Interaction between aromatic residues plays an important role in determining the structure and dynamics of protein folding. In the case of stacked residues, the extensive overlap of their aromatic rings gave rise in the past to speculations about the physical origin of stacking and the possible role of delocalized electrons in stacked aromatic π systems, leading to some confusion about the issue.

The application of quantum-mechanical approaches with inclusion of electron correlation effects in the case of TRP-HIS pair[14] and, in the present study, for all other relevant aromatic amino acid pairs, provide a better qualitative and semiquantitative description of the phenomenon of π–π complexation. In the case of stacked complexes, we do not find any evidence of important charge transfer or polarization contribution to the overall binding energy. The binding energy in stacked complexes is dominated by repulsion-dispersive interactions that can be effectively parameterized with conventional atom–atom Lennard–Jones centers. For hydrogen bond and T-shaped complexes, the extent of charge transfer and induction contribution to the binding energy is larger with respect to the stacked systems but still small compared to electrostatic interactions. The binding energy in T-shaped complexes is dominated by electrostatic contributions that can be effectively parameterized by using standard atomic charges.

For the DNA bases interactions[44,45] and for aromatic residues, we found that hydrogen bonded, nonclassical hydrogen bonded and stacked aromatic amino acid pairs can all be modeled reliably by the inclusion of three effects: dispersive attraction, electrostatic interaction, and short-range repulsion. All these effects are accounted for in a satisfactorily way by the AMBER force field for proteins,[34] as shown by the excellent agreement between MM/AMBER and single-point MP2/6-31Gd(0.25) BSSE corrected calculations for the binding energies of the complexes. For complexes involving TRP, the agreement between MM and ab initio is poorer. We found that a modest reparameterization of the MM repulsive Lennard–Jones contribution for the polar hydrogen of tryptophan is sufficient to obtain a better energetic (and geometric) agreement with the ab initio data.

ACKNOWLEDGMENTS

F.L.G. is grateful to Menarini S.p.A. for partial financial support of his Ph.D. studentship. NWChem Version 4.0.1, was developed and distributed by Pacific Northwest National Laboratory, P.O. Box 999, Richland, WA 99352 USA and was funded by the U.S. Department of Energy and was used to obtain some of these results.

REFERENCES

1. Hunter CA, Sanders JKM. J Am Chem Soc 1990;112:5525–5534.
2. Evans DA, Chapman KT, Hung, KT, Kawaguchi AT. Angew Chem Int Ed Engl 1987;26:1184–1185.
3. Muehldorf AV, Engen DV, Warner, JC, Hamilton AD. J Am Chem Soc 1988;110:6561–6562.
4. Hunter CA. J Mol Biol 1993;230:1025–1054.
5. Saenger W. Principles of nucleic acid structure. New-York: Springer-Verlag; 1984.
6. Hobza P, Selzle HL, Schlag EW. Chem Rev 1994;94:1767–1785.
7. Burley SK, Petsko GA. Science 1985;229:23–28.
8. Singh J, Thornton JM. FEBS Lett 1985;191:1–6.
9. Blundell T, Singh J, Thornton J, Burley SK, Petsko GA. Science 1986;234:1005–1005.
10. Mitchell JBO, Nandi CL, McDonald IK, Thornton JM, Price SL. J Mol Biol 1994;239:315–331.

11. Hunter CA. Chem Soc Rev 1994;23:101–109.
12. Burley SK, Petsko GA. FEBS Lett 1986;203:139–143.
13. Levitt M, Perutz MF. J Mol Biol 1988;201:751–754.
14. Gervasio FL, Procacci P, Cardini G, Guarna A, Giolitti A, Schettino V. J Phys Chem B 2000;104:1108–1114.
15. Gervasio FL, Chelli R, Marchi M, Procacci P, Schettino V. J Phys Chem B 2001;105:7835–7846.
16. Cozzi F, Cinquini M, Annunziata R, Siegel JS. J Am Chem Soc 1993;115:5330–5331.
17. Gervasio FL, Chelli R, Procacci P, Schettino V. J Phys Chem A 2002;106:2945–2948.
18. Hobza P, Selzle HL, Schlag EW. J Phys Chem 1996;100:18790–18794.
19. Hall D, Williams DE. Acta Crystallogr 1975;A31:56–58.
20. Kratochvil M, Engkvist O, Sponer J, Jungwirth P, Hobza P. J Phys Chem A 1998;102:6921–6926.
21. Karlin S, Zuker M, Brocchieri L. J Mol Biol 1994;239:227–248.
22. Brocchieri L, Karlin S. Proc Natl Acad Sci USA 1994;91:9297–9301.
23. McGaughey GB, Gagné M, Rappé AK. J Biol Chem 1998;273:15458–15463.
24. Hobza P, Sponer J. Chem Rev 1999;99:3247–3276.
25. Alagona G, Ghio C, Monti S. J Phys Chem A 1998;102:6152–6160.
26. Chelli R, Gervasio FL, Procacci P, Schettino V. J Am Chem Soc 2002. In press.
27. Mitchell JBO, Laskowski RA, Thornton JM. Proteins 1997;29:370–380.
28. Kratochvil M, Sponer J, Hobza P. J Am Chem Soc 2000;122:3495–3499.
29. Ryjacek F, Kratochvil M, Hobza P. Chem Phys Lett 1999;313:393–398.
30. Kratochvil M, Engkvist O, Vacek J, Jungwirth P, Hobza P. Phys Chem Chem Phys 2000;2:2419–2424.
31. Ryjacek F, Engkvist O, Vacek J, Kratochvil M, Hobza P. J Phys Chem A 2001;105:1197–1202.
32. Boys SF, Bernardi F. Mol Phys 1970;19:553–556.
33. Chipot C, Jaffe R, Maigret B, Pearlman DA, Kollman PA. J Am Chem Soc 1996; 118:11217–11224.
34. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA. J Am Chem Soc 1995;117:5179–5197.
35. Press WH, Flannery BP, Teukolsky SA, Vetterling WT. Numerical recipes. Cambridge, MA: Cambridge University Press; 1986.
36. Procacci P, Darden TA, Paci E, Marchi M. J Comput Chem 1997;18:1848–1862.
37. Bayly CI, Cieplak P, Cornell WD, Kollman PA. J Phys Chem 1993;97:10269–10280.
38. Hutter J, Ballone P, Bernasconi M, Focher P, Fois E, Goedecker S, Parrinello M, Tuckerman M. CPMD 4.1. Stuttgart: MPI fur Festkrperforschung and IBM Zurich Research Laboratory; 2000.
39. Becke AD. Phys Rev A 1988;38:3098–3100.
40. Lee C, Yang W, Parr RG. Phys Rev B 1988;37:785–789.
41. Trouillier N, Martins JL. Phys Rev B 1991;43:1993–2006
42. Group HPCC. NWChem, a computational chemistry package for parallel computers, version 4.0.1. Richland, WA: Pacific Northwest National Laboratory; 2001.
43. Hobza P. Sponer J, Leszczynski J. J Phys Chem B 1997;101:8038–8039.
44. Sponer J, Hobza P. J Phys Chem A 2000;104:4592–4597.
45. Sponer J, Berger I, Spackov N, Leszczynski J, Hobza P. J Biomol Struct Dyn 2000;Conversation 11:383.
46. Alkorta I, Rozas I, Elguero J. Chem Soc Rev 1998;27:163–170.
47. Pulay P. Chem Phys Lett 1980;73:393–398.